Genomics of drought tolerance in flax (*Linum usitatissimum* L.)

by

Demissew Sertse Desta

Thesis submitted to the University of Ottawa in partial fulfilment of the requirements of the degree

Doctor of Philosophy

Department of Biology

Faculty of Science

University of Ottawa

Abstract

Flax (*Linum usitatissimum* L.), one of the eight founder crops of agriculture, is a versatile plant that has been a source of fiber and oil for millennia. Flax production, however, is constrained by various factors, including moisture stress. Yet, variations in the level of tolerance to different moisture regimes exists among flax germlines, thereby providing opportunities for breeders.

To understand the global scale eco-geographic adaptation and to further gain insight into the genetic architecture of drought-related traits leading towards marker-assisted breeding and/or genomic selection in flax, a genome-wide investigation using high density single nucleotide polymorphism (SNP) data was performed in two sets of flax genotypes: a core (n=407) and a mini-core (n=115) collection, both representing the major world flax growing regions. The core collection was used to demonstrate haplotype distribution and possible driving forces that have shaped the population structure of the crop. The mini-core was used to perform genotype-phenotype association for drought-related traits through genome-wide association studies (GWAS) based on phenotypic values generated in two experiments. In the first experiment, 16 early root and shoot traits were evaluated using a semi-hydroponic pouch system in a controlled environment, while in the second experiment 11 drought-related traits were evaluated under irrigated and non-irrigated conditions. For the latter, GWAS was performed using the adjusted values of six of the measured traits that significantly responded to the watering regime, as well as six calculated stress indices for each trait.

The genetic structure analysis clustered the core collection into four major groups: Temperate, South Asian, Mediterranean and Abyssinian. The Temperate group comprised the majority of the core collection and was dominated by accessions from Eurasia, the Americas and Oceania. Despite their limited representation, the other three groups harbored a high concentration of endemic haplotypes that can be attributed to the long history of flax cultivation

in these regions hypothesized to be the centers of diversification the crop. More in-depth samplings from these regions is thus expected to broaden the genetic diversity of the flax core collection substantially. Genetic variation within morphotypes is more pronounced than between the fiber and oil types, reflecting selection of genotypes for dually elite traits i.e. both for seed and fiber. The overall result suggests the contribution of use-directed human selection and eco-geographic variations along the latitudinal gradient as major forces in shaping the genetic structure of the crop.

The two GWAS experiments, i.e., early root-shoot growth and field-drought-tolerance-related traits, yielded 228 and 144 quantitative trait nucleotides (QTNs), respectively, that were associated with at least one trait. Most loci, 100 kb up- and downstream of the associated QTN, harbored genes predicted to play roles in modulating the associated trait or traits. In the root-shoot GWAS, genes predicted to encode GRAS transcription factors, mitogen-activated protein kinases (MAPKs), and auxin related lateral organ boundary proteins were present at root trait QTN loci, while QTN loci associated with shoot traits mainly harbored genes involved in photo-morphogenesis and plant immunity. QTN loci of drought tolerant traits consistently harbored genes known to mediate responses to abiotic stress, including drought, heat and salt, suggesting the importance of these loci for stress tolerance to be considered as candidate QTL of the traits associated with.

Keywords: Flax, drought tolerance, eco-geographic adaptation, root phenotyping, GWAS

Acknowledgements

I sincerely wish to thank the following persons, who have been instrumental in the success of my Ph.D. project. Dr. Sylvie Cloutier (my supervisor) for having allowed me to pursue a research project in her lab and for having guided me throughout this journey. Prof. Xia Xuhua, my co-supervisor for his encouragement, and his help, especially with the admission to the graduate program at the University of Ottawa. My advisory committee members Drs. Owen Rowland and Nicolas Rodrigue, at Carleton University, for their constructive comments and advice, from my first proposal to my final thesis.

The lab colleagues in Dr. Sylvie Cloutier's lab. Tara Edwards, Madeleine Lévesque-Lemay, Dr. Sridhar Ravichandran for their support throughout my studies, especially with field work and plant phenotyping. Dr. Ravichandran has provided invaluable help and encouragement with many facets of my time in the lab. Shanise Reddekopp, Lindsey Cameron, Gabriel Sullivan and other students have helped me in the field and growth cabinet/greenhouse experiments.

Dr. Scott Duguid and Mary Meiklejohn, for their collaboration with the field experiment at Morden, Manitoba, Canada. Dr. Malcom Morrison for his support and guidance during my field phenotyping experiment and to his team for their assistance with the installation of the irrigation system every Spring. Ron Stanley and the farm management team for their cooperation during my field experiment. Alain Saumure for his help in designing and construction of the root phenotyping system.Dr. Frank M. You for his guidance and valuable suggestions during the data analysis.

Drs. Blen Beyene and Wubshet A. Bekele (Wubi) what so ever, you had to do; including Wubi's technical advices all throughout! Yet, I have to say only thank you! Sal and Dr. Zerihun A. Demissie (Zerish) for all values and events we shared with but the date I said "Wegen Alegn"

Acknowledgements

will remain in my memory printed! Thank you! I also have to mention the group *Altebelashenim* for your inspiring events.

My Parents Sertse Desta (ሠርፀ ደስታ) and Kebebush Belayneh (ከበቡሽ በላይነህ), and my siblings Eniye, Solomon, Yared for constant support and prayers. Yaredo, thank you for visiting me twice, crossing the vast Atlantic and realms (Ethiopia-Canada) and thank you for planning to attend my graduation.

My funding sources: Agriculture and Agri-food Canada (AAFC) for employing me through as a Research Affiliate Program student; the University of Ottawa for offering me teaching assistance jobs during all Fall and Winter terms; and the International Doctoral Scholarship Tuition Fee Support since September 2018. Expenses to run the experiments were covered by Agri-Innovation Project (J-000279) funded by the Flax Council of Canada, and once again AAFC.

Finally, I wish to thank God for His Grace throughout my Ph.D.

ቅዱስ አባት እግዚአብሔር ሆይ ላደረክልኝና ለምታደርግልኝ ሁሉ አመሰግንሀለሁ!

Table of contents

Table of contents

Table of contents

Table of contents

List of Tables

List of Figures

Acronyms

AAC	ADP, ATP carrier protein
ABYS	Abyssinian
ADP	Adenosine diphosphate
AF	Activation factor (for NAC-domain)
AGL	Agamous-like
AGSNP	Annotation-based genome-wide SNP
ALA	Alpha linolenic acid
ANOVA	Analysis of variance
ARD	Average root diameter
ARF	Auxin response factor
ARM	Armadillo repeat
ATP	Adenosine triphosphate
BAM	Binary alignment map
BDLWt	Bundle weight
BSK	Brassinosteroid-signaling kinase
BWt	Boll weight
CA_RU	Canadian_Russian
CANC	Canadian cultivars
CDC	Crop Development Center (Canada)
CEPR	C-terminally encoded peptide receptor
Chr	Chromosome
CI	Confidence interval
CIPK	Calcineurin B-like interacting protein kinase
COP	Constitutively photomorphogenic

Acronyms

CSI	Cellulose synthase interactive
CT	Canopy temperature
CYP	Cytochrome P
DNA	Deoxyribonucleic acid
DSI	Drought susceptibility index
DTE	Drought tolerance efficiency
EGR	Clade E growth-regulating
EMB	Embryo
EOD	Enhancer of DA; DA= Big in Chinese
FAO	Food and Agricultural Organization of the United Nations
FAOSTAT	Food and Agriculture Organization of the United Nations Statistics
FDR	False discovery rate
FIB	Fiber
FLD	Flowering date
GDP	Gross domestic product
GEA	Genome-environment association
GRAS	Gibberellic acid insensitive (GAI), repressor of GA1 (RGA) and Scarecrow (SCR) transcription factors
GWAS	Genome wide association study
H^2	Heritability
HS	Heat shock
HT	Plant height
IMPA1	Importin subunit alpha-1
ITN	Increased tolerance to NaCl (sodium chloride)
KCS	Ketoacyl-CoA synthase

KEA	Potassium ion efflux antiporter
KFB	A kelch domain-containing F-box protein
LD	Linkage disequilibrium
LDB	Linkage disequilibrium block
LFMM	Latent factor mixed model
LOB	Lateral organ boundary
LRR	Leucine-rich repeat (receptor)
MAD	Modified augmented design
MADS (four genes)	M = Minchromosome maintainance factor (MCM);
	A = Agamous (AG);
	D = Deficient (DEF);
	S = Serum response factor (SRF)
MAF	Minor allele frequency
MAPK	Mitogen-activated protein kinase
MATE	Multi and toxic compound extrusion
MaxR	Maximum number of roots
Mb	Mega base
MedR	Median number of roots
MEDT	Mediterranean
MIND	Mini_Indian
MLM	Mixed linear model
MN	Manitoba
MTD	Maturity date
NAC (three genes)	N = NAM (no apical meristem);
	A = ATAF (*Arabidopsis thaliana* activation factor);

	C = CUP2 (cup-shaped cotyledon)
NAP	Nucleosome assembly protein
NF_YA	Nuclear transcription factors Y subunit A
NJ	Neighbor joining (phylogenetic tree)
NWA	Network area (roots)
NWAD	New World admixture
NWDep	Network depth (root depth)
NWDist	Network distribution (root)
NWL	Network length (root)
NWPeri	Network perimeter (root)
NWSA	Network surface area (root)
NWV	Network volume (root)
NWW	Network width (root)
NWW_Dep	Network width to network depth ration (root)
ON	Ontario
OWAD	Old World admixture
PC	Principal component
PCA	Principal component analysis
PGRC	Plant Gene Resource Canada
PILS	PIN likes (PIN=Pin-formed; PIN is not an acronym)
PPR	Pentratriconpeptide repeat
PVE	Phenotypic variance explained
QTL	Quantitative trait locus/loci
QTN	Quantitative trait nucleotide
RDWt	Root dry weight

REF	Relative of early flowering
RPPR	Ribosomal pentatricopeptide repeat protein
RR	Response regulators
RTM-GWAS	Restricted two-stage multi-locus multi-allele GWAS
SA	South Asian
SAIN	South Asian-India
SAP	Stress-associated protein
SAPK	South Asian-Pakistan
SCR	Scarecrow (gene)
SDF	Stromal cell-derived factors
SDWt	Shoot dry weight
SHN	Shine
SHR	SHORTROOT
SL	Shoot length
SNP	Single nucleotide polymorphism
SPA	Suppressor phytochrome A
SPB	Seeds per boll
SRL	Specific root length
SSI	Stress susceptibility index
SSSI	Schneider’s stress severity index
STD	Plant stand
STI	Stress tolerance index
SWtPB	Seed weight per boll
TBL	Trichome birefringence-like
TEMP	Temperate

TIC	Translocon at the inner envelope membrane of chloroplasts
TOL	Tolerance against stress
TPR	Tetratricopeptide repeat
TPR	Tetratricopeptide repeat
TSW	Thousand seed weight
UN	United Nations
USD	US dollar
UV	Ultra violet
WAKL	Wall-associated receptor kinase-like
WD40R	Tryptophan-aspartic acid (W-D) dipeptide terminating ~40 amino acids tandem repeats
WEUR	Western Europe
WIN	Wax induce
WRKY	Amino acids: W = tryptophan, R = arginine, K = lysine and Y = tyrosine
XTH	Xyloglucan endotransglucosylase/hydrolase
YLD	Yield (grain)
YUC	YUCCA (named after the genus *Yucca*, in the family *Asparagaceae*)

Note: acronyms of proteins and/or genes include only those in the main body of the text and not those of the appendices.

1. Introduction

1.1. Agriculture and drought

According the latest United Nations Department of Economic and Social Affairs report, the world population has been projected to exceed 9.5 billion by 2050 (UN 2019). To meet the demand for food and other agricultural products of the projected 2050 global population, agriculture is required to increase its production by 70% (Hunter et al. 2017). At the current production rate, 100 million hectares of additional land would be needed to achieve the anticipated demand (Pastor et al. 2019). Although this is globally achievable, most of the currently available arable lands are concentrated in a few less densely populated regions while land is already scarce in regions with dense populations (FAO 2009). The 2050 production plan faces other critical challenges with climate change being at the top of the list (Asseng et al. 2015).

Drought is one of the foremost climate-related factors confronting agriculture and it can have serious global economic repercussions. Drought does not have a single clear and comprehensive definition, although all agree that its definition includes deficit of water as a core element (Dracup et al. 1980; Lloyd-Hughes 2013). The phenomenon is technically classified into four categories: meteorological (deficit in precipitation), hydrological (deficit of water in stream flow, aquifers and other natural reservoirs such as lakes), socio-economic (deficit in water supply) and agricultural (soil moisture deficit) (Wilhite and Glantz 1985). All combined, drought is estimated to cause annual losses of up to 0.25% of the world gross domestic product (GDP) (Jenkins and Warren 2015) which amounts to ~ $194.7 billion USD based on the 2014 world GDP (World-Bank 2015).

Drought impacts are more severe in agriculture than in other sectors because agriculture accounts for more than 80% of the overall impact (Baas et al. 2015) and, as such, drought has become a critical challenge of the sector (Hao et al. 2014; Zhao and Dai 2015). In the past 50

years, drought has been the leading natural disaster challenging agriculture. Causing havoc to the entire farming system through severe losses in both livestock and crop production, droughts significantly affect human's livelihood around the globe (Baas et al. 2015). The effect of drought is more severe with high temperature (Lesk et al. 2016). The recent drought-temperature effects were estimated to result in approximately 10% global yield losses for major crops such as maize, wheat and soybeans (Matiu et al. 2017).

Drylands are defined by a scarcity of water. It is not simply a consequence of precipitation but a balance between precipitation and evaporation from the soil surface and in the form of plant evapotranspiration. Dry areas account for approximately 41% of the earth's surface with

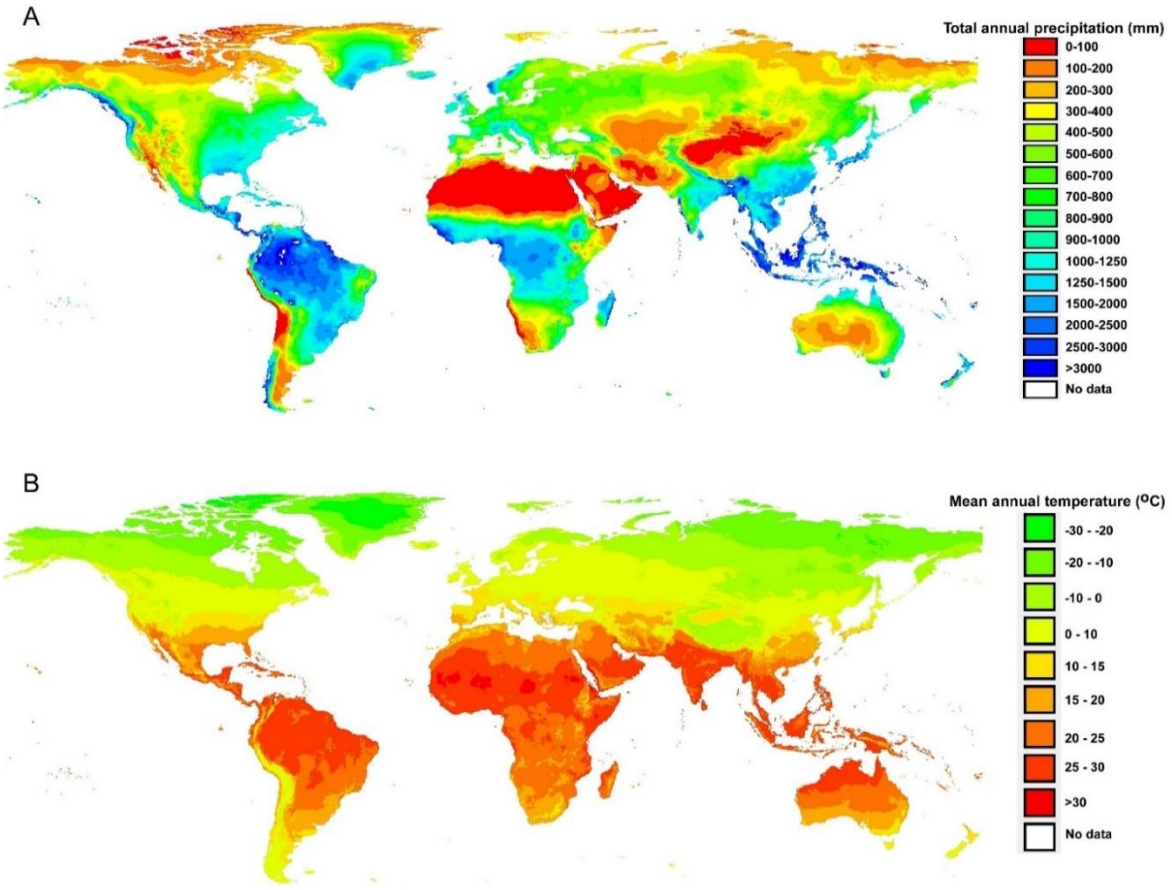


Figure 1.1 *Annual precipitation (**A**) and mean annual temperature (**B**) of the world terrestrial surface.*

The figure was created using the GIS software based on Worldclim version 2 data downloaded from http://worldclim.org

over 2 billion inhabitants (Solh and van Ginkel 2014) of which severe to extreme drought-prone areas cover 18% (Kogan et al. 2013). The drylands are expanding and are expected to increase to 50% of the land surface by the end of this century (Zhao and Dai 2015).

The severity of drought spells is also predicted to increase (Prudhomme et al. 2014). Although the major low-precipitation regions are high-latitude cooler zones (Figure 1.1A&B), the global temperature increases due to climate change have become alarming because they affect food production (Asseng et al. 2015). Against these critical alerts, agriculture needs to find appropriate mechanisms to meet the anticipated 2050 production (Hunter et al. 2017; Tilman et al. 2011). To cope with this challenge and feed the growing world population, breeding for drought-tolerant crop varieties has been suggested as one of the strategies (Dar and Laxmipathi Gowda 2013; Tester and Langridge 2010). Plant breeding for drought tolerance has been accelerated by the current advances in molecular biology that enable targeting of genes associated with drought-tolerance related traits (Foley et al. 2011; Tester and Langridge 2010).

1.2. Coping strategies of plants against drought

In a broad sense, plants' strategies to cope under drought conditions can be categorized into drought escape, drought avoidance, and drought tolerance (Delzon 2015). Drought escape is the completion of the life-cycle before a drought condition occurs. This is followed by a waiting period until moisture is available before initiation of the next life cycle (Shavrukov et al. 2017). This is in contrast to the other two strategies which are founded on survival through dry periods (Kooyers 2015). Drought avoidance entails maintaining the water status of tissues by avoiding loss of water and/or through increase water uptake (Blum 2005). Drought tolerance (hardiness) refers to endurance of tissue while under water deficit (Basu et al. 2016). All strategies involve morphological, physiological and molecular adaptive mechanisms (Farooq et al. 2012).

These coping strategies and their corresponding adaptive mechanisms vary between and within species as well as between developmental stages. Generally, drought escape is more relevant to seasonal/annual plants than to cross-seasonal plants whose life cycle require multiple seasons or years. Drought escape is an important mechanism considered by breeders in varietal selection for dry areas (Kooyers 2015; Shavrukov et al. 2017). However, early maturing cultivars are usually associated with lower yield compared to their late maturing counterparts (Mason et al. 2018).

Drought avoidance strategy mainly relies on morphological and physiological adaptive mechanisms. Modification of organs for water storage in succulent xerophytic or Crassulacean Acid Metabolism (CAM) plants are examples of morphological adaptive mechanisms that minimize drought effects (Ogburn and Edwards 2010). Root architectural features are another adaptive mechanism of the drought avoidance strategy. Some plants have deep roots that may allow them to access deep moisture or even the water table. Others have extensive shallow root systems adapted to efficiently and quickly absorb available surface water and to store it in their modified organs. This is particularly effective in hot dry areas where soil surface moisture can be rapidly lost by evaporation (Nobel and North 1996). Reduced plant size and number of leaves, stomatal openings, and efficient photosynthesis pathways (C4 and CAM) (Lara et al. 2011) are also adaptive mechanisms of drought avoidance (Farooq et al. 2012). Most of these strategies have been characterized at the species level but have not been extensively exploited in varietal improvement even though they are key to select drought-resistant plant types for dry areas. However, variations in root system, leaf area and stomatal conductance have been actively selected for in crops such as wheat (Fischer et al. 1998; Friedli et al. 2019; Quarrie and Jones 1979), rice (Gowda et al. 2011; Liu et al. 2012; Uga et al. 2013) and maize (Bänziger et al. 2000).

Drought tolerance/hardiness that enables plant to withstand low turgescence is mainly achieved by osmotic adjustment (Blum 2005). Osmotic adjustment keeps the cellular water balance against the dry surrounding environment by accumulating solutes in the cytoplasm, thereby creating a low cellular water potential. This maintains cell turgor and physiological processes. Osmotic adjustment enables plants to tolerate not only drought, but also other abiotic stresses such as salinity (Zhang et al. 1999) and low temperature (Zhu 2001). These three abiotic stresses can be considered osmotic stresses and breeding for tolerance to one can improve tolerance to the others.

The molecular adaptive mechanisms of plants to drought and other abiotic stresses is complex and appears to involve all three moisture deficit coping strategies in plants. Crop breeding is mainly designed to exploit the available genetic variation within the targeted genepool. Although breeding for drought using whole plant phenotyping seems simple, the genetic bases of the trait is complex (Blum 2011a). Current advances in molecular genetics and data sciences enable the dissection of complex traits of through genome-wide association studies that may provide new tools for accelerating or improving the efficiency of breeding better drought-resistant crops.

1.3. Genome-wide association study (GWAS) and its application

Genome-Wide Association Study (GWAS) refers to a methodology that associates genetic variants to phenotypes using a set of diverse individuals. This approach has now evolved for nearly two decades with thousands of successful phenotype-genotype association studies in many species. GWAS initially started with the human genome and focussed mainly on the genetic bases of complex diseases. To date, GWAS has identified ~3700 genetic variants associated with human health risk factors (Mills and Rahal 2019). GWAS has also been applied to several species, especially those of economically valuable plants and livestock (Gupta et al.

2019). In plant species, these studies identified genetic variants associated with important yield and yield-related traits, including biotic and abiotic stress resistance (Challa and Neelapu 2018).

The GWAS approach requires two sets of data: 1) genome-wide markers, usually single nucleotide polymorphisms (SNPs) and 2) precise phenotypic scores/measurements of all the individuals under study. The main idea is to use the SNP variants to identify regions of the genome significantly associated to the phenotypes measured. For instance, an association between a chromosomal location and a trait (resistance to a disease for example) would be declared if all or most of the individuals carrying a cytosine on chromosome 1 at position 500 for example were also resistant to the disease being evaluated. Compared to the traditional bi-parental QTL analysis, GWAS has the potential to identify markers that are closer to the gene responsible for the phenotype because GWAS relies on a much larger number of recombination events as represented by the historical recombination of the germplasm collection (Myles et al. 2009).

Trait-associated variants explain large portion of phenotypic variation in plants compare to variants identified in humans (Brachi et al. 2011) suggesting the usefulness of GWAS in detecting genomic regions in crop plants underlying the complex yield and yield-related traits. Unlike the traditional quantitative trait locus (QTL) mapping that relies on the limited allelic diversity provided by the bi-parental crosses, GWAS relies on historical recombination of a diverse set of individuals; therefore, the larger number of haplotypes provide power to GWASs. In addition, the genotyping data often tends to be denser in GWASs compared to bi-parental crosses. The genome-wide high density of markers also lends greater power of association to GWASs. Finally, the diversity of the germplasm in GWAS translates into segregation for more phenotypic traits and sometimes a broader range of diversity for given traits, rendering this type of studies advantageous over bi-parental QTL mapping.

The number of SNPs required to detect association depends on linkage disequilibrium (LD). LD refers the non-random allelic association of loci in a population (Slatkin 2008). The LD size is the distance between loci (referring here to SNPs) harboring non-randomly associated alleles. SNPs within the same LD block are assumed to be co-inherited during recombination; hence, SNPs in LD can be considered somewhat "redundant". In high LD region, fewer markers are required to detect association and vice-versa, but identification of the gene or causal variant in these regions is more challenging because they can be physically quite large. LD varies across genomes, species, level of fixation and inbreeding (Bush and Moore 2012). The most commonly used measurements of LD are the ratio of the deviation from the Hardy-Weinberg equilibrium to the maximum allelic frequency (D') and, the coefficient of determination/squared coefficient of correlation (r^2) (Mueller 2004; Soto-Cerda and Cloutier 2012). To avoid the high LD between rare and abundant alleles, most plant GWAS studies use the r^2 values which take into account allelic frequencies (Hill and Robertson 1968; VanLiere and Rosenberg 2008). The threshold r^2=0.1 is widely used, considering SNPs with r^2>0.1 to be in LD (Soto-Cerda and Cloutier 2012; Vos et al. 2017). This value is also often used to suggest the size of the region to be investigated in order to identify putative genes responsible for a trait identified by an associated marker (often SNP). This guideline should not be overstated considering that a recent study showed that causal gene(s) may be located much further away from the SNP(s) than physical distance assumed in many reports (Brodie et al. 2016).

1.4. Flax (*Linum usitatissimum* L.)

Flax (*Linum usitatissimum*) is a member of the *Linaceae* family, and is the only cultivated species of the *Linum* genus that comprises ~230 diploid species (Gray 1988) with 2n ranging from 16 to 60 chromosomes (Diederichsen and Richards 2003). Flax has 30 chromosomes (2n = 2x = 30) and a genome size of approximately 370 Mb (Wang et al 2012). Its closest relative is

pale flax or *L. bienne* Mill. *(syn. L. angustifolium* Hunds.*)* (Diederichsen and Fu 2006) with which flax can readily be hybridized.

1.4.1. Origin and historical use of flax

The center of origin of flax is not clear. Vavilov (1951) hypothesized one of the following five regions as the center of origin: Indian subcontinent, Central Asia, Near-East, Mediterranean and Abyssinian regions. Flax is among the earliest domesticated plants and is one of the founding members of the fertile crescent crops (Zohary and Hopf 2000). Archeological records estimate the oldest cultivation ~8,500-9,000 years ago at Tell Ramad in Syria (Van Zeist and Bakker-

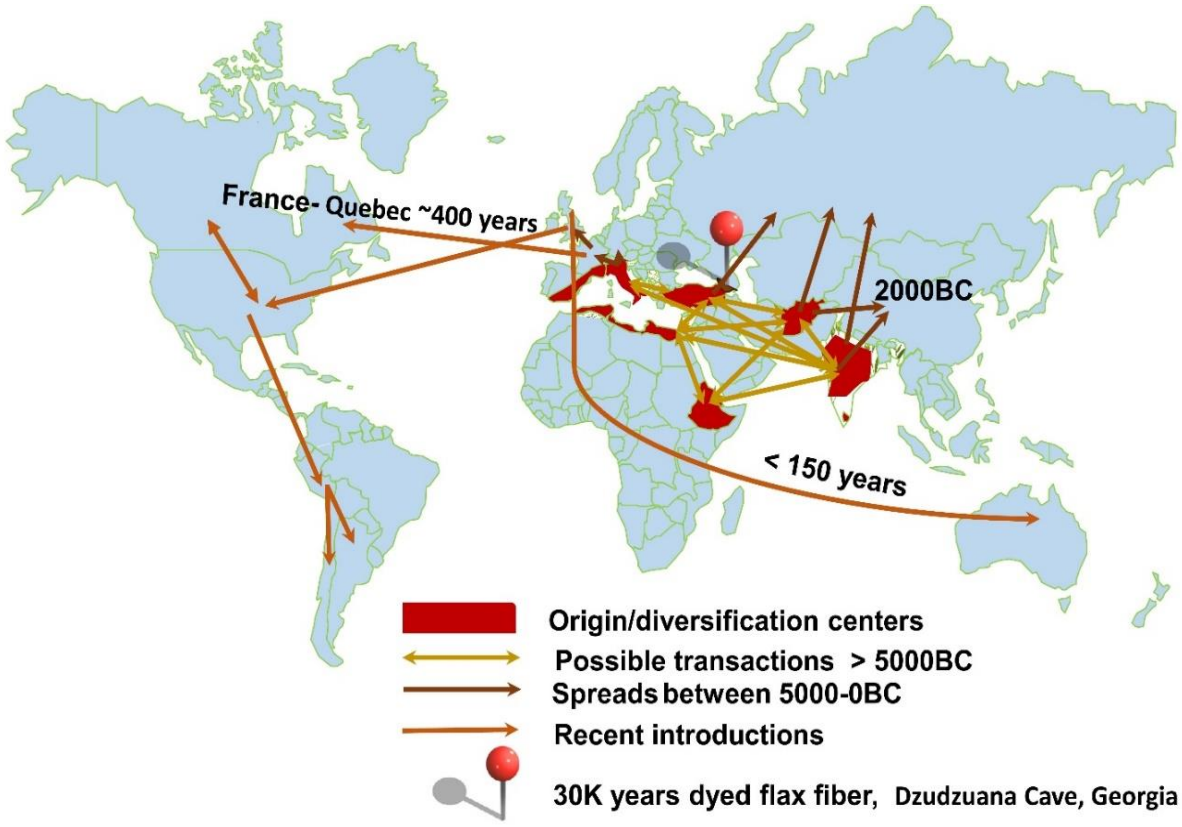


Figure 1.2 *Postulated centers of origin and diversification, historical use and spread of flax.*

Heeres 1975) where it was possibly first domesticated as an oilseed (Allaby et al. 2005; Van Zeist and Bakker-Heeres 1975). However, paleontological evidence suggests that it had been used as a fiber source long before its domestication because fibers were discovered as early as

30,000 years ago (Kvavadze et al. 2009). In its post-domestication era, flax has been grown as both a fiber and an oilseed crop for nearly the full length of its cultivation history (Herbig and Maier 2011). Historical evidence indicate its spread to most of its current growing regions before the first century AD (Herbig and Maier 2011; Wang et al. 2018; Zohary and Hopf 2000). It was one of the crops of ancient agrarian areas such as Egypt, Mesopotamia (Zohary and Hopf 2000) and the Indus Valley (Vaisey-Genser and Morris 2003). Flax spread to its current production regions: Americas and Oceania, following human migrations to these regions (Figure 1.2).

Flax is a versatile crop with multiple uses, including industrial oils (Singh et al. 2011b). *Usitatissimum* means most useful, which reflects the versatility of the crop. Anthropogenic selection resulted in two morphotypes of flax genotypes: fiber and oilseed. The oilseed types have short, branchy stems, large bolls and seeds with higher oil content, whereas the fiber types are tall, less branched and bear small bolls and seeds of low oil content (Herbig and Maier 2011). Flax is the oldest source of fiber used by mankind (Kvavadze et al. 2009). Flax remains the earliest crop used in the textile industry. Indeed, linen was used to wrap mummies throughout the history of ancient Egypt (Davey 2017) and, the production of high-quality linen clothes continues today (Hall et al. 2016). Flax fiber is also used in paper industries to produce high-quality products including banknotes (Carter 1993; Marques et al. 2010). Flax oil is a quality industrial oil widely used in the production of paints, varnishes, inks, and linoleum flooring (Green and Marshall 1984). Flaxseed (linseed) has recently gained some attention by the food, feed and even pharmaceutical industries because of its nutritional and medicinal values and as a source of functional foods (Goyal et al. 2014; Rabetafika et al. 2011; Singh et al. 2011b). Its high alpha linolenic acid (ALA) content, which usually accounts for >55% of its total fatty acids (Green and Marshall 1984; Przybylski 2005), makes flax the richest plant source of omega-3 fatty acids (Morris 2007). Consumption of flax seeds and oil has been associated with health benefits such as a preventative role of several cancers (Chen et al. 2002). Flax has

traditionally been consumed for its medicinal values in countries such as China, Ethiopia (Geleta et al. 2002) and India (Tolkachev and Zhuchenko 2004).

1.4.2. Recent production status of flax

Despite its versatility with all the aforementioned uses, flax production is constrained by different biotic and abiotic limiting factors, including drought, that push growers away from flax who shift to other better paying crops. According to the recently available 10-year data (2008-2017 FAOSTAT, http://www.fao.org/faostat/en/#data/QC), flax is grown on a world total area of approximately ~2.4 million hectares, producing ~2.4 million tons annually. More than 90% of the total production of the crop is the oilseed type. Canada has continuously been the leading producer since 1993 with the exception of 2011 and the last two years. The Russian Federation and China are also long-time top producers. Recently, Kazakhstan has become the lead producer. Flax seed yield rarely exceeds 2.5 tons per hectare under farm conditions (Wittkop et al. 2009) and the average is a little less than a ton per hectare over all the growing areas of the world (FAOSTAT 2017). The low yield emanates from non-capitalization of the yield potential of the crop and the impact of various biotic and abiotic stresses. The global production, and production areas of flax, have been declining with only half a ton/ha in yield gain since 1961 (Figure 1.3).

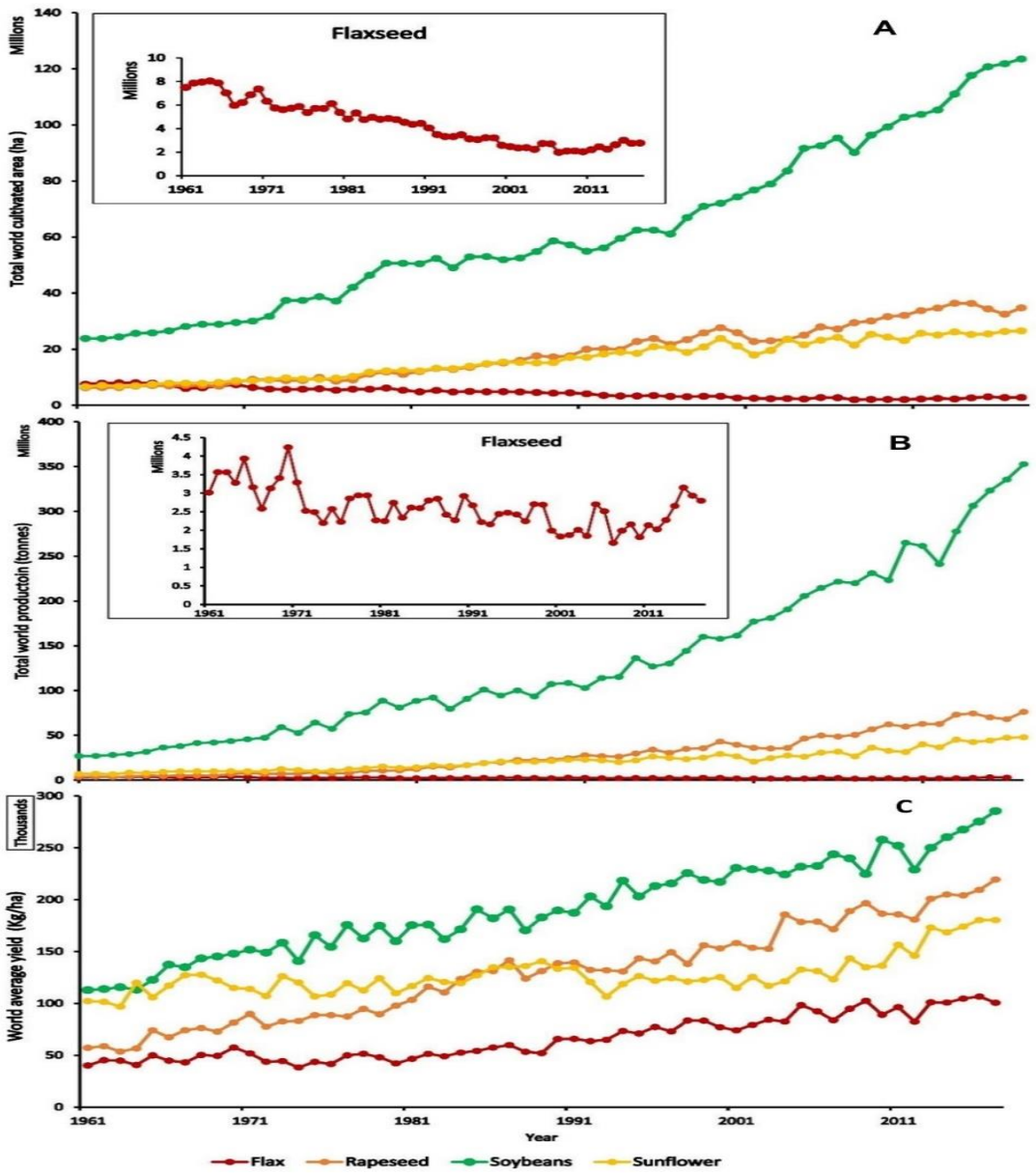


Figure 1.3 *Global production and productivity statistics of flax compared to the major oilseed crops.*
*(**A**) Total world cultivated area (million hectares (ha)), (**B**) Total world production (million tonnes) and (**C**) World average yield (thousand Kg/ha).*

1.4.3. Flax and moisture stress

Moisture stress is one the major challenges of flax production (Lisson and Mendham 2000) as it is in agricultural production at large (Solh and van Ginkel 2014). Flax performs better under moist and well-drained soil conditions (Casa et al. 1999). A few reports claimed flax to be drought tolerant, especially in the presence of mycorrhizae (Augé 2000; Von Reichenbach and Schönbeck 1995), which enhance health and tolerance to diseases (Dugassa et al. 1996). In contrast, numerous reports on the susceptibility of the crop to moisture limitations are far more abundant (Hocking et al. 1997; Lisson and Mendham 2000). Flax is especially sensitive to water stress during early growth (Guo et al. 2012) and reproductive stages (Casa et al. 1999; Lisson and Mendham 2000; Mishra and Singh 2010). To attain maximum yield, moisture availability during flowering and the following ripening stage is critical (Mirshekari et al. 2012). At this stage, the plant requires the most water while root development nearly ceases (Rowland 1998). Compared to other oilseed crops such as canola, sunflower, and safflower, flax has a shallow root system and, as such, it mainly relies on moisture available in the top layer of the soil (Hocking et al. 1997; Kar et al. 2007). Flax was also reported to be more sensitive to water deficit than cereal crops like wheat (Morillon and Lassalles 2002).

In many experiments, flax performed poorly under weedy conditions (Alessi and Power 1970; Bell and Nalewaja 1968; Klimek-Kopyra et al. 2015). Flax yields better in pure stand than when intercropped with legumes (Zając et al. 2013). Nevertheless, higher yields were obtained when it was intercropped with chickpea in furrow-irrigated fields, i.e., when chickpea was planted on the upraised bed/ridge and flax was planted in the furrow where it had better access to water (Ahlawat and Gangaiah 2003). However, these authors recorded low flax yield when it was intercropped with the same chickpea on flat fields. Alessi and Power (1970) attributed yield losses in flax under weedy field conditions to moisture scarcity. Klimek-Kopyra et al. (2015) also

noticed poor flax performance when intercropped under drier conditions, implying the inability of flax to perform well under water-stressed conditions.

Flax is highly responsive to irrigation (Gabiana 2005). Lisson and Mendham (2000) reported significant positive yield response to irrigation with declining precipitation. Gabiana (2005) reported >60% seed and >77% oil yield increases in response to irrigation. Townshend and Bolyend (2008) recorded an average seed yield >4 tons/ha in a year when rainfall doubled in the last two months of the growing season compared to <1 ton/ha in dry periods. Availability of adequate moisture, especially during the flowering to the ripening stages, extends flowering time and enables plants to produce more bolls with better quality seeds (Kar et al. 2007; Rowland 1998). With a continuous supply of water, plants may even continue to produce new branches during the post-flowering period (Hocking 1995).

Although flax is generally considered a drought-sensitive crop, it can grow in a wide range of eco-geographical environments (Sertse et al. 2019b), suggesting the adaption of the genepool to varying environmental factors including amplitudes of moisture conditions (Diederichsen et al. 2006). A wide range of variations for important agronomic traits such as yield, thousand seed weight, and plant height were observed among 96 flax accessions grown under drought conditions in Canada and Russia (Diederichsen et al. 2006). Similarly, Lisson and Mendham (2000) reported significant yield response variations among flax varieties exposed to different water stress levels. Diederichsen et al. (2008) observed disease and drought resistance of North American compared to Eurasian germplasm. Foster et al. (1998) suggested short and early flowering flax varieties for dry environments as a mechanism of drought avoidance. Differences in the degree of tolerance to water stress are a function of numerous drought-related traits such as photosynthetic ability, stomatal conductance and root systems (Cattivelli et al. 2008).

Currently, gene banks around the world housed approximately 17,000 flax accessions (https://www.genesys-pgr.org). Diederichsen et al. (2013) characterized a collection of 3,378 flax accessions from the Plant Gene Resource Canada (PGRC) and identified 381 accessions representing the breadth of the diversity of the collection. This subset formed the original Canadian core collection (Diederichsen et al. 2013). A total of 26 accessions of interest to Canadian breeders were since added, forming the current collection of 407 accessions. Genetic characterization of this augmented core collection revealed high genetic diversity, one of the characteristic that contributes to the suitability of the collection for association mapping (Soto-Cerda et al. 2013).

Recent developments in molecular genetics, coupled with advances in computational techniques (Pirooznia et al. 2015), have enabled the deployment of high resolution and robust approaches to explore the genetic diversity and adaptation spectrums across environmental gradients of the entire genepool and to identify specific genomic region(s) underlying a trait of interest using GWAS (El-Soda et al. 2014; Zhu et al. 2008a). Assessing genetic diversity of the existing genepool is an important way of understanding the available adaptation potential of a crop to gradients of environmental conditions (Onda and Mochida 2016). The current sequencing technologies have enabled whole-genome sequencing of many crop species including large and complex genomes at a fraction of prior cost and/or at much higher resolution than used to be (Michael and VanBuren 2015). To date, the complete genome sequences of 387 land plants including flax have been made available (https://www.ncbi.nlm.nih.gov/genome/browse accessed on July 13, 2019); and flax was the 12th one sequenced (Wang et al. 2012). This has created an opportunity to explore genome-wide polymorphism, allelic and/or haplotype distribution patterns across entire habitats of the species to discover potential adaptation signals. GWAS is performed in order to identify important loci or genomic regions underlying important traits responsible for the performance of

plants under different environmental conditions through phenotype-genotype association analysis. GWAS has indeed been applied to several crops to identify quantitative trait loci (QTL) associated with traits including drought tolerance in wheat (Sukumaran et al. 2015; Tricker et al. 2018), rice (Guo et al. 2018; Muthukumar et al. 2015), maize (Farfan et al. 2015; Wang et al. 2016b), soybean (Sonah et al. 2015), sorghum (Badigannavar et al. 2018) and sesame (Dossa et al. 2019). GWAS has also been used in flax to identify loci associated with seed and oil quality traits (Soto-Cerda et al. 2018; Soto-Cerda et al. 2014; You et al. 2018a), disease resistance (He et al. 2018) and agronomic traits such as plant height and fiber length (Xie et al. 2018a). However, GWAS for drought-tolerance-related traits in flax has not been performed to date.

1.5. The scope and purpose of this study

This study was designed to assess global-scale population structure and eco-geographic distribution patterns of flax genetic diversity, and to perform GWAS for drought-tolerance-related traits. The work provides insight into the historical events, such as anthropogenic and eco-geographic selection pressures, that shaped the current population structure and gave rise to the genetic diversity of the crop, thereby generating relevant information for breeders and conservationists (Chapter 2). It also investigates loci underlying various traits of early shoot and root development, and identifies candidate genes for the targeted traits (Chapter 3). Finally, it explores genomic regions responsible for important drought-tolerance traits, including yield, and proposes candidate genes mediating the makeup of these traits (Chapter 4).

To achieve the above outcomes, the following major activities were conducted:

1. Generating single nucleotide polymorphism (SNP) data and data quality control;
2. Performing genetic structure and haplotype distribution analyses of the core collection;

3. Evaluating early root and shoot development of the mini-core collection in controlled environments;
4. Phenotyping the mini-core collection for drought tolerance related traits under field conditions;
5. Performing phenotype-genotype association analyses for experiments described in 3 and 4 based on the SNP data of the mini-core and identifying loci associated with the traits under consideration;
6. Performing gene mining of the predicted genes harbored at loci associated with the significant traits and further investigating their potential role based on previously reported role of their orthologues in other species;
7. Providing general remarks and recommendations based on discoveries.

1.6. Hypotheses

1. The flax population genetic stratification and variation are independent of variations in eco-geographic and anthropogenic factors across the growing regions of the crop.
2. The early root and shoot traits variations among flax accessions can be measured using hydroponic pouch system, environmental effect and their interactions for drought-related traits and indices can be measured in irrigated and non-irrigated field experiments conducted at multiple locations and years using a mini-core collection of 115 flax individuals.
3. The genetic variation, environmental effect and their interactions for drought-related traits and indices can be measured in irrigated and non-irrigated field experiments conducted at multiple locations and years using a mini-core collection of 115 flax individuals.

4. Regions of the flax genomes associated with the genetic variance for early root & shoot, drought-related traits and drought-indices can be identified using single- and multi-locus models to associate genome-wide single nucleotide polymorphisms to the phenotypic trait data, and their contribution to the genetic variance can be quantified.

1.7. Material and methods

In this introductory chapter, only highlights of the materials and methods used in the subsequent chapters are described, pointing out specific details of some of the techniques. More extensive and specific descriptions can be found within each chapter.

1.7.1. Plant materials

In this study, two sets of flax genotypes were used: a core collection (n=407) and a selected subset thereof called the mini-core collection (n=115). Both comprised genotypes from nearly all major flax growing regions of the world (Figure 1.4). The mini-core represents approximately 95% of the genetic diversity of the core collection and it was selected for trait characterization. The entire core collection was sequenced to generate a large number of SNPs. The core collection was used to evaluate the worldwide genetic diversity, eco-geographic adaptation, and haplotype distribution of flax. The mini-core was phenotyped for 11 drought-related agronomic traits under two watering conditions in the field, as well as for two shoot and 14 root traits, using an indoor platform designed to study these traits.

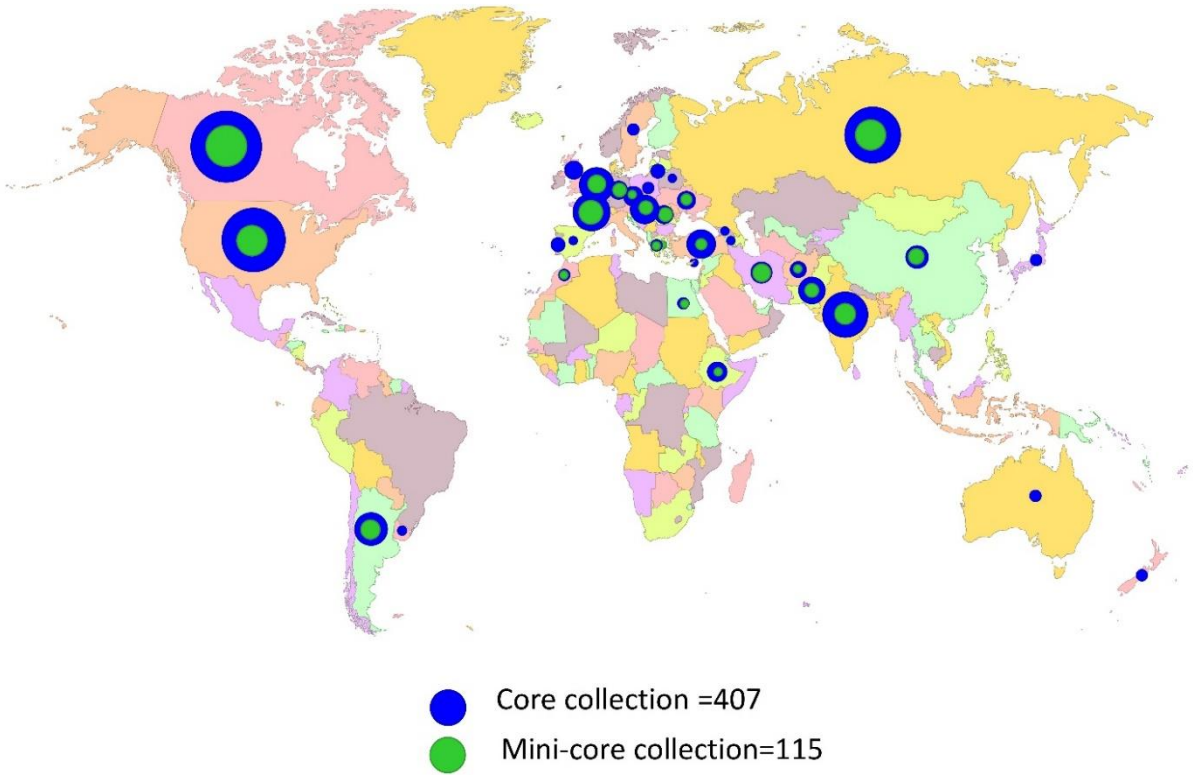


Figure 1.4 *The global distribution of the core and mini-core collection.*
The size of the circles is proportional to the number of samples in the collection.

1.7.2. Phenotyping

High-quality phenotypic data is crucial for GWAS (Araus and Cairns 2014; Harvey and Strauss 2009). Although drought-resistance traits are claimed to be simple from physiological, agronomic, and whole-plant perspectives (Blum 2011a) , their genetic control is complex (Salekdeh et al. 2009; Yue et al. 2006). Maximum yield in drought conditions depends on the time, duration, and intensity of the drought (Mishra and Singh 2010). Traits that maximize yield in one environment may not play a role in water-limiting environments because environments are variable for many other factors (Richards 1996). For this study, two experiments were conducted. The field experiment looked at plant agronomic traits under irrigated and non-irrigated conditions. The replicated controlled environment experiment strictly characterized the

germplasm for early root and shoot development because roots are important organs for drought tolerance.

a) Early root and shoot traits

Roots are important organs that play a vital role in mediating the interaction between the plant and the bio-physico-chemically dynamic rhizosphere, especially in terrestrial vascular flora. The root distribution and pattern along the soil layer determines the volume of soil accessible to the plant and somewhat defines its efficiency for water and nutrient uptake (Yu et al. 2006). Variations in root architecture were reported to result in significant differences in yield production under drought conditions (Henry et al. 2011). Dense root systems have been found to be more important than maximum root depth (Yu et al. 2006), as extensive depth might be associated with an energy cost (Reynolds et al. 2012). Seedling vigor that is manifested by early shoot and root performance is important for the survival and subsequent growth and final yield performance of crops. In this study, two shoot (shoot dry weight and shoot length) and 14 root traits were evaluated (see Chapter 3 for details)

b) Field experiment

The field-based phenotyping mainly focused on the traditional yield-related agronomic, phenological, and physiological traits. Yield is the ultimate trait and the most important objective of breeding programs (Araus et al. 2008; Blum 2011b). Yield, a polygenic trait, is a function of multiple yield components that are themselves polygenic in nature, i.e., associated with a large number of QTL of small to large effects and with epistatic-interactions. In order to consider drought-related traits in relation with yield, high correlation and heritability are required (Araus et al. 2012; Tuberosa 2012). The selected traits were based on this important assumption. Apart from the commonly-known agronomic and phenological traits, canopy temperature (CT) was included. CT is an adaptive and heritable trait associated with the performance of plants under

drought conditions (Olivares-Villegas et al. 2007), and consequently linked to crop yield (Amani et al. 1996). CT is influenced by radiation (sunshine/cloud), relative humidity, soil moisture in the root zone and wind (Pietragalla 2012). Variations in CT can reflect differences in morphology or physiology such as root (Lopes and Reynolds 2010), leaf size and shape (Pask and Pietragalla 2012) or stomatal opening (Reynolds et al. 2005). Phenological traits such as flowering and maturity are not only important yield-related traits but are also important indicators of wide-range eco-geographic adaptation, drought escape and of response to drastic environmental changes, especially temperature and light.

c) Experimental designs

The field experiment was conducted using a modified augmented design (MAD) type 2 (Lin and Poushnisky 1985) using a four row by four column design for each of the irrigated and non-irrigated field blocks (see Fig. 4.1). In this design, only the main plot and subplot controls were replicated. Each field block was divided into 16 main plots and each main plots into nine subplots for a total of 144 subplots for each of the irrigated and non-irrigated blocks. The central subplot of each main plot was allocated to the main plot control CDC Bethune and 16 randomly-selected subplots were assigned the two subplot controls, namely Prairie Thunder and Macbeth. The test lines were un-replicated and randomly assigned to the remaining subplots (see detail in Chapter 4). The experiment was conducted at two locations over three years. The early root and shoot experiment was performed in a growth chamber using a hydroponic pouch system modified from Hund et al. (2009) (see details in Chapter 3).

1.7.3. Genotyping

SNP data was generated by resequencing each of the 407 genotypes of the core collection on the Illumina HiSeq 2000 platform using the 100-bp paired-end mode following the manufacturer's instructions. The sequencing yielded an average of 69.5 million reads per

genotype which is approximately 18× coverage given the genome size of flax of ~370 Mb (Wang et al. 2012). Reads were aligned to the chromosome-scale pseudomolecules (You et al. 2018b) of the flax reference genome using BWA v0.5.9 (Li and Durbin 2009) and SNPs were called using the AGSNP pipeline package (You et al. 2011). Prior to quality-control filtering, a total of 4.4 M candidate SNPs were recovered by the pipeline. Upon removal of candidate SNPs with minor allele frequency (MAF) <0.01, a total of 1.77 M SNPs were retained. Finally, exclusion of SNPs with MAF <0.05, call rate <0.8 and SNPs located in transposable elements resulted in a subset of 570,443 SNPs. This dataset was stored and further filtering and quality-control criteria were applied to generate the various datasets used for the different analyses conducted (see individual chapters for details).

1.7.4. Linkage disequilibrium (LD) and minimum required SNPs

Different numbers of SNPs are used for the experiments described in chapters 2-4. This section provides the necessary background information to support the decision-making processes that determined the filtering criteria and the size of the target regions for the candidate gene searches. The LD of all SNPs on each chromosome was computed based on a total ~51K SNPs using TASSEL v5.2 (Bradbury et al. 2007). LD decay was constructed by fixing the threshold at r^2=0.1 (Soto-Cerda and Cloutier 2012; Vos et al. 2017) for each chromosome (Figure 1.5). The average physical distance where LD falls below the indicated threshold was considered the maximum region size to be investigate for the presence of a causative gene of a SNP-trait association.

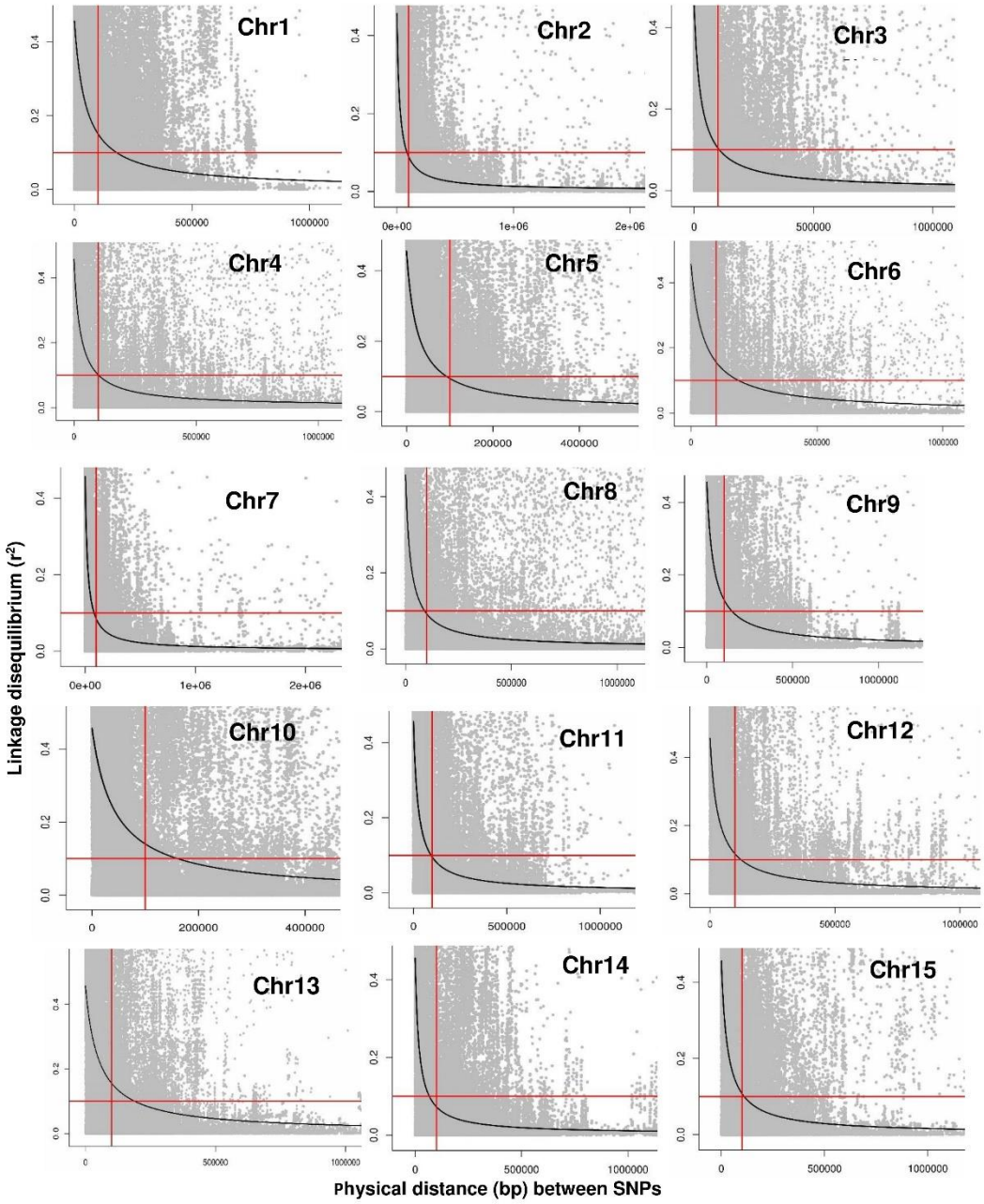


Figure 1.5 *Linkage disequilibrium (LD) decay on each chromosome.*

The horizontal and vertical red lines indicate r^2=0.1 and a physical distance between two SNPs of 100 Kb, respectively

2. The genetic structure of flax illustrates environmental and anthropogenic selections that gave rise to its eco-geographical adaptation

Demissew Sertse [a,b], Frank M. You [a], Sridhar Ravichandran [a], Sylvie Cloutier [a,b,*]

[a] Ottawa Research and Development Center, Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Ontario, Canada K1A 0C6

[b] Department of Biology, University of Ottawa, 30 Marie Curie, Ottawa ON Canada K1N 6N5

This chapter has been published in Molecular Phylogenetics and Evolution 137(2019): 22-32

DOI: 10.1016/j.ympev.2019.04.010

Authors' contribution

DS conceived the idea of the study, obtained climate and environment-related data from different sources, performed the data analyses and prepared the manuscript. SR contributed ideas for part of the data analyses. SC and FY produced the genetic data. SC, BS, FY and SR read and edited the manuscript. All authors read and approved the final version.

2.1. Abstract

Flax, one of the eight founder crops of agriculture, has been cultivated for its oil and/or fiber for millennia. Understanding genetic divergence and geographic origins of germplasm in line with their cultivation history and ecological adaptation are essential for conservation and breeding. Here we performed analysis of global-scale flax-population genetic structure and eco-geographic adaptation based on more than 51,000 single nucleotide polymorphic sites and 383 flax accessions from a core collection representing 37 growing countries. The genetic structure analysis resulted in a total of 12 populations that were pooled into four major groups: Temperate, South Asian, Abyssinian and Mediterranean. The vast majority (n=335) belonged to the Temperate group that comprised eight populations including one dominated by fiber flax accessions. Genetic variation between fiber and oil morphotypes was less pronounced than variation within morphotypes. The genetic variation among groups and populations was attributed in part to eco-geographic and anthropogenic factors. Analysis for adaptation signals yielded loci under strong selection by environmental factors such as day length. A high concentration of private haplotypes as observed in the South Asian, Mediterranean and Abyssinian populations despite their low genotype representation, hinting at the long history of the crop in these regions. The addition of genotypes from these three regions would enrich the core collection through the capture of a wider genetic breadth for breeding and conservation.

Keywords: Flax core collection; population structure; eco-geography; haplotype distribution

2.2. Introduction

Linum usitatissimum (L.), commonly known as linseed (flaxseed) or fiber flax depending on its main use as oilseed or fiber crop, is used in the industrial oil, feed, food and textile industries (Singh et al., 2011). Recently, linseed has gained attention for its nutritional and medicinal values mainly attributed to its high alpha linolenic acid content which is an essential omega-3 fatty acid (Gebauer et al. 2006; Morris 2007; Przybylski 2005).

As one of the founder crops of agriculture in the Fertile Crescent, flax has been important throughout agrarian history (Weiss and Zohary, 2011). *Linum usitatissimum* is the only cultivated species of the genus *Linum* which comprises over 200 species (Diederichsen, 2007). The diploid crop (2n=2x=30) has a genome size of ~370 Mb (Wang et al., 2012) and it is believed to have been domesticated as a single event (Allaby et al., 2005). Archeological evidence suggests the earliest cultivation to have occurred in the present day Syria (van Zeist and Bakker-Heeres, 1975). Seed size investigations of early germplasm indicated that the crop might have first been domesticated as an oilseed (Allaby et al., 2005; Van Zeist and Bakker-Heeres, 1975). However, fiber flax cultivation period was estimated to coincide largely with the entire cultivation history of the crop with evidence of selection for fiber flax as early as the fourth millennium BC (Herbig and Maier, 2011). The history of the crop indicates that it spread to most of the Old World many centuries ago. Flax was present in ancient Mesopotamian and Egyptian irrigated fields of ca 7000 years (Zohary and Hopf, 2000) and, around 3000 BC, it reached as far east as present China (Wang et al., 2018).

The successes of the crop into diverse environments and independent selection processes throughout its cultivation history have resulted in substantial genetic divergence. Some of the early flax-growing regions eventually became secondary centers of diversification (Vavilov, 1951). The ecological resilience of the crop is exemplified by its adaptation to warm regions in

the Indian subcontinent (Rajwade et al. 2010) and to cool temperate regions of Eurasia and the Americas (Casa et al., 1999). The current geographic distribution of flax morphotypes also seems to reflect anthropogenic end-uses. While South Asian and East African countries have been growing flax exclusively as an oilseed (Chandrawati et al. 2014; Zeng et al. 2015), the cool regions of Europe and the Russian Federation continue to grow flax for both its oilseed and stem fibers (Pavelek et al. 2014). The historically hemp- and silk-producing countries such as China were also predominantly growing linseed (oilseed flax) throughout the crop's history until fiber type was introduced ~100 years ago (Liu et al. 2011) suggesting that use-directed selection of genotypes in different regions contributed to the genetic divergence of the crop. Contrary to its ecological elasticity and use versatility, flax acreage is globally shrinking, in part, because yield gains during the last 50 years have been small compared to most field crops including oilseeds such as rapeseed (http://www.fao.org/faostat/en/#data/QC).

Outcrossing in flax is less than 6% (Dillman, 1938). However, artificial crossing by breeding programs has resulted in the emergence of new recombinants. Genebanks currently house more than 17,000 flax accessions (https://www.genesys-pgr.org) and *in situ* reservoirs are expected to harbor even more diversity.

Advancements in sequencing technologies and data sciences have provided opportunities to apply genomics to decipher the wide-range of genetic diversity of historical parental lineages, and to better understand gene flow and driving forces that played major roles in shaping the present genetic structure in crops such as wheat (Cavanagh et al. 2013), rice (McNally et al. 2009), and the model plant *Arabidopsis thaliana* (The 1001 Genomes Consortium 2016). Such advancements have encouraged plant breeders to access broader germplasm pools using molecular markers (Nadeem et al., 2018). Here, we performed a genome-wide genetic diversity analysis of 383 flax genotypes representing the major flax-growing areas of the world: 1) to

trace the ancestral background and extent of admixture in this core collection; 2) to determine linkages between genetic structure of the crop and climatic information and/or historical uses; 3) to identify eco-geographical regions harboring the greatest genetic diversity and to define their unique haplotypes for conservation and breeding considerations.

2.3. Materials and methods

2.3.1. Plant materials

In this study, the flax core collection (n=407) (Diederichsen et al. 2013) that comprises 91 fiber, 287 oil and 30 unassigned genotypes collected from 37 countries representing the major flax growing regions of the world was used. Of the genotypes of the core collection, ~37% (n=151) are from New World regions including the Americas and Oceania. The remaining originated from Old World flax-growing regions including accessions from the proposed centers of origin of the crop (Vavilov, 1951) and regions with a long history of flax cultivation. These genotypes have been characterized for major taxonomic, agronomic, disease and quality traits at multiple locations over many years (You et al., 2017).

2.3.2. DNA extraction and genotyping

Genomic DNA of each accession was extracted from 75-100 mg of young leaf tissue of a single plant using the Qiagen DNeasy 96 plant kit (Qiagen Sciences, Maryland, USA). Genomic libraries constructed based on the Illumina (Illumina, San Diego, CA) plate-based method were sequenced by the Michael Smith Genome Sciences Centre of the BC Cancer Agency (Vancouver, BC, Canada). A total of six samples were indexed per Illumina HiSeq 2000 lane and the sequencing was performed in 100-bp paired-end mode following the manufacturer's instructions.

2.3.3. Single nucleotide polymorphism (SNP) calling and filtering

Sequence reads were processed as previously described (Kumar et al. 2012). Processed reads from each accession were aligned to the chromosome-scale pseudomolecules (You et al., 2018) of the flax reference genome using BWA v0.5.9 (Li and Durbin, 2009). SAMtools 0.1.0 (Li et al., 2009) was used to convert the aligned reads to a sorted binary alignment map (BAM) file and SNP calling was carried out using the AGSNP pipeline package (You et al., 2011). Candidate SNPs occurring at least three times in any accession were retained.

Initially, loci with a minor allele frequency (MAF) < 0.01 were removed and then, in a secondary filtering, SNPs with a MAF < 0.05, call rate < 80%, and those located in annotated transposable elements were removed. Finally, a stringent filtering that retained only loci with a call rate exceeding 95% was applied. At this step, genotypes that scored over 10% missing were omitted. This final data set was used for further analysis.

2.3.4. Population structure analysis

Individuals' ancestral estimate (Q) matrix together with cross-validation error to decide the appropriate number of populations (K) (Alexander and Lange 2011) ranging from K = 2 to 20 was generated using ADMIXTURE v1.3 with default parameters (Alexander et al. 2009). K with the lowest cross-validation error (Alexander and Lange 2011) was selected as an appropriate number of populations. Here K=11 was the suggested appropriate number of populations. Apart from the case in two strongly admixed populations, individuals sharing high ancestral estimate (Q ≥ 50) were assigned to a cluster except for three scenarios. First, if individuals exceeded 45% in their ancestral proportion in a cluster and if their remaining ancestral proportion was distributed to others in each with a maximum of 40%, they were assigned to a cluster where >45% of their ancestral proportion originated. Second, groups of individuals that showed consistent clustering with Q > 90 in all K = 12–20 but not at K = 11 were assigned to independent

populations. Third, individuals that did not meet the 50% ancestral value at K = 11 but that clustered together with members of known clusters at K = 9, 10, 12 and 13 were assigned to their respective populations. Independent clusters had at least one individual with Q > 90 except for clusters representing strongly admixed individuals. Population names were decided based on the geographic origin of member individuals, and where at least three individuals had Q > 60 and/or one individual had Q > 90. Accordingly, clusters as populations were named after historically known flax growing regions of the Old World with the exception of those dominated by recent New World cultivars and admixtures. Following the assignment of individuals to populations, the program package CLUMPACK (Kopelman et al. 2015) was used to summarize the ADMIXTURE results into structure plots.

To substantiate the results from admixture analysis, principal component (PCA) and neighbor-joining phylogenetic analyses were performed using TASSEL v5.2 (Bradbury et al., 2007). PCAs were graphically summarized using R and a neighbor-joining phylogenetic tree was visualized using Tree of Life (iTOL) v3 (Letunic and Bork, 2016). The populations were further clustered into major groups based on their relationships as visualized by the first three PCs.

2.3.5. Population genetic variation

To estimate the population pairwise and locus-level variations, fixation indices with 95% confidence intervals (CIs) were calculated using the R package diveRsity (Keenan et al. 2013). Outputs were generated based on different algorithms: G_{ST} (Nei and Chesser 1983), Gg_{ST} (Hedrick 2005), GG_{ST} (Meirmans and Hedrick 2011), D_Jost (Jost 2008), F_{ST} (Weir and Cockerham 1984). In subsequent analysis and discussion, the traditional F_{ST} estimate by Weir and Cockerham (1984) was used. SNPs with fixation indices ≥ 0.75 were extracted as high F_{ST}

SNPs for separate further analyses. Total variation and pairwise PhiPT (analogue of F_{ST}) was calculated on these SNPs using GenAlEx 6.5(Peakall and Smouse 2012).

Frequencies and distribution of haplotypes among the populations were generated from the high F_{ST} SNPs using the web-based tool SNiPlay v3 (Dereeper et al. 2011) http://sniplay.southgreen.fr/cgi-bin/analysis_v3.cgi). Populations were putatively assigned to geographic locations based on the passport data of individuals representing likely origin of their assigned population. In the absence of collection site information, we performed a thorough literature search (cited in Appendix I-1) to trace historically dominant flax growing regions of the countries of origin and, applied coordinates of these areas to map the geographic distribution of haplotypes. The haplotypes were geo-referenced using the web-program PhyloGeoViz (Tsai 2011) http://phylogeoviz.org/. To assess the rate of unique (private) haplotypes among the populations, haplotypes that occurred at least in two individuals were selected from each chromosome and frequencies of these haplotypes were calculated for each population.

To reduce discrepancies in the rate of private haplotypes in a population due to differences in sample sizes, the frequencies per population were normalized using the following equation

$$PhapR = \frac{\left(\sum_{chr=1}^{chr=15} \frac{n}{N}\right)}{15}$$

Where $PhapR$ = private haplotype rate; n = number of individuals in a population with unique haplotype at each chromosome; N = total number of individuals in that particular population.

2.3.6. Geographic origin, climate data and mapping

To understand the contribution of climatic factors to the genetic divergence among populations, representative of geographic coordinates of the populations were overlaid on high-density temperature and precipitation grid data (1km^2 or 30 sec) of worldClim2 (Fick and Hijmans 2017) downloaded from http://worldclim.org/version2 using DIVA-GIS 7.5 (Hijmans et al. 2001).

Assuming that climatic variations among regions with long-history of growing flax impacted the genetic divergence, the locations from Old World that better represent the major group were defined based on visual observation of the map. Populations within a group that showed noticeable climate differences were retained as independent representatives even when they were geographically closely located. Vologda (Russia), Allahabad Uttar-Pradesh (India), Faisalabad (Pakistan), Lalibela (Ethiopia) and Braga (Portugal) were selected as representative locations of the major groups. Indian and Pakistan represented the same major group but they were maintained due to substantial variation in their precipitation. For these five representative locations, location-specific 22-year average climate data, including temperature, precipitation and day length were downloaded from https://power.larc.nasa.gov. Correlation of spatial variations of these climate parameters and latitude with pairwise genetic variation of populations that match the above five selected locations were performed following the approach outlined by Mantel (1967) using GenAlEx 6.5 (Peakall and Smouse, 2012). Considering the relevance of climate variables during the cropping months, average daily temperature, day length and total precipitation during the cropping periods of the target location were applied separately. To define cropping months of the target locations, FAO crop calendar chart http://www.amis-outlook.org/amis-about/calendars/en/ was used as guide. To justify the possible residual effect of off-cropping months climatic conditions, Mantel correlations between annual temperature, precipitation with pairwise genetic variation were computed.

2.3.7. Genome-environment association and loci under selection

To identify possible genetic signatures linked to local adaptation, genome-environment association (GEA) was performed based on the latent factor mixed model (LFMM) using the R package lfmm (Frichot et al. 2013). For GEA, genotypes with a higher ancestral proportion (Q >90%) in clusters affiliated to the five location-representative groups and their corresponding climate data, i.e., temperature, precipitation and day length, were used. The genome data was

filtered to retain only loci with ~97.7% call rate and genotypes with <5% missing. The number of ancestral populations, considered as number of latent factors (K), was decided following a cross-validation technique (Alexander and Lange 2011) using the R package LEA (Frichot and François, 2015). Assuming the effects of temperature and precipitation conditions in off-cropping period, especially during the early and late cropping months, GEA was also performed based on annual data of these variables. Bonferroni correction at 0.05/n was applied as threshold *p* value.

2.4. Results

2.4.1. SNP discovery and filtering

An average of 65.9 M reads per genotype were produced, equivalent to an average of 17.8× genome coverage, and 82% mapped to the reference genome (Wang et al., 2012). Initially, 4.4 M candidate SNPs were identified. Removal of loci with minor allele frequency (MAF) <0.01 yielded a subset of 1.77 M SNPs. Exclusion of SNPs with MAF <0.05, call rate < 80% and SNPs in transposable elements resulted in a subset of 570,443 SNPs. In the final filtering step, 51,575 SNPs with a call rate exceeding 95% were recovered. A total of 24 accessions with more than 10% missed calls were discarded; hence, a total of 383 accessions were retained for subsequent analysis. For GEA, 11018 SNPs and 44 genotypes that passed the quality criteria were obtained.

2.4.2. Population genetic structure

ADMIXTURE results indicated a most likely number of populations was K=11 (Figure 2.1A). However, a noticeably aberrant small cluster represented by genotypes of Mediterranean origin starting from K=12 in ADMIXTURE and in all other population structure analyses (Figure 2.1B-E) was observed. This cluster was considered an independent population for a total of 12. These 12 populations were re-clustered into four major groups reflecting the eco-geographical

origins of the genotypes: Temperate (TEMP), South Asian (SA), Abyssinian (ABYS) and Mediterranean (MEDT). Eight populations representing more than 87% of the entire collection clustered into the TEMP group which included the majority of the fiber flax and New World genotypes. Most, but not all, fiber types clustered in populations dominated by oil types. Two populations: fiber one (FIB1) and fiber two (FIB2) contained fiber-types 79% and 32% of their members respectively, and together they accounted for approximately 36% of the 85 fiber genotypes in the final retained core collection (n=383) (Appendix I-2). FIB1 was dominated by Eurasian accessions, while FIB2 contained accessions of diverse geographic origins. As their names indicate, the Old World Admixture (OWAD) and New World Admixture (NWAD) populations contained extensive admixtures. The remainder of the TEMP group included Canadian cultivars (CANC), West and Central European (WEUR), Temperate1 (TEM1) and Temperate 2 (TEM2) populations. Genotypes from traditional flax growing regions such as China fell within the TEMP group (Appendix I-2.). The SA group comprised the Indian (SAIN) and Pakistani (SAPK) populations. ABYS and MEDT groups each encompassed a single population (Figure 2.1B).

PCA results concurred with the previous analyses in that PC1 and PC2, which explain a total of 34.43% of the variance, clearly separated the four groups (Figure 2.1C). PC2 and PC3 discriminated the ABYS and MEDT groups as outliers and fairly illustrated the sub-clustering within the TEMP group but the SA group were confounded within the TEMP group (Figure 2.1D). The neighbor-joining phylogenetic tree assembled more than 75% of the fiber genotypes in a single clade. The SA group also formed a separate clade. The ABYS and MEDT groups were wedged in TEMP population clades (Figure 2.1E-F). ABYS and MEDT individuals however were clearly distinguished by long branches in their respective clades. Generally, the SA, ABYS and MEDT tended to have longer branches (Figure 2.1E).

2.4.3. Genetic variation

Pairwise population differentiation using the 51K SNP data set showed a similar pattern of population variations based on the fixation and related indices from the different algorithms (Appendix I-3). Based on the traditional F_{ST} estimate (Weir and Cockerham, 1984), the highest differentiation (F_{ST} = 0.84) was observed between ABYS and MEDT (Table 2.1). The MEDT appeared to be the most divergent population; its closest population being WEUR with F_{ST} = 0.48. Populations of the TEMP group showed low pairwise variation with maximum F_{ST} of 0.3 observed between the TEM1 and TEM2 populations (Table 2.1).

Differentiation at each locus resulted in 1253 SNPs $F_{ST} \geq 0.75$. The pairwise population PhiPT (analogue to F_{ST}) with this high F_{ST} data set did not differentiate most of the populations within the TEMP group but significantly ($P < 0.05$) discriminated all others from all TEMP group populations. The two SA populations were also significantly differentiated by this estimate. The highest PhiPT value of 0.987 was computed for the MEDT and TEMP subgroups FIB1 and CANC (Table 2.2).

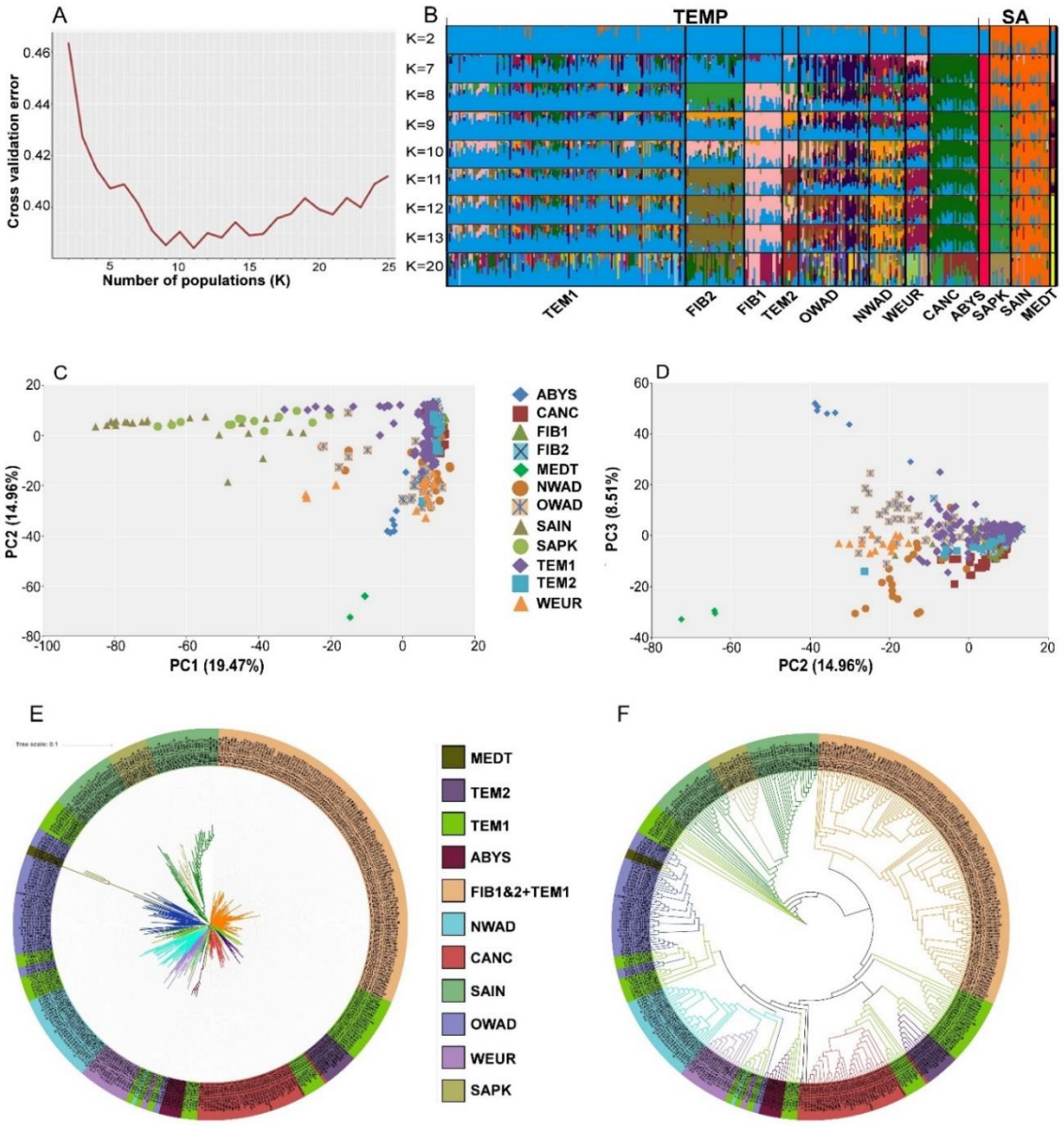


Figure 2.1 *Population genetic structure of the core collection.*

*(**A**) Cross-validation error showing the likely appropriate number of populations K to be 11 (**B**) Population structure based on genetic admixture for K = 2, 7–13 and 20; (**C**) Principle component analysis (PCA) plot of the first two principal components (PCs); (**D**) PCA plot of the second and third PCs. Percentages in brackets indicate the variance explained by the PCs; (**E**) Neighbor-joining phylogenetic tree with branch length displayed; (**F**) Topological view of the neighbor-joining phylogenetic tree. Individual names in phylogenetic trees: the first letter refers to the morphotypes (O = Oil, F = Fiber, U = Unknown), the next three letters indicate the country of origin, the next letter indicates the breeding status (C = cultivar, L = landrace, B = breeding material, U = unknown), followed by the accession name where accessions prefixed by CN corresponds to the accession codes at Plant Gene Resources of Canada (PGRC).*

Table 2.1 *Pairwise population differentiation F_{ST} based on Weir and Cockerham (1984)*

Population	TEM1	FIB2	FIB1	TEM2	OWAD	NWAD	WEUR	CANC	ABYS	SAPK	SAIN	MEDT
TEM1		(0.049-0.073)[2]	(0.111-0.190)	(0.244-0.388)	(0.061-0.112)	(0.109-0.213)	(0.214-0.329)	(0.11-0.166)	NA	(0.471-0.609)	(0.545-0.639)	NA
FIB2	0.0606[1]		(0.113-0.192)	(0.203-0.359)	(0.056-0.100)	(0.104-0.192)	(0.199-0.294)	(0.107-0.164)	NA	(0.454-0.609)	(0.514-0.6327)	NA
FIB1	0.1449	0.1495		(0.271-0.415)	(0.098-0.151)	(0.127-0.202)	(0.221-0.309)	(0.19-0.257)	NA	(0.504-0.653)	(0.538-0.648)	NA
TEM2	0.3034	0.2731	0.3348		(0.119-0.212)	(0.138-0.231)	(0.160-0.272)	(0.223-0.369)	NA	(0.476-0.663)	(0.508-0.650)	NA
OWAD	0.0827	0.0774	0.1237	0.1578		(0.035-0.106)	(0.063-0.143)	(0.103-0.143)	NA	(0.297-0.425)	(0.382-0.479)	NA
NWAD	0.1613	0.1446	0.1634	0.1795	0.0684		(0.044-0.143)	(0.097-0.174)	NA	(0.299-0.449)	(0.386-0.501)	NA
WEUR	0.2603	0.2386	0.259	0.2159	0.0948	0.0849		(0.197-0.281)	NA	(0.33-0.469)	(0.383-0.522)	NA
CANC	0.1382	0.1314	0.2196	0.2916	0.1204	0.1308	0.2297		NA	(0.457-0.594)	(0.513-0.610)	NA
ABYS	0.6479	0.6472	0.701	0.6976	0.4209	0.4399	0.4405	0.6241		NA	NA	NA
SAPK	0.5406	0.5374	0.5756	0.5622	0.3614	0.3759	0.4008	0.5219	0.729		NA	NA
SAIN	0.5903	0.57	0.5919	0.5788	0.428	0.4475	0.4494	0.5593	0.6915	0.2985		
MEDT	0.7622	0.7333	0.7578	0.6852	0.5319	0.4973	0.4834	0.6983	0.8402	0.7252	0.7131	

[1] Values below the diagonal are F_{ST}, s[2] Values in parentheses above the diagonal represent the 95% confidence interval (CI), the shaded area shows populations under TEMP group.

Table 2.2 *Pairwise population differentiation calculated for the high FST SNP data set*

Population	ABYS	MEDT	SAPK	SAIN	CANC	WEUR	FIB2	FIB1	NWAD	OWAD	TEM2	TEM1
ABYS		0.012	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MEDT	0.952		0.005	0.001	0.001	0.006	0.001	0.001	0.002	0.001	0.004	0.001
SAPK	0.705	0.580		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
SAIN	0.837	0.769	0.291		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
CANC	0.981	0.987	0.863	0.913		0.001	0.002	0.125	0.001	0.025	0.082	0.246
WEUR	0.409	0.625	0.678	0.812	0.212		0.001	0.001	0.110	0.003	0.069	0.001
FIB2	0.952	0.977	0.872	0.917	0.016	0.218		0.410	0.001	0.010	0.254	0.279
FIB1	0.986	0.987	0.838	0.900	0.016	0.173	0.003		0.001	0.085	0.141	0.233
NWAD	0.515	0.729	0.742	0.846	0.066	0.027	0.069	0.048		0.181	0.246	0.005
OWAD	0.652	0.834	0.831	0.891	0.027	0.098	0.032	0.018	0.009		0.285	0.001
TEM2	0.954	0.961	0.750	0.862	0.065	0.068	0.010	0.066	0.000	0.000		0.218
TEM1	0.891	0.958	0.943	0.960	0.000	0.352	0.002	0.000	0.121	0.056	0.000	

PhiPT values below diagonal; *P* values shown above diagonal. The shaded part shows the variation among populations within TEMP

2.4.4. Distribution of haplotypes

The total number of haplotypes per chromosome derived from the high F_{ST} SNP data set ranged from 12 on chromosome 4 to 98 on chromosome 1. The most abundant haplotype in a chromosome was shared among 67 to 87% of the 383 genotypes. Most genotypes in TEMP were confined to these haplotypes. Discordantly, these haplotypes were present at very low frequency in SA (Table 2.3) which sheltered the highest number of private (unique) haplotypes and, where the two sub-groups together accounted for more than 50% of the average private haplotypes observed on all chromosomes (Figure 2.2). MEDT and ABYS included 22 and 17% of the private haplotypes, respectively. The largest TEMP populations harbored only 10% of them. Within the TEMP group, WEUR contributed the highest percentage of private haplotypes whereas none were observed in CANC, FIB1 and TEM2.

Table 2.3 *Percent frequency of the most common haplotype and total number of all haplotypes at each chromosome.*

Chromosome	SAIN	SAPK	MEDT	ABYS	WEUR	NWAD	OWAD	TEM1	TEM2	FIB2	CANC	FIB1	Overall (%)	Total
Chr1	0.0	0.0	0.0	0.0	7.1	47.8	66.7	78.0	90.0	89.2	100.0	100.0	67.1	98
Chr2	0.0	0.0	0.0	0.0	78.6	78.3	75.6	84.7	90.0	97.3	93.8	100.0	75.5	65
Chr3	0.0	0.0	0.0	100.0	71.4	69.6	84.4	88.7	100.0	94.6	100.0	100.0	79.6	51
Chr4	0.0	15.4	0.0	100.0	71.4	95.7	91.1	99.3	100.0	100.0	100.0	100.0	87.2	12
Chr5	0.0	0.0	100.0	100.0	78.6	87.0	88.9	98.0	100.0	100.0	96.9	100.0	86.2	25
Chr6	4.0	0.0	66.7	100.0	71.4	82.6	82.2	94.0	100.0	78.4	100.0	95.8	81.2	20
Chr7	0.0	0.0	0.0	100.0	50.0	87.0	73.3	93.3	90.0	100.0	90.6	95.8	79.6	46
Chr8	4.0	0.0	0.0	100.0	92.9	87.0	97.8	99.3	100.0	100.0	96.9	100.0	87.7	15
Chr9	0.0	7.7	0.0	0.0	57.1	69.6	66.7	92.0	80.0	97.3	100.0	87.5	75.7	45
Chr10	4.0	0.0	0.0	0.0	78.6	87.0	64.4	94.7	90.0	91.9	100.0	100.0	78.9	44
Chr11	0.0	7.7	0.0	0.0	64.3	78.3	80.0	96.0	90.0	100.0	100.0	100.0	80.9	38
Chr12	4.0	0.0	0.0	100.0	71.4	87.0	84.4	97.3	90.0	100.0	96.9	100.0	84.3	35
Chr13	0.0	0.0	0.0	0.0	35.7	69.6	77.8	88.0	90.0	100.0	100.0	100.0	75.7	59
Chr14	0.0	0.0	0.0	14.3	64.3	69.6	80.0	94.0	100.0	100.0	100.0	100.0	79.9	25
Chr15	4.0	0.0	33.3	100.0	78.6	91.3	84.4	96.7	100.0	100.0	100.0	100.0	85.4	18
Average	**1.3**	**2.1**	**13.3**	**54.3**	**64.8**	**79.1**	**79.9**	**92.9**	**94.0**	**96.6**	**98.3**	**98.6**	**80.3**	**37**

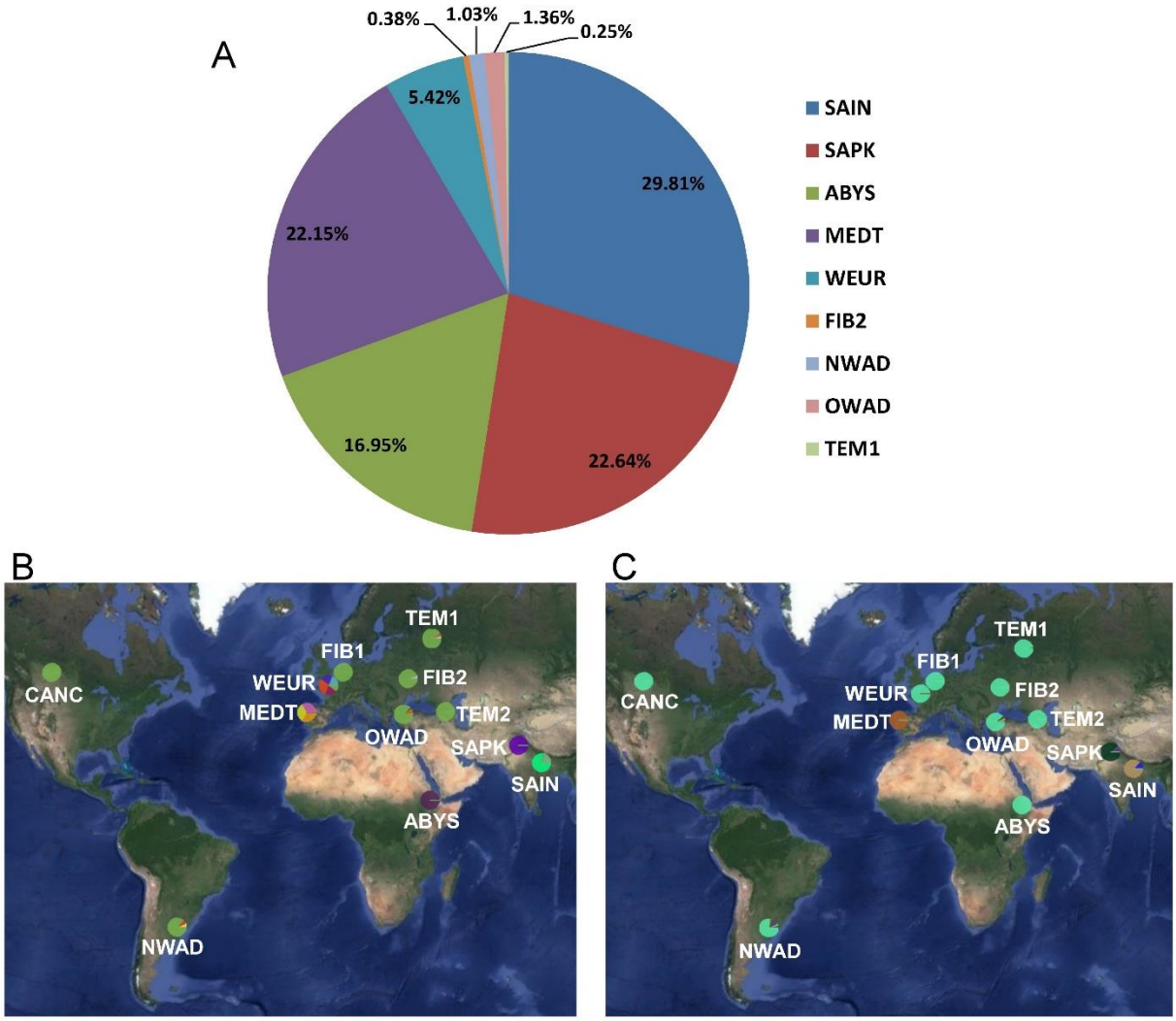


Figure 2.2 *Haplotype distribution by population.*
*(**A**) Percentage of private haplotypes observed in at least two individuals and summarized for all chromosomes based other rate per chromosome and individual in each population; (**B**) Haplotype distribution on chromosome 1; in MEDT haplotypes occurred only in single individual and each appeared unique; (**C**) Haplotypes distribution on chromosome 3; pie chart colors represent haplotypes*

2.4.5. Genetic and eco-geographic variations

The representative flax-growing regions span a wide range of eco-geography from warm South Asia to cool temperate Eurasia. A noticeable annual precipitation difference was observed between SAIN and SAPK (Appendix I-4A) despite their geographic proximity with nearly the

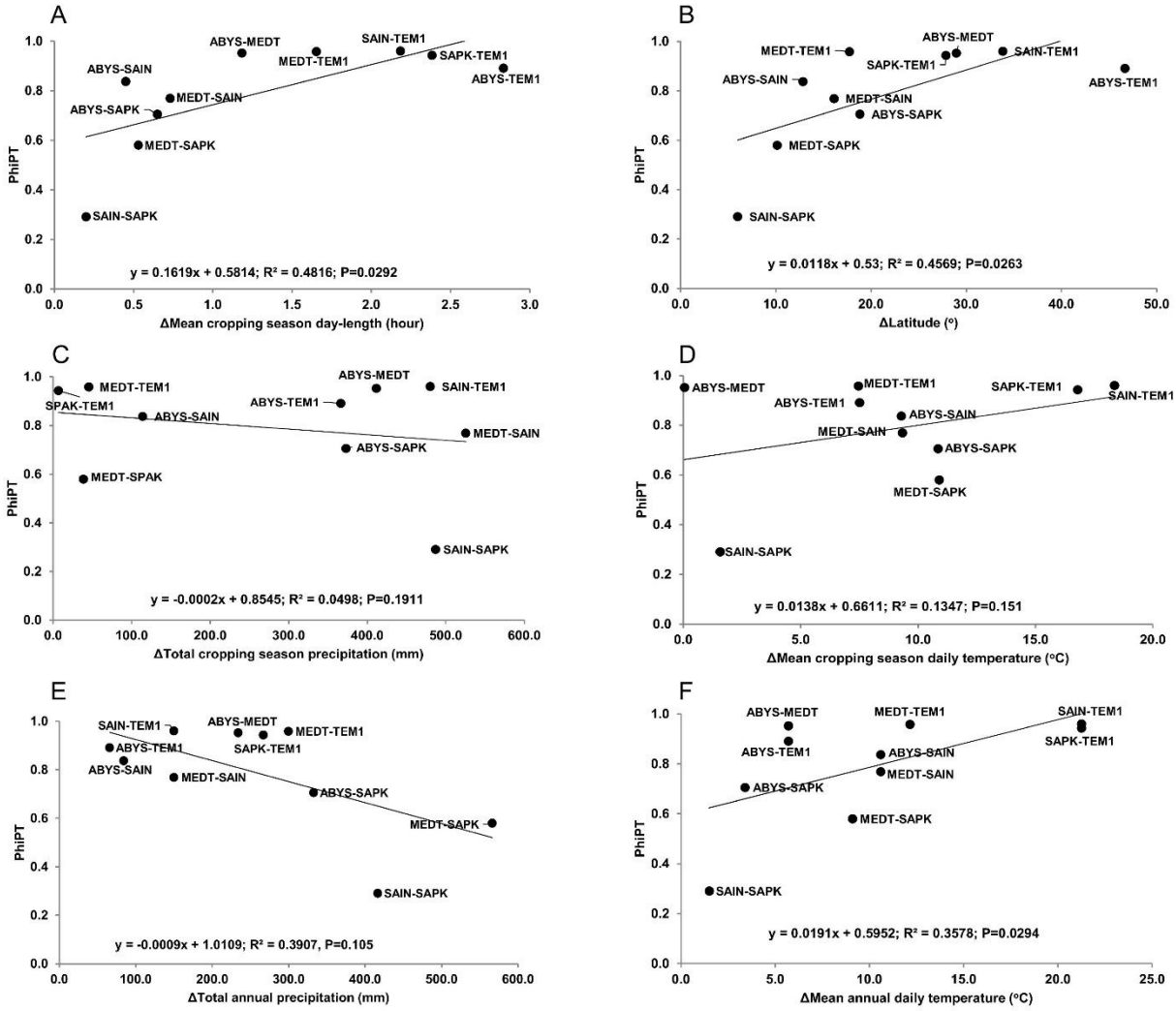


Figure 2.3 *Mantel spatial auto-correlation between the pairwise genetic differentiation based on high F_{ST} SNPs and differences in site parameters.*
*(**A**) cropping season total precipitation; (**B**) cropping season average daily temperature; (**C**) cropping season average day length; (**D**) latitude; (**E**) total annual precipitation; (F) mean annual daily temperature. TEM1 represents all the populations in group TEMP.*

same annual temperature (Appendix I-4B). The Mantel test revealed significant auto-correlation of spatial variation in cropping season day length ($P < 0.0292$; $R^2 = 0.4816$) and latitude ($P = 0.0263$; $R^2 = 0.4569$) with pairwise genetic-differentiations (Figure 2.3A, B). While spatial variation in annual precipitation, cropping-season precipitation and temperature did not correlate to pairwise genetic differentiation (Figure 2.3C-E), significant ($P = 0.0294$, $R^2 = 0.3578$)

correlation was computed between spatial variation in annual average temperature and the genetic variation (Figure 2.3F).

2.4.6. Genome-environment association and adaptation signatures

Since the genotypes selected for GEA were those with high ancestral proportion in their respective population, they fell into five clear clusters (Appendix I-5A, C-D) which is also illustrated in the suggested number of populations (K) (Appendix I-5B). For the 11018 SNPs

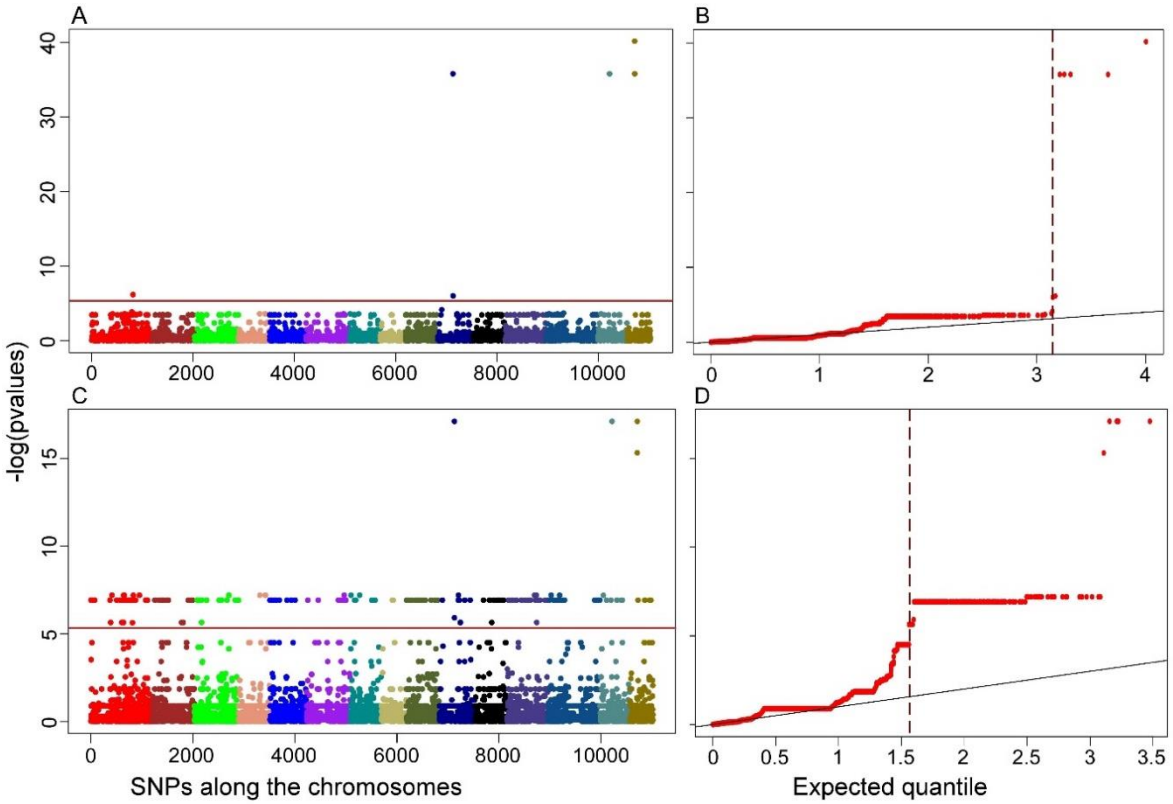


Figure 2.4 *Manhattan and QQ plots for genome-environment association for day length and latitude.*

*(**A**) Manhattan plot for Genome-day length association; (**B**) QQ plot for Genome-day length association; (**C**) Manhattan plot for Genome-latitude association; (**D**) QQ plot for Genome-latitude association. The horizontal brown line in Manhattan plots indicates the threshold p-value. Each color in Manhattan plot represent a chromosome where chromosomes are in their order from left to right. P-values to the right of the broken vertical line in QQ plots show the outliers below the threshold.*

used for GEA, a *p* value threshold $4.54X10^{-06}$ was obtained at α=0.05. The total number of loci significantly associated ($P < 4.54X10^{-06}$) for each environmental factor ranged from seven for

day length during the cropping season to 271 for latitude (Appendix I-6). Five loci, one from chromosome 10 and two each for chromosomes 14 and 15, had the lowest *P* values for association with both latitude and day length (Figure 2.4A-D). These loci were also significantly associated with both mean cropping season and annual temperature where they still docked the outlying position for mean annual temperature (Appendix I-7). However, none of them were associated with precipitation variables. These loci accounted for relatively high differentiation among all populations (n=12) in 51K SNP dataset following the different algorithms employed (Appendix I-8). Chromosomes 1, 5 and 13 had a particularly high number of loci associated with temperature variables (Appendix I-6, Appendix I-7A-D). Loci strongly association with precipitation variables were scattered throughout the genome, although a higher concentration was observed on chromosomes 1, 9 and 11 (Appendix I-7A-B).

2.5. Discussion

The population structure of the flax core collection reflects the geo-ecological and historical anthropogenic influences on the crop. The adaptive genetic signatures imply the importance of an eco-geographic gradient in shaping the genetic structure of flax. Population structure also revealed the dominance of closely-related temperate-origin genotypes in the current flax core collection. Whereas, a high concentration of haplotype endemism was observed in genotypes from the flax centers of diversity regions, especially South Asia, implying the importance of these regions as flax genetic reservoir (Vavilov, 1951). The long branch length for populations from the crop center of diversity regions (Figure 2.1E) also suggest a high rate of substitution and consequently high diversity given that neighbor joining is one of the efficient methods to estimate evolutionary rate (Rusinko and McPartlon 2017). Addition or substitution with accessions from these regions would be beneficial to capture the available genetic diversity in breeding and conservation schemes of the crop.

2.5.1. Flax population structure and eco-geographic resilience

The flax population structure can be associated in large part to the eco-geographic heterogeneity among the major growing regions of the crop that span the warm Indian subcontinent, Mediterranean, Eurasian temperate and East African highlands (Casa et al. 1999; Rossini and Casa 2003). The appreciable auto-correlation between genetic differentiation and day length stresses the importance of photoperiodism in shaping the genetic structure of this crop. Day length is one of the vital environmental factors that determine adaptation of crop varieties to different eco-geographic locations (Hung et al. 2012; Itoh et al. 2010). Recently, Gutaker et al. (2019) reported allelic variants associated with flowering time and adaptation to high latitudes in flax that were possibly acquired via post domestication gene flow from Eurasian wild flax. On the other hand, the lack of auto-correlation between genetic differentiation and spatial variations in temperature and precipitation during the cropping period could plausibly be attributed to the complex edaphic and atmospheric conditions limiting their impacts (Moro et al. 2015). The significant auto-correlation of genetic differentiation to the spatial variation in annual mean temperature might reflect a carry-over effect of winter conditions, especially on soil temperatures (Zhang 2005), during the early cropping season which can correlate with the genetic structure among populations (Robinson et al. 2018). The substantially higher correlation of latitude to the genetic variation can be associated with differential environmental factors including day length and temperature along the latitude (Foll and Gaggiotti 2006).

The deviation of the MEDT population, especially from the nearby Old World TEMP populations can be attributed to the unique ecological and vegetation patterns of the Mediterranean habitats (Stella et al. 2013). The Mediterranean zones are characterized by heterogeneous environmental conditions sheltering rich biodiversity and endemism (Medail and Quezel 1997). Flax germplasm from Mediterranean regions had peculiar inflorescence with open large petal bearing large seeds (Diederichsen and Richards, 2003). Large seeds were

also observed for beans and chickpea of the region (Perrino 1988), hinting that this morphological trait might be adaptive to these unique ecological conditions (Ackerly 2004). The Mediterranean germplasm is somewhat under-represented (few individuals) in our flax core collection but the three individuals that were consistently outliers in all analyses came from that population, in line with the previously reported unique features of the germplasm from this region (Diederichsen and Richards, 2003). The Mediterranean region is one of proposed centers of diversification of flax and its most likely center of origin (Vavilov, 1951). As such, and in light of our results, introduction of additional Mediterranean germplasm would enrich the flax core collection.

The distinct partitioning of the SA group was previously noted by (Soto-Cerda et al., 2013). The large number of private haplotypes and the clear separate clustering starting from K= 2 infer the unique adaptation of the populations in this group to the warm South Asian climate. Genotypes from this region are characterized as short, and large seeded with high oil content (Diederichsen 2001), traits considered adaptive to the warm climate of the region (Rajwade et al. 2010; Zajac et al. 2012). Differences in annual precipitation between India and Pakistan may explain the distinction between the two SA populations (Appendix I-5A). The relatively close relationship between the SA and TEMP populations may be a consequence of the high genetic diversity of the SA (Allaby et al., 2005). The Indian subcontinent, which includes India and Pakistan, is also a center of diversification for flax (Vavilov, 1951).

The relative genetic peculiarity of ABYS can be allied with adaptation of flax to elevation-induced environmental factors of the Abyssinian highlands which are major producers of flax in Africa (http://www.fao.org/faostat/en/#data/QC). Abyssinian highlands that embrace the flax-growing niches are described as high altitude, cool environment with rugged terrain that create substantial agro-climatic heterogeneity (Samberg et al. 2010). These highlands harbored

noticeable plant genetic diversity and endemism in many plant species including flax (Vavilov, 1951). Unlike those of the Mediterranean and South Asian, genotypes from Abyssinia have small seeds (Diederichsen and Fu, 2006). These genotypes are described as forage type with short stem and dense leafy crown (Dillman 1953), adapted to windstorms and resistant to lodging (Vaisey-Genser and Morris, 2003). Given the proximity of the Abyssinian highlands to the equator, genotypes in this regions are likely adapted to short day conditions differing from the commonly known long day flax genotypes (Domantovich et al. 2012).

2.5.2. Genetic signatures for differential eco-geographic adaptation

Loci that showed significant association to environmental factors may represent signatures of local adaptations. The five loci that showed consistent strong association with both day length and latitude suggest differential selection of loci along the latitude that impacts day length and temperature. The two loci of chromosome 15 are part of the flax gene Lus10017823 which is orthologous to *Arabidopsis thaliana* gene AT1G01790 that encodes one of the potassium ion efflux antiporters (KEA). KEAs are reported to be important for chloroplast osmoregulation (Kunz et al. 2014) and may play an important role in abiotic stress tolerance, especially drought (Sheng et al. 2014). The target Lus10017823 locus did not show association with precipitation variables but still had significant association with temperature variables, especially with mean annual temperatures. As a mainly high latitude and altitude crop, low temperature exposures cause stresses including moisture limitation that might be an important environmental factor for flax. Hence, Lus10017823 might play a role in high latitude and altitude adaptation. KEAs might also have differential effects with regards to photoperiod sensitivity (Dana et al. 2016). The strong association of the loci in Lus10017823 with day length points to an added role in adaptation to day length gradients across flax habitats.

The strong day length and latitude associated locus on chromosome 10 is flanked by flax genes Lus10011959 and Lus10011957/6 that are orthologous to AT2G19480.1 and AT2G19480.3, respectively, and that both encode a nucleosome assembly protein1 (NAP1). NAPs are involved in DNA repair and ultra violet (UV) irradiation immunity (Liu et al. 2009) one of the important environmental factor in local adaptation (Molinier 2017). UV radiation varies with latitude and day length, strengthening the argument proposing a role for NAP1 in adaptation of flax to its eco-regions. More than 15 genes were found in the 100kb downstream of the two co-located loci on chromosome 14 but none were predicted in the 100kb upstream. The nearest downstream gene to these loci Lu10013180 is orthologous to AT1G54070.1 which encodes a dormancy-related protein that might be considered as a signature of adaptation along latitudinal and thus day length gradients.

2.5.3. Flax genetic divergence and anthropogenic influence

The population structure (Figure 2.1) and pairwise variation (Tables 2.1) patterns also reflect the historical human interventions in the different regions. The low genetic divergence among the populations within TEMP regardless of the morphotypes likely resulted from selection of elite flax genotypes having both fiber and seed merits from closely-related genetic origins (De Haan 1952; Rachinskaya et al. 2011; You et al. 2016). Most of the genotypes in the flax core collection exhibited a combination of traits characteristic to both fiber and seed morphotypes (You et al., 2017). Fu (2005) reported low genetic distance among flax accessions from Europe, the Americas, Oceania and the Mediterranean that matches the geographic range of our TEMP group (Appendix I-2) except for the unique MEDT group. The close genetic relationship of these genotypes, as illustrated by the small number of private haplotypes, in such wide geographic locations is also suggestive of gene flow via human movements following discovery of the New World s and/or recent germplasm exchanges by breeders and collectors.

Compared to the other populations in the TEMP group, the WEUR population, which is mainly dominated by West and Central Europe and Canadian genotypes, harbored more unique haplotypes (Figure 2.2). The Western European region is one of the ancient flax growing regions and, as such, it may be a sanctuary of some old genetic lineages (Herbig and Maier, 2011).

The large number of private haplotypes in the SA group suggests thousands of years of selection for oilseed type and against fiber traits (Diederichsen and Fu, 2006). The use of flax as a source of fiber is likely restricted to its early cultivation history in some parts of India because it was displaced a long time ago by cotton and other fibers; hence, there likely was a period when linseed, fiber and/or dual flax were grown in the region (Barnes 2004). The relatively lower differentiation between the SA and TEMPs populations (Table 2.1) may suggest gene flow by contemporary explorers, collectors and flax breeders.

The genetic distinctness of the ABYS group can also be related to anthropogenic influences that shaped the crop for the traditional oil uses and the environment (Engels et al. 1991). Flax had also likely been cultivated as a fiber source before cotton was introduced to the region in the first century (Gervers 1990). The Abyssinians were in cultural and geographical immediacy with the traditionally fiber flax growers of Egypt (Vaisey-Genser and Morris, 2003). Though Abyssinia was postulated to be one of the possible centers of origin and diversification of flax (Vavilov, 1951), the crop was likely introduced with other crops such as barley by early immigrants approximately 5000 years ago (Pagani et al. 2012; Phillipson 1993). The uniqueness of the MEDT population, especially compared to the Old World TEMP populations, would be intuitively in disagreement with their geographic proximity and the demographic flow between these regions, suggesting a more likely eco-geographical than anthropogenic influence. Preservation of unique genes may be intensified in this region for adaptation to its

unique environment. This region has a long tradition of flax production (Carvalho et al. 2013) as a source of both fiber and oil (Zapata et al. 2004). The small sample size of the MEDT and ABYS groups in our core collection is likely insufficient to provide a good representation of the breadth of the genetic diversity from these regions (Vavilov, 1951).

The inclusion of genotypes from China and Japan (Figure 2.1E, Appendix I-2) into the TEMP group indicates that the genotypes were likely exotic to these regions. Most were fiber types closely related to the Eurasian collections. Fiber flax in countries such as China has less than 120 years of history (Wang et al., 2018). China is among the oldest flax-growing regions of the world exclusively focused on linseed until fiber types were introduced ~100 years ago (Liu et al., 2011) where silk and hemp have long been used as their major traditional fiber sources (Liu 2010; Lu and Clarke 1995). Traditionally, five indigenous flaxseed ecotypes (Loess Plateau, Grassland, Yellow River Valley, Xinjiang and Quighai) are grown mostly in the dry and relatively high elevation areas of China (Marchenkov et al. 2003).

2.5.4. Genetic structure and its implication to breeding

The genetic structure and the haplotype distribution indicate that the flax core collection possesses a broad representation of the diversity of genotypes of temperate origin but the more diverse groups are underrepresented, particularly considering their distinctness. Although the genetic structure clearly distributed the accessions of the core collection into four major groups, more than 87% belonged to the TEMP group, of which 39% fell in the TEM1 population. Likewise, our haplotype frequency analysis using high F_{ST} data at each chromosome showed that as much as 87% of the genotypes were represented by unique haplotypes. Most private haplotypes came from the less represented groups SA, MEDT and ABYS which are all proposed as centers of genetic diversity of the crop (Vavilov, 1951). Addition of genotypes from

these regions would widen the breadth of genetic diversity to make it more representative of the species, an attribute that could be harnessed by breeding programs.

2.6. Conclusions

The genetic diversity of flax was mainly shaped by eco-geographic factors and use-directed selections within the different flax growing regions of the world. Day length is one of the ecological factors that significantly impacted genetic variations among the different flax growing regions. The majority of the flax genotypes in the temperate zone that extends from Europe including the Russian Federation to the New World such as the Americas and Oceania are genetically closely related. The core flax collection comprised a total of 12 populations that are pooled in four major groups: Temperate, Mediterranean, South Asian and Abyssinian. Regions that are postulated to have been centers of origin of the crop such as Mediterranean, South Asian and Abyssinian harbor more unique haplotypes than the temperate pool. More than 87% of the core collection genotypes belong to the temperate group that encompasses both fiber and oilseed types subdivided into eight populations, and where variation between seed and fiber morphotypes is less pronounced than variation within the morphotype. The remaining fraction of the core collection genotypes distributed into the other three groups are all oilseed type. However, there is significant variation among these groups and populations within the groups. Domination of the closely related genotypes of the TEMP group should be balanced with additional genotypes from the other three regions to achieve higher diversity for improvement of the crop through breeding and to dependably secure the conservation endeavors.

3. The complex genetic architecture of early root and shoot traits in flax revealed by genome-wide association analyses

Demissew Sertse[1,2], Frank M. You[1], Sridhar Ravichandran[1], Sylvie Cloutier[1,2,*]

[1]Ottawa Research and Development Center, Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, ON, Canada

[2]Department of Biology, University of Ottawa, 30 Marie Curie, Ottawa ON Canada

This chapter has been published in Frontiers in Plant Sciences 10 (2019): 1483

doi.org/10.3389/fpls.2019.01483

Authors' contribution

DS and SC conceived the experiment. DS, SC and SR participated in designing and implementing the phenotyping experiment. SC and FY produced the genetic data. DS performed the phenotypic and GWAS data analyses. DS prepared the manuscript. SC, FY and SR corrected and edited the manuscript. All authors read and approved the final version.

3.1. Abstract

Roots are fundamental organs for water and nutrient uptake as well as for signal transduction in response to biotic and abiotic stresses. Flax has a shallow tap root system that relies mostly on top-soil nutrient and moisture resources. The crop can easily be outcompeted by weeds or other crops in intercropping systems, especially in moisture deficit conditions. However, there is a wide range of variation among genotypes in terms of performance under scarce resources such as moisture limitation. Here we phenotyped two shoot and 14 root traits on 115 flax accessions grown in a hydroponic pouch system and performed a genome-wide association study (GWAS) based on seven different models to identify quantitative trait loci underlying these traits. Significant variation among genotypes was observed for the two shoot and 12 of the 14 root traits. Shoot dry weight was correlated with root network volume, length, surface area and root dry weight *(r* > 0.5, *P* < 0.001) but not significantly correlated with root depth (*r* = 0.033, *P* > 0.05). The seven GWAS models detected a total of 228 quantitative trait nucleotides (QTNs) for the 16 traits. Most loci, defined by an interval of 100 kb up- and downstream of the QTNs where LD falls below r^2=0.1 on average (Figure 1.5), harbored genes known to play role(s) in root and shoot development, suggesting them as candidates. Examples of candidate genes linked to root network QTNs included genes encoding GRAS transcription factors, mitogen-activated protein kinases and auxin-related lateral organ boundary proteins while QTN loci for shoot dry weight harbored genes involved in photo-morphogenesis and plant immunity. These results provide insights into the genetic bases of early shoot and root development traits in flax that could be capitalized upon to improve its root architecture, particularly in view of better withstanding water limiting conditions during the cropping season.

Keywords: flax, root, shoot, genome-wide association study (GWAS), quantitative trait nucleotides (QTNs), candidate genes

3.2. Introduction

Roots are vital organs in terrestrial higher plants for acquisition of essential nutrients and water. Because roots function in a bio-physico-chemically dynamic rhizosphere, they play an important role in controlling and regulating the impacts of various edaphic factors through internal physiological adjustments (Hodge 2004; Jackson et al. 1990) and signal transduction (Batool et al. 2018). Adaptation of plants to a scarcity of resources and associated edaphic factors is therefore governed by their root system. The architectural features of root systems are crucial in efficiently tapping the available resources such as water and nutrients in the rhizosphere (Yue et al. 2006) and inadequate development of the root system may lead to significant yield losses in water-limiting conditions (Henry et al. 2011). Understanding root traits and resource use efficiencies of the root system is key to yield improvement in crops (Kell 2011). However, the inaccessibility of the rhizosphere has made studies of root traits challenging; hence, these traits have been scantly considered in varietal improvement.

Flax, one of the founder crops of agriculture (Weiss and Zohary 2011), has been grown as both fiber and oilseed crops for nearly the entire span of its cultivation history (Herbig and Maier 2011). The crop is adapted to diverse ecologies, from the warm Indian subcontinent to the cool temperate areas in Eurasia (Casa et al. 1999; Sertse et al. 2019b). A wide range of uses are derived from its stem fibers and its oil-rich seeds (Singh et al. 2011b). Flax production, however, is constrained by low yield (Wittkop et al. 2009). The meagre improvement in flax yield of ~0.5 ton/ha since the 1960s, obtained through breeding and agronomic practices, has not been sufficient to impact its production in a major way and production continues to decline as growers are shifting to higher yielding crops (http://www.fao.org/faostat/en/#data). High and stable yielding cultivars are urgently needed to rekindle growers' interest and meet market potential.

The complex genetic architecture of roots in flax

Flax, like other oilseed crops, is a tap-rooted plant. Compared to canola, sunflower and safflower, flax has a shallower root system and, as such, it mainly relies on moisture and nutrient resources available in the soil's top layers (Hocking et al. 1997; Kar et al. 2007). Under limited water, flax can be easily outcompeted by many weeds, exacerbating the competition for resources (Alessi and Power 1970; Gruenhagen and Nalewaja 1969; Klimek-Kopyra et al. 2015). Flax is also less competitive than fibrous rooted cereals such as wheat (Morillon and Lassalles 2002). However, a wide range of performance among flax genotypes under different moisture regimes (Diederichsen et al. 2006; Foster et al. 1998) can be attributed to variations of their root system (Cattivelli et al. 2008).

Prompted by the current large sets of genetic data from genome-wide association studies (GWAS) and advancements in imaging and data processing, high-throughput digital root phenotyping techniques have become attractive. Hund et al. (2009) used a pouch system to phenotype root traits in maize and, this system has subsequently been applied in other crops such as wheat (Atkinson et al. 2015), Brassica (Thomas et al. 2016) and barley (Canto et al. 2018) for examples.

Here we applied this technique to study the early root and shoot development of a flax mini-core collection (n=115) that comprises representative genotypes from all major flax-growing regions of the world. A GWAS was performed using a set of single nucleotide polymorphism (SNP) markers obtained from shotgun short-read re-sequencing data of the germplasm collection. The objectives of this research were 1) to assess the extent of the variation in root traits among genotypes, 2) to identify quantitative trait nucleotides (QTNs) associated with the genetic architecture of selected root and shoot traits, and, 3) to identify candidate genes for the traits harbored at the QTN loci.

3.3. Materials and methods

3.3.1. Plant materials

A flax mini-core collection (n = 115) that comprised >95% of the genetic diversity (Soto-Cerda et al. 2013) of the flax core collection (n = 407) (Diederichsen et al. 2013) was used. The 75 linseed, 33 fiber flax and seven accessions of unknown type of the mini-core collection were collected from 20 countries that represent all major flax-growing regions of the world.

3.3.2. Phenotyping

Early development root and shoot phenotyping was performed in a hydroponic pouch system modified from Hund et al. (2009). The system comprised two large plastic bins each containing 108 L of nutrient solution that was transferred every 3 h from one bin to the other using timer-operated peristaltic pumps. Aluminum frames mounted on top of the two opaque bins were used to hang the pouches. The bins, but not the plants, were covered with black polyethylene sheets to prevent algal growth. This system was installed in a growth chamber (Conviron PGC20, Serial No. 150342, Controlled environment Ltd, Canada) maintained at 21/18^{o}C with a 16-h day/8-h night photoperiod.

Sterile 24x30 cm blue germination blotting papers (SGB1924B, Anchor Paper Company, St Paul, MN, USA) were inserted in Ziploc bags. The assemblies were attached to rust-proofed rods using fold back clips that were suspended in the bins with the open side of the Ziploc bags down to allow blotting up of the nutrient solution. Assemblies were labeled and randomized for each of the three consecutive biological replicates performed (Figure 3.1A).

The Hoagland nutrient solution (HOP1, Hoagland's No. 2 Basal Salt Mixture, Caisson Labs, Smithfield, UT, USA) was made at 25% strength with deionized water and adjusted to pH ~6.3. The pouches were randomly distributed in the two bins. After complete moistening of the

blotting paper, three seeds were directly put on slight scratch in the paper using tweezers, ~2cm below the top edge and slightly spaced out near the center.

Figure 3.1 *Early root phenotyping experiment set-up and representative root images. (**A**) Close-up of the experimental set-up showing a partial view of the upper section of the blotting papers, labels and plantlets, (**B**) extensive root system of TMP-2530 (U_MAR_C_CN98193) and (**C**) root system of flax cultivar Hanley.*

The siphons that transferred the nutrient solution were set ~3 cm from the bottom of the bins, i.e. touching the bottom of the blotting papers, in order to keep them moist at all time. Five days after seeding, germinated seeds were thinned out to one seedling per pouch. The retained seedlings were allowed to grow for 18 days.

On the 19th day, plants were carefully removed from their assemblies for imaging and measurements. Shoots, cut at the shoot and root junction, were measured with a ruler to estimate shoot length prior to being transferred to labeled envelopes and dried for three days at 60°C to measure dry weight. The root system of each plant was imaged using a Canon camera (EOS Rebel T5i) mounted on a custom stand to ensure a consistent 40 cm distance between lens and roots. After imaging, the roots were gently pealed from the paper and processed as the shoots to measure their dry weight. Three biological replicates were thus performed consecutively.

Root images were processed using the General Image Analysis of Roots (GIA Roots) software (Galkovskyi et al. 2012). Each image was scaled using the 30 cm edge of the paper as reference. Scaled images were then analyzed for 13 root traits (Table 3.1).

Shoot dry weight, shoot length, root dry weight and the 13 root traits measured by the GIA software were analyzed. Basic statistics of the 16 traits were computed for 111 genotypes (Table 3.1). Four genotypes with a single replicate were not included in the analysis. A one-way analysis of variance (ANOVA) was performed for each trait using R. Pearson pairwise correlation coefficients between traits were calculated and summarized using R package sjPlot (Lüdecke 2017).

Table 3.1 *Phenotypic traits and their basic statistics of the 111 accessions of the mini-core collection*

Trait	Abbreviation	Description	Unit[1]	Range	Median	Mean ± SD[2]
Average root diameter	ARD	Average individual root diameter	cm	0.01 - 0.04	0.03	0.035 ± 0.0001**
Maximum number of roots	MaxR	After sorting the number of roots crossing a horizontal line from smallest to largest, the maximum number is considered to be the 84th-percentile value (one standard deviation)	count	2 - 19	8.33	8.94 ± 3.37**
Median number of roots	MedR	The result of a vertical line sweep in which the number of roots that crossed a horizontal line was estimated, and then the median of all values for the extent of the network was calculated.	count	1 - 10	3.67	3.99 ± 1.65**
Network area	NWA	Total area covered by all roots	cm^2	1.92 - 14.67	7.31	7.27 ± 2.15**
Network depth	NWDep	The maximum depth reached by the root	cm	10.95 - 24.91	20.29	20.04 ± 2.51
Network distribution	NWDist	The fraction of root network in the lower two third of the network (analogy of root depth density)	na	0.26 - 1.98	0.79	0.865 ± 0.331
Network length	NWL	Total length of the entire network (~half perimeter)	cm	156.2- 513.7	255.9	245.3 ± 77.5**
Network perimeter	NWPeri	Total length following all root surfaces	cm	111.7 - 1061.7	539.1	510.2 ± 163.8**
Network surface area	NWSA	The sum of surface area of all roots in the network	cm^2	7.03 - 53.93	26.79	26.49 ± 7.94**
Network volume	NWV	The total volume of all roots in the network	cm^3	0.08 - 0.48	0.25	0.252 ± 0.074*
Network width	NWW	The maximum linear width attained by the root	cm	4.17 - 21.33	12.42	12.43 ± 3.12**
Network width to depth	NWW_Dep	Ratio of NWW to NWDep (NWW/NWDep)	na	0.203 - 1.316	0.64	0.632 ± 0.159*
Root dry weight	RDWt	Oven dried root weight	g	0.008 - 0.045	0.02	0.024 ± 0.006
Shoot dry weight	SDWt	Oved dried shoot weight	g	0.015 - 0.064	0.03	0.028 ± 0.008*
Shoot length	SL	Length of shoot from root collar to tip of the shoot	cm	4.95 - 13.27	7.67	7.91 ± 1.30*

3.3.3. Genotyping and data quality control

SNP data of the entire flax core collection (n = 407) was previously generated after resequencing each genotype using an Illumina HiSeq 2000 platform in 100bp paired-end mode to achieve an average coverage of 17X. The alignment, SNP calling and quality control such as removal of SNPs in long terminal repeat regions were performed as previously described (Sertse et al. 2019b) . The SNP data for the mini-core collection was extracted from this core collection dataset.

For this study, only SNPs with no missing data were used. Because the minor allele frequency (MAF) of a SNP could differ between the core and the mini-core collections, we used two datasets for GWAS. The first dataset (7K) included all SNPs with no missing data regardless of their MAF in the mini-core collection because these had already previously been included based on MAF>5% criteria in the core collection. The second dataset was smaller (3K) because SNPs with MAF<5% in the mini-core collection *per se* were removed. The two datasets were analyzed separately and results were compared.

3.3.4. Genetic structure and variation analysis

To estimate the possible number of ancestral populations (K), a cross-validation technique (Alexander and Lange 2011) was applied. Analysis of the ancestral proportion of each genotype (Pritchard et al. 2000) was performed for K values ranging from 2 to 20 using sparse non-negative matrix factorization sNMF (Frichot et al. 2014) of the R package LEA (Frichot and François 2015) with default parameters except that the number of runs was increased from ten to 20. The K value producing the lowest cross-validation error was accepted as the number of ancestral populations. The same package was used to visualize the cross-validation and to generate the structure plot. Neighbor-joining (NJ) phylogenetic and principal component (PC) analyses were performed using TASSEL v 5.2 (Bradbury et al. 2007). Results of the two

analyses were summarized using Tree of Life (iTOL) v3 (Letunic and Bork 2016) and R, respectively. Genotypes were assigned to the suggested ancestral populations based on their Q-matrix. Populations were named based on the passport data of their members indicating geography of origin.

3.3.5. Phenotype-genotype association analyses and mapping

To assess the genetic variants underlying each root and shoot trait, the multi-locus GWAS methods FASTmrEMMA (Wen et al. 2016), FASTmrMLM (Zhang and Tamba 2018), ISIS EM-BLASSO (Tamba et al. 2017) , mrMLM (Wang et al. 2016a), pKWmEB (Ren et al. 2018) and pLARmEB (Zhang et al. 2017) included in the R package multi-locus random-SNP-effect mixed linear model (mrMLM) (Wen et al. 2017) were applied. The single locus genome scan method latent factor mixed linear model (LFMM) in the R package lfmm (Frichot et al. 2013) was also used. To control the type I error in multiple comparison, false discovery rate (FDR) correction at = 0.05 was applied (Benjamini and Hochberg 1995) for all models to identify significant quantitative trait nucleotides (QTNs). For LFMM, Bonferonni correction factor at α = 0.05 (0.05/n, n = the total number of SNPs) was also used as a comparative methods. Quantitative trait loci (QTL) regions spanning 100 Kb up- and downstream of all associated QTNs were examined for the predicted coding genes they harbored using the flax reference genome (You et al. 2018b). Predicted functions of genes identified within each QTL were investigated based on their Arabidopsis orthologues (www.arabidopsis.org). Strongly-associated SNPs and their putative underlying genes were illustrated on the flax pseudomolecules (You et al. 2018b) using MapChart 2.3 (Voorrips 2002). The favorability of alleles at QTNs detected by at least two of the multi-locus models with phenotypic variance explained (PVE) (R^2 > 5%), was illustrated using box plot based on mean phenotypic value of genotypes with each allele.

For genes previously identified to encode interacting proteins, protein interaction networks were constructed using the tool STRING V11 (https://string-db.org) (Szklarczyk et al. 2018). The interaction networks were constructed based on protein matchings searches in flax (*Linum usitatissimum*) with a minimum regulatory confidence of 0.95.

3.4. Results

3.4.1. Phenotypic variation

From the two shoot- and 14 root-traits, significant variation (P<0.05) among genotypes was detected for all except for root dry weight, network depth and network distribution (Table 3.1). In multiple comparisons, genotype TMP-2530 (U_MAR_C_CN98193) significantly outperformed at least one genotype for 10 of the 13 traits (Appendix II-1). TMP-2530 was also the only genotype that had significantly higher shoot dry weight and it displayed a distinctive heavy root network (Figure 3.1B) compared to other genotypes such as Hanley for example (Figure 3.1C).

Several traits were significantly correlated (Appendix II-2). Shoot length and dry weight were strongly correlated with root dry weight and root network volume; the latter being highly correlated with root and shoot dry weights with r values of 0.78 and 0.58, respectively.

3.4.2. SNP datasets and genetic structure

A total of 7,707 SNPs with MAF>5% in the core collection had no missing data and, 3,243 of these had a MAF > 5% in the mini-core *per se*. These datasets are henceforth referred to as the 7K and 3K datasets, respectively. Phenotypic data was incomplete for 4 genotypes and additional 10 genotypes had poor SNP calls; hence the following analyses were carried out on 101 accessions of the mini-core collection using the 7K and 3K datasets independently. Analysis of population structure analysis based on both SNP datasets clustered the genotypes into six ancestral populations (Figure 3.2A) grouped as follows: Canadian cultivars (CANC), Canadian Russian (CA_RU), Temperate (TEMP), Asian (ASIA), Admixture (ADM) and

mini_Indian (MIND) with only two accessions (Figure 3.2B). Principal component analysis (PCA) with PC1/PC2 and PC1/PC3 produced similar population structure patterns (Figures 3.2C-D).

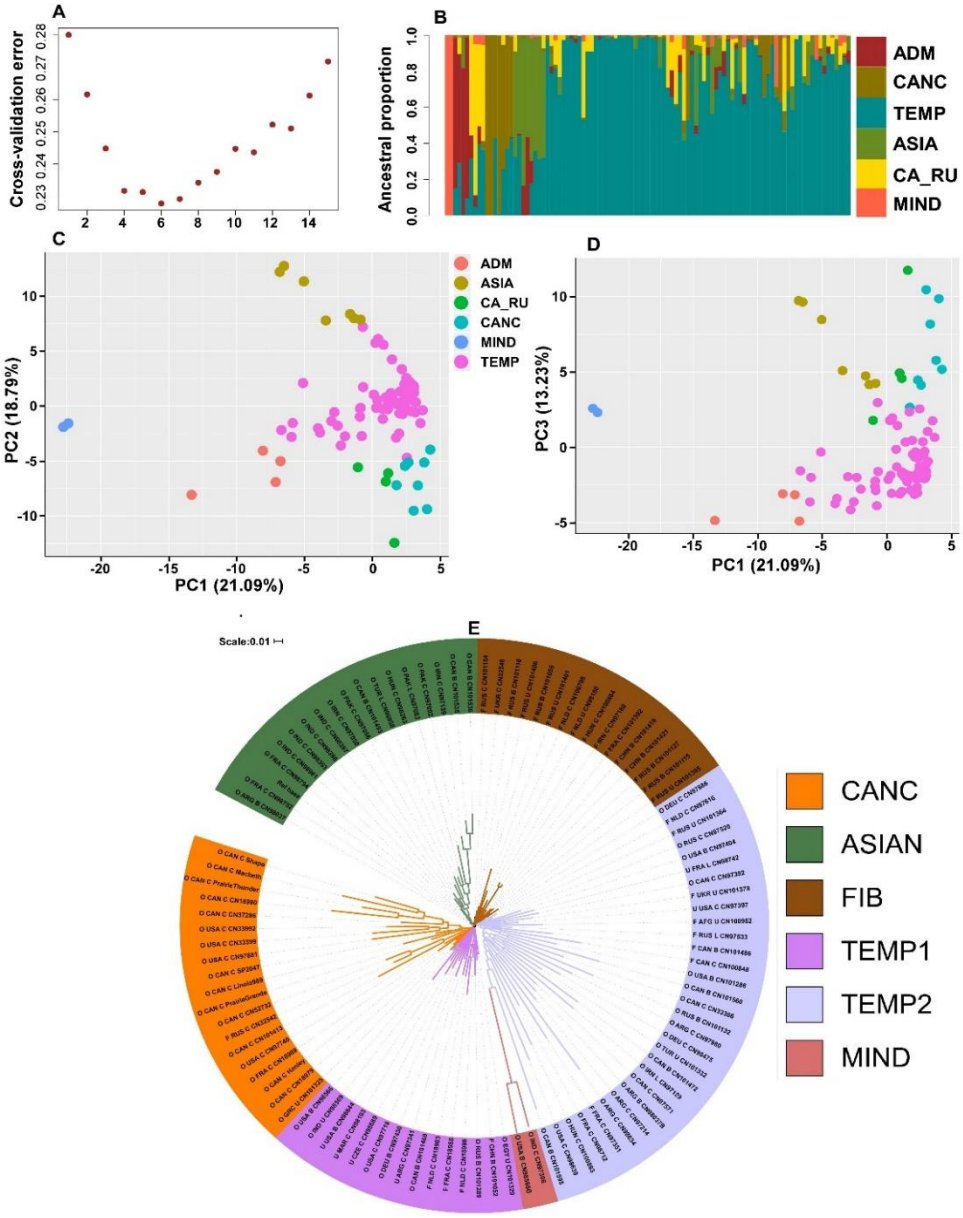


Figure 3.2 *Population genetic structure of minicore.*

*(**A**) Estimates of the number of ancestral populations indicating six as the best fit, (**B**) population structure plot showing the six populations in different colors, (**C**) principal component analysis (PCA) plot of the first two principal components (PCs) where the percentages in parentheses represent the variance explained by the PCs, (**D**) PCA plot of the first and third PCs, (**E**) neighbor-joining (NJ) phylogenetic tree where naming convention indicates the type (O = oil, F = fiber, U = unknown), the country of origin, the breeding status*

The NJ clustering slightly deviated from this pattern by assigning more than half of the fiber-type accessions to a clade (FIB), splitting the TEMP into two clades (TEMP1 and TEMP2), distributing the CAN_RU, ADM and MIND into the TEMP clades but with distinguishable long branch lengths of MIND. CANC and ASIA were each in a separate clade for a total of five clades and six distinguishable populations (Figure 3.2E).

3.4.3. QTN-trait association

From the two datasets, the six multi-locus (mrMLMs) and the LFMM methods identified a grand total of 228 QTNs associated with at least one of the 16 traits, of which 33 QTNs were detected in both datasets (Appendix II-3). A total of 35 large-effect QTNs (high PVE) with $R^2 > 5\%$ for at least one trait were discovered, of which, 15 were identified by two or more models (Table 3.2). Overall, 14 QTNs were significantly associated with at least one trait with the LFMM model and the stringent Bonferroni 0.05/n threshold (Appendix II-4). The extent of the phenotypic variations for the traits at these 15 QTNs is illustrated (Figures 3.3A-B). Large effect QTNs at Chr4:17242614 and Chr5:15312783 were consistently associated with root network depth explaining 22.7 and 19.3% of variation for this trait respectively. Both were detected by at least four of the six multi-locus models and the LFMM model where the latter was also significant in the 3K dataset using the Bonferroni correction threshold (Appendixes II-4, II-5). A QTN associated with network perimeter with the highest PVE ($R^2 = 24.20$) at Chr9:19061342 using the multilocus FASTmrMLM (Table 3.2, Appendix II-3) was also significantly associated with the same trait based on the Bonferroni criterion in both dataset (Appendix II-4, Appendix II-5, Appendix II-6).

Table 3.2 *QTNs with high phenotypic variance explained (R^2>5%)*

Traits	Model[1]	QTN[2]	R^2 (%)[3]	LOD[4]	MAF[5]
NWW; NWW_Dep; RDWt	3,6; 3; 5	Chr1:756854*	6.32-13.07	3.1-6.0	6.06
RDWt	5	Chr1:4908649	8.50	4.8	7.07
MedR	4	Chr1:11064283	7.82	3.9	7.07
SDWt	4	Chr1:18970469	10.65	3.6	8.08
SDWt	5	Chr1:20356976	8.04	6.7	6.06
NWDis	3	Chr2:4513304	11.49	3.4	5.94
NWW_Dep	2	Chr2:5963452	14.88	3.9	9.9
NWDis	4	Chr2:7095057	10.54	3.4	10.1
MaxR; MedR; NWL; NWPer; NWA; NWSA	2; 3; 3,6; 6; 3 6	Chr3:6925560	9.00-13.34	3.7-4.4	3.96
NWW	3,6	Chr3:16939026	8.89	4.2-4.6	3.96
SDWt	4	Chr3:17343476	7.48	3.7	8.08
NWDis	3	Chr3:18772054	9.68	4.206	4.95
NWDis	4	Chr3:25380098	12.72	3.7	13.13
NWDep	2,3,4,6	Chr4:17242614*	7.18-22.68	3.6-5.1	7.07
SRL	2,3,5,6	Chr4:18399285	10.65-15.60	3.3-4.9	10.1
MedR	4	Chr5:1375386	12.94	4.0	9.09
RDWt	3,7	Chr5:2645287	21.88	4.2	6.93
MedR	4	Chr5:11019409	10.32	3.4	6.06
NWDep; NWL; NWSA	1,2,3,4,6; 6; 4	Chr5:15312783*	5.61-19.26	3.0-7.4	15.15
SDWt	5,6	Chr6:3310382*	11.45-17.64	3.3	13.13
SRL	3	Chr6:7732273	5.47	3.2	9.9
MedR	4	Chr7:4774423	7.72	3.025	6.06
NWW_Dep	3,4	Chr7:6346464	7.32-13.46	3.7-3.9	10.1
SRL	3,6	Chr8:21825897	5.00	3.4	13.86
RDWt	5	Chr9:15946848	9.16	5.3	7.07
NWPer	3,7	Chr9:19061342*	24.20	4.0-9.8	11.88
MedR	2	Chr11:5382629	8.27	3.2	5.94
SDWt	3,7	Chr11:8176149	16.53	4.8	7.92
NWV	3,6	Chr12:255713	11.34-11.40	3.1-3.2	5.94
SL	3	Chr12:3690290	12.58	3.3	5.94
SDWt	4	Chr12:12200657	10.05	3.5	18.18
NWL; NWPer; NWA; NWSA	4,6; 4; 4; 4	Chr14:13363192*	5.33-10.69	3.2-4.6	9.09
NWW_Dep	3	Chr14:15462441	14.77	5.0	2.97
MedR	4	Chr15:10531332	8.52	3.8	6.06
NWL; NWPer; NWA; NWSA; NWV; NWW	3,4,6; 3,4,6 3,4,6; 2,3,4,6 2,3,6; 2,4	Chr15:11371216*	5.36-14.17	3.0-4.6	6.06

[1]Models are 1=FASTmrEMMA, 2=FASTmrMLM, 3=ISIS EM-BLASSO, 4=mrMLM, 5=pKWmEB, 6=pLARmEB; LFMM=7 (only QTN that passed the Bonferroni correction criterion); [2]QTN=Quantitative trait nucleotide, chromosome number and position are indicated, asterisks (*) indicate QTNs detected by multiple models from both 7K and 3K datasets; [3]R2=coefficient of determination explaining phenotypic variation due to allelic effect; [4]LOD=logarithm of odds; [5]MAF=minor allele frequency

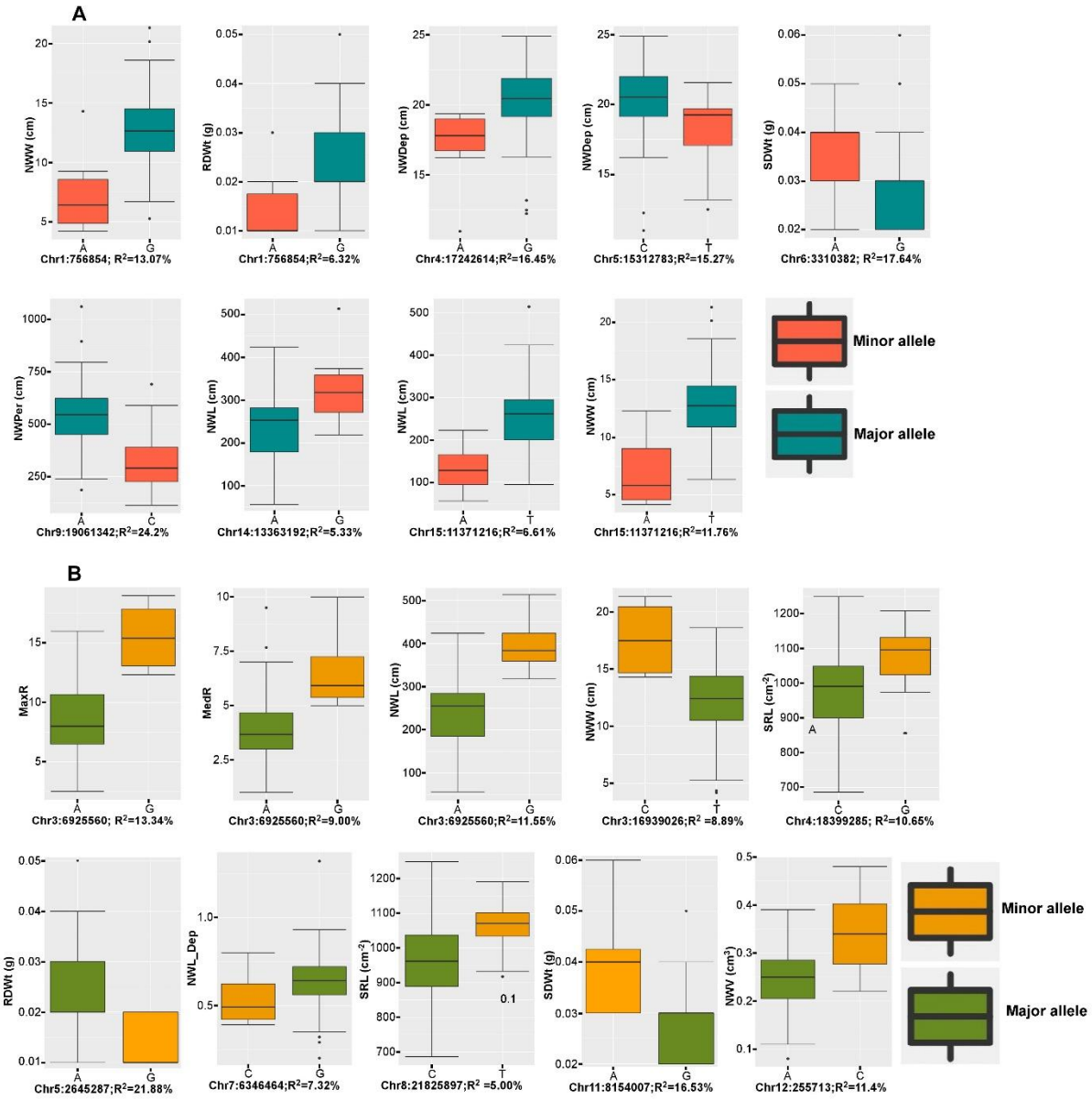


Figure 3.3 *Phenotypic variations between alleles at large effect QTNs.*
*(**A**) Phenotypic variations at large effect QTNs consistently detected in both 3k and 7k datasets, and (**B**) phenotypic variation at large effect QTNs that were detected by at least two models in at one or both datasets. Trait abbreviations are listed in Table 1.*

As expected, QTN associations with multiple correlated traits were observed. For example, a QTN at Chr15:11371216 was associated with the following six correlated root network traits: NWA, NWL, NWPer, NWSA, NWV, and NWW (Appendix II-2). This QTN was one of the large-effect QTNs detected by multiple models for all the traits in both datasets (Tables 3.2 and Appendix II-3). Two QTNs (Chr6:3310382 and Chr11:8176149) were associated with shoot dry weight (SDWt), each explaining more than 15% of the phenotypic variance of the trait (Table 3.2). QTN Chr11: Chr11:8176149 was also significant for SDWt in both 3K and 7K datasets using the Bonferroni correction criterion (Figure 3.4).

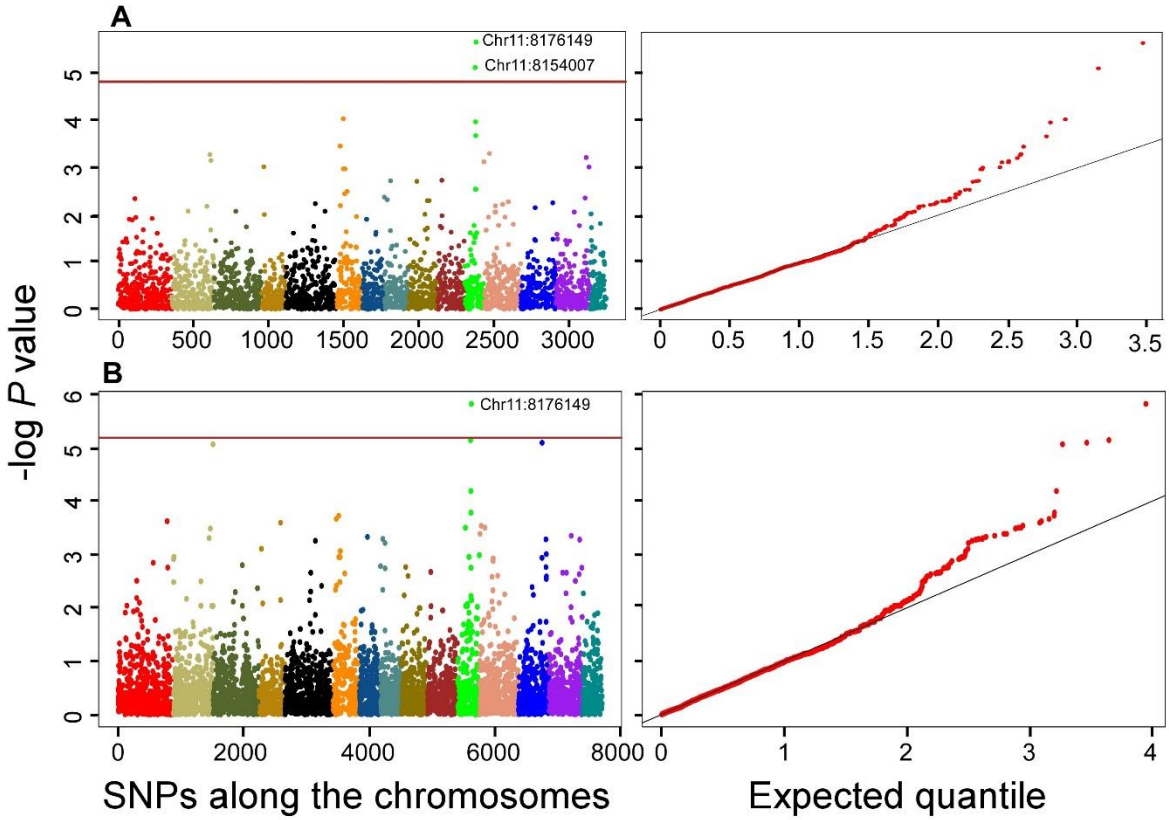


Figure 3.4 *Manhattan and QQ plots showing QTNs that are significantly associated with shoot dry weight using the stringent Bonferroni criterion for both datasets.*
*(**A**) 3K dataset; (**B**) 7K data set. The horizontal brown lines indicate the threshold* P = *0.05/n = 0.05/3243 = 1.54178E-05 and 0.05/7707 = 6.48761E-06 for the 3K and 7K datasets, respectively. Position of the significant QTNs on chromosome 11 are indicated. Colors in Manhattan plot indicate the 15 chromosomes of flax in order from 1 to 15.*

3.4.4. Genes linked to QTNs

Most loci within 100 kb (Figure 1.5) up- and downstream of the detected QTNs harbored genes that had previously been reported to play role(s) in root and/or shoot development in plants (Figure 3.5, Appendix II-3). The genes were primarily related to auxin efflux, nutrient transport and plant immunity. Loci corresponding to large-effect QTNs detected by multiple methods harbored genes for the growth and development of plant organs. For instance, the locus defined by QTN Chr15:11371216 harbored genes predicted to encode a lateral organ boundary (LOB) protein and a mitogen-activated protein kinase (MAPK). Network analysis indicated that this predicted MAPK gene likely interacts with other MAPK genes typical of MAPK cascades (Appendix II-7). QTN locus Chr5:15312783, also detected by multiple methods but in this case for its association with root depth, comprised genes predicted to encode GRAS (collective name for gibberellic acid insensitive (GAI), repressor of GA1 (RGA) and Scarecrow (SCR)) transcription factors (Gao et al. 2004). This locus also harbored genes that were predicted to encode ARM repeats, a GATA-type zinc finger transcription factor family protein and YUCCA6 (YUC6). The other large-effect QTN associated with root depth (Chr4:17242614) was linked to Arabidopsis orthologue genes AT5G37020 and AT5G10360 that encode auxin response factor-8 (ARF8) and sucrose synthase-6 (SUS6), respectively.

Most QTNs associated with shoot dry weight were at loci containing genes predicted to encode pentatricopeptide repeat (PPR), photo-morphogenesis, ubiquitin-related and plant defence proteins such as multi and toxic compound extrusion (MATE) (Figure 3.5, Appendix II-3). The large-effect QTNs Chr6:3310382 and Chr11:8176149 associated with shoot dry weight were located in relatively high gene density regions involved in plant immunity, development processes and plant growth regulation (Figure 3.5, Appendix II-3). Among others, a gene predicted to function as a suppressor of phytochrome A-105 (SPA3) and assumed to regulate plant growth by suppressing photo-morphogenesis (Laubinger and Hoecker 2003), was

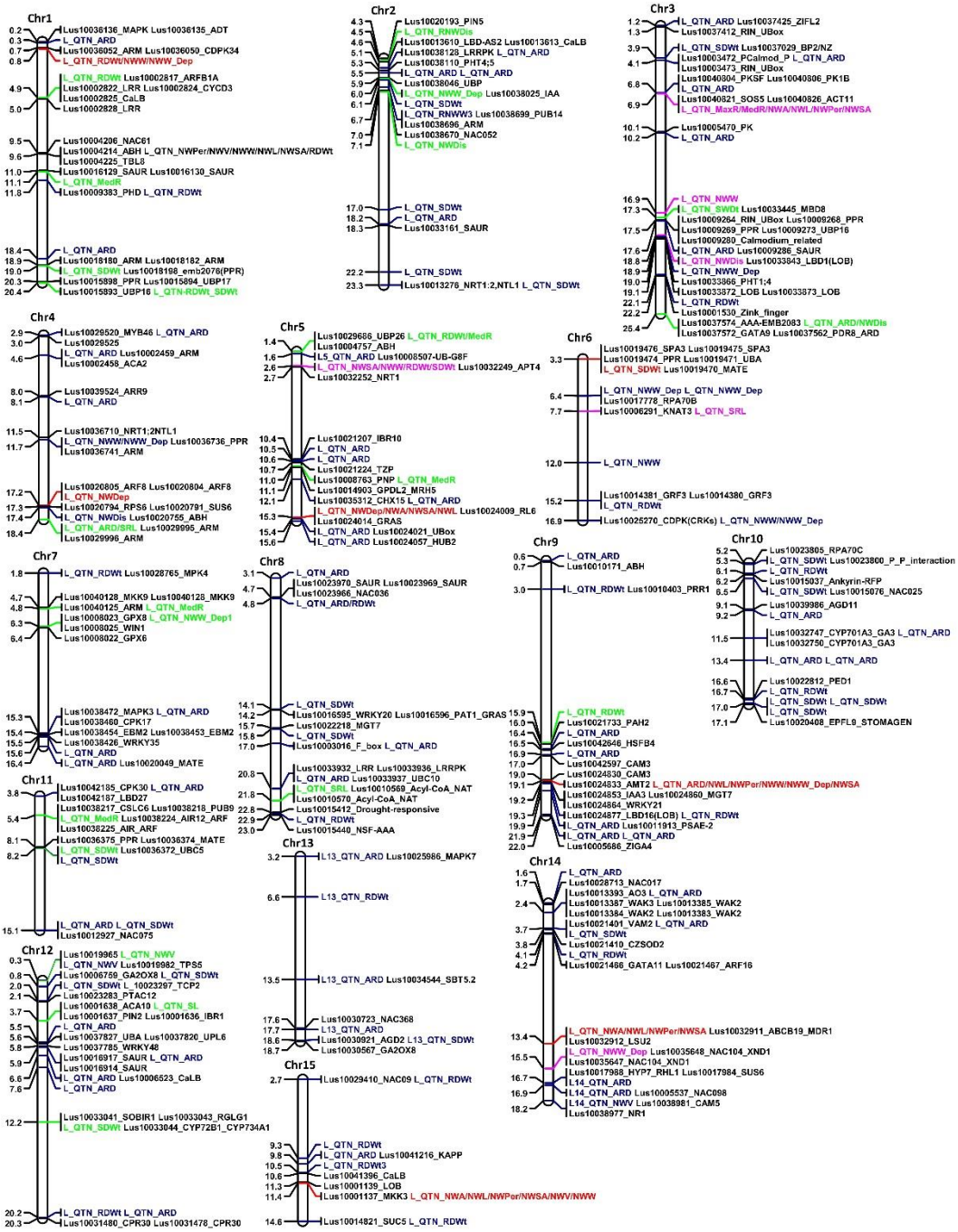


Figure 3.5 *Physical map of the 15 chromosomes of flax illustrating the position of the QTNs.*
Proximal candidate genes (right of the chromosomes). QTNs with R^2>5% in the 3K, 7K or in both datasets are in green, purple and red, respectively. QTNs indicated in blue have R^2<5% and were detected in the 3K dataset. Numbers on the left of the chromosomes represent physical distances in megabases

duplicated at QTN locus Chr6:3310382. The SPA3 and UBA proteins (Appendix II-3) were predicted to interact via COP1 (Appendix II-7).

3.5. Discussion

Root-trait measurements provide essential information to facilitate varietal improvement in breeding programs to select superior genotypes especially in nutrient and moisture deficit areas (Comas et al. 2013; Ndour et al. 2017). Root traits have already been used in breeding schemes to select elite genotypes in crops such as wheat for example (Wasson et al. 2014). Applications of root phenotype-genotype association through GWAS has enabled the identification of important QTL for root traits (Hochholdinger et al. 2018) that impact shoot traits including yield (Reinert et al. 2016). Our study provides insights into phenotype-genotype associations for early root and shoot traits of flax genotypes from over 20 countries by identifying QTL and proposing plausible candidate genes for further investigations.

3.5.1. Phenotypic variation

The significant variation observed among genotypes for most of the early root system and shoot development traits evaluated points to the genetic diversity of flax for such traits and the potential for genetic improvement. The highest and lowest morphometric values reflect the level of diversity within the gene pool, promising useful materials for improvement through breeding. The wide range of variation for most agronomic traits among flax genotypes (Richards and Diederichsen 2003) is well represented in the core collection (Diederichsen et al. 2013; You et al. 2017). Most genotypes in the flax core collection exhibit both fiber and oilseed features that could be attributable to selection processes for dually elite ones (You et al. 2017).

Correlations between root and shoot traits reflect the notion of balance between roots and shoots referring to plants partitioning their resource allocations between the two plant parts (Davidson 1969; Garnier 1991). A high correlation between shoot and root dry weights has also

been reported in rice (Zhao et al. 2019); this is not surprising considering the role of roots in supplying nutrients to the above-ground parts. The higher correlations between root network traits such as network perimeter, network length, network surface area and network volume with shoot traits compared to that with root depth suggest the crop's reliance on the top layer rooting system for resource uptake (Hocking et al. 1997; Kar et al. 2007).

The inverse relationship between average root diameter and root network related traits such as total network length, perimeter and surface area have been reported not only in flax (Soto-Cerda et al. 2019) but also in other plants such as *Arabidopsis thaliana* (Qian et al. 2015). The strong negative correlation between specific root length and average root diameter agrees with similar trade-offs between these traits in plants such as maize (Zhu and Lynch 2004) and various tree species (Bauhus and Messier 1999; Kramer-Walter et al. 2016).

3.5.2. Genome-wide association and candidate genes.

The genetic structure of the mini-core collection used herein is in accordance with previous reports describing the extent of the genetic and agronomic trait diversity of the flax core collection and indicating its suitability for GWAS (Soto-Cerda et al. 2013). Most QTN loci harbored genes predicted to play a role in organ development. QTN Chr15:11371216 was identified based on its association with multiple root traits, its large effect and its detection by multiple models. The LOB and MAPK orthologous genes at this locus have already been shown to contribute to root development in *Medicago trunculata* (Ariel et al. 2010; Han et al. 2014; Liu et al. 2005). LOBs are plant-specific proteins known for their involvement in lateral organ development (Shuai et al. 2002) including lateral root formation (Jeon et al. 2017). LOB proteins not only mediate a number of root and shoot development processes but can also respond to environmental stimuli (Shuai et al. 2002; Xu et al. 2016). MAPKs regulate several physiological processes of all eukaryote organisms including microbes and metazoans (Nishihama et al.

1995). In plants, they are involved in diverse developmental processes including response to abiotic stresses (Nakagami et al. 2005) but they have also been reported to play a role in regulating plant root growth via auxin signaling (Zhao et al. 2013).

The consistent association of Chr5:15312783 with root depth based on most of the models used and the occurrence of multiple genes predicted to function as GRAS, ARM, ATA and YUC genes at its locus hint at a possible role in determining root depth. GRAS family genes play important roles in plant organ development (Bolle 2016). The predicted GRAS gene at this locus appeared to be an orthologue of *Arabidopsis thaliana's* AT5G66770 that encodes SCR (Gao et al. 2004), thereby supporting the candidacy of this gene for the QTL. SCR and SHORTROOT (SHR) proteins of the GRAS family are known for their regulatory functions of root development in Arabidopsis (Benfey et al. 1993; Kamiya et al. 2003; Sbabou et al. 2010). These two proteins interact and work in tandem, i.e., SCR regulates the movement of SHR (Cui et al. 2007). Arabidopsis orthologues ARM (Coates et al. 2006), GATA (Behringer and Schwechheimer 2015), and YUC (Cha et al. 2015; Woo et al. 2007) at this locus could also be important for the development of plant organs in general and roots in particular. The specific YUC6 orthologue gene at this locus was reported to have important role in drought tolerance in Arabidopsis (Atkinson et al. 2015). The large effect Chr4:17242614 QTN for root depth may be attributable to an ARF8 orthologous gene that can regulate auxin-mediated process and influence primary root elongation (Blilou et al. 2005) and possibly lateral roots as well (Lee et al. 2019; Wang et al. 2015). ARF8 has been demonstrated to have clear effect on root growth habit where the roots of wild type plants grew slanted and those of mutants had a vertical downward elongation (Tian et al. 2004). This gene can also regulate flower maturation in later developmental stage, affecting fertility and seed production (Ghelli et al. 2018; Nagpal et al. 2005). The existence of a SUS orthologous gene as a regulator of primary root development (Sturm et al. 1995) may also be entertained as a candidate QTL for flax root depth.

The consistent occurrence of PPR, MATE and ubiquitin related orthologous genes at large effect QTN loci associated with shoot traits, especially with shoot dry weight, make them strong candidates. The PPR genes are involved in plant growth and stress tolerance in several plant species (Laluk et al. 2011; Wu et al. 2016; Xing et al. 2018). MATE efflux proteins on the other hand play a vital role in plant immunity against different toxins (Diener et al. 2001) including secondary metabolites (Gomez et al. 2009), xenobiotics such as heavy metals (Li et al. 2002) and aluminum (Li et al. 2017). Ubiquitin orthologues linked to shoot dry weight QTNs are candidate loci on the ground of their known roles in shoot development (Stirnberg et al. 2002; Yang et al. 2007). The SPA3 orthologous gene at the Chr6:3310382 locus, the QTN associated with shoot dry weight, is a member of a gene family that acts as a suppressor of phytochrome A and regulates photo-morphogenesis (Hoecker et al. 1999; Laubinger and Hoecker 2003). In Arabidopsis, SPA3 has a pronounced effect on seedling elongation (Laubinger and Hoecker 2003) and, interestingly, it is expressed in all above ground tissues while showing no detectable expression in roots (Zhu et al. 2008b). The SPA3 protein in Arabidopsis and rice has a conserved DWD motif that reflects its role in modulating a protein involved in photo-morphogenesis repression and in activation of etiolation through CONSTITUTIVELY PHOTOMORPHOGENIC1 (COP1); the latter also possesses E3 ubiquitin ligase activity (Lee et al. 2008). A secondary role for SPA3 that would be related to UBA at this locus through COP1 can therefore be considered (Appendix II-7).

Most genes linked to the large effect QTNs associated with root and shoot traits are responsive to abiotic stresses such as drought, salt and temperature. For instance, some GRAS family genes have useful roles in drought and salt tolerance (Ma et al. 2010; Xu et al. 2015). Some members of MAPK genes positively regulate low-temperature tolerance while decreasing drought and salt resistance (Jia et al. 2016). The role played by auxin-related genes (e.g. LOBs) in various abiotic stresses, including drought, have been reported for many plant species

(Huang et al. 2016; Jain and Khurana 2009; Jung et al. 2015; Wang et al. 2010). The QTN trait associations in these studies may further implicate variation in local adaptation to different environmental conditions given that our germplasm represented all flax growing regions of the world and represented more than 95% of the genetic diversity of whole flax core collection.

3.6. Conclusions

Early root and shoot trait phenotyping and GWAS of flax have provided insights into the complex relationships of these traits and their associated QTN loci were mined to hypothesize candidate genes. Early root network traits are interrelated, positively impact shoot traits and, consequently, seedling vigor. As such, early root establishment may also affect downstream yield performance. Our results suggest that, in flax, the extent of the root network is more important than root depth *per se* during the early growth stages. Root development studies spanning all growth stages would be necessary to quantify the relative importance of both traits over the entire cropping season and under different moisture regimes.

The GWAS yielded QTNs associated with most of the root traits and both shoot traits and the candidate genes identified at the major loci provide grounds for further investigations, particularly as they relate to stress tolerance. Some of the QTNs have pleiotropic effects that can either stem from linked genetic features at the loci or from single genes affecting multiple root traits. Shoot dry weight associated loci harbored genes that regulates physiological processes in above ground plant parts such as photosynthesis. Genes expressed at different stages and in different tissues may be tested for their specific role in controlling agronomic traits or imparting stress tolerance. Therefore, some of these loci may co-locate with QTL for other traits not measured herein. Given the polygenic nature of several agronomic traits, consideration must be given to QTL of small effects because their cumulative impact is

important in pre-breeding and positive selection can be achieve through genomic selection and other marker-assisted breeding schemes.

4. Genome wide association for drought tolerance related traits in flax using multiple models

This chapter will be submitted to a peer-reviewed journal.

4.1. Abstract

In the current climate change context, drought is one of the critical phenomena challenging today's agricultural sector. To cope with this challenge, crop varieties adapted to moisture deficit conditions are becoming vital. Flax is a crop that can be greatly affected by limiting moisture conditions, especially during the early development and reproductive stages. In this study, 115 flax genotypes from more than 20 major growing countries were evaluated for eleven drought-related traits in irrigated and non-irrigated fields for three years. Six of the 11 traits showed significant ($P < 0.05$) variation between the irrigated and non-irrigated conditions. A genome-wide association study (GWAS) was performed for these six traits and six stress indices based on 106 genotypes and 12,316 SNPs using a single-locus and six multi-locus models. In addition, SNPs were reduced to 8,050 linkage disequilibrium (LD) blocks to which a restricted two-stage multi-locus multi-allele GWAS was applied. A total of 148 quantitative trait nucleotides (QTNs) and LD blocks were associated with at least one trait or stress index. Of these, 16 had large effect, i.e., explaining more than 15% of the genetic variance. Most loci, 100 kb down and upstream of the QTNs, harbored gene(s) previously predicted to play role(s) in their associated traits. Genes mediating responses to abiotic stress including drought, heat and salt resided at loci associated with stress indices. Some of the candidate genes seem to have a pleiotropic effect. Tetratricopeptide repeat (TPR), WRKY DNA-binding, leucine-rich receptor (LRR), heat shock (Hs), GRASS, MAPK, MADS-box transcription factor genes were the major genes at these loci.

Keywords: Drought tolerance, flax, GWAS, QTL, candidate genes

4.2. Introduction

Drought is one of the most severe natural phenomena that can result in whole ecosystem and socio-economic disasters (Zhao and Dai 2015). It is estimated to cause annual losses of 0.01 to 0.25% of the world GDP (Jenkins and Warren 2015), amounting to $7.8-194.7B USD based on the 2014 world GDP (World-Bank 2015). Supporting over 2 billion people, drought-prone areas constitute approximately 41% of the world's surface (Solh and van Ginkel 2014) and, severe to extreme drought-prone areas cover 18% (Kogan et al. 2013). Drought frequently leads to catastrophic agricultural losses (Hao et al. 2014; Zhao and Dai 2015). Drought leads to reduction in harvest areas and yield that can be exacerbated when coupled with high temperatures (Lesk et al. 2016). The recent drought-temperature effects were estimated to result in global yield losses of ~11.5, 12.5 and 9% for maize, soybeans and wheat, respectively (Matiu et al. 2017).

Linum usitatissimum (L.), commonly known as flax, is a versatile crop used as industrial, feed, food and fiber crops (Singh et al. 2011b). Flax, one of the founder crops, has been cultivated since the beginning of agriculture (Zohary and Hopf 2000). Its cultivated area spans from the warm South Asia to the cool temperate regions of Eurasia and the Americas (Casa et al. 1999). Moisture stress is one of the major challenges of flax production (Lisson and Mendham 2000; Rowland 1998). Although a few reports claimed flax to be drought tolerant in association with mycorrhizae (Augé 2000; Von Reichenbach and Schönbeck 1995), many more reports evidenced the susceptibility of the crop to moisture limitations (Hocking et al. 1997; Kar et al. 2007; Lisson and Mendham 2000).

Moisture availability is particularly critical at its early growth and reproductive stages (Lisson and Mendham 2000; Mishra and Singh 2010). Flax is more sensitive to moisture deficit than cereal crops such as wheat (Morillon and Lassalles 2002). Compared to other oilseed crops

such as rapeseed, sunflower and safflower, flax has a shallow root system (Hocking et al. 1997; Kar et al. 2007). As a result, it relies on nutrients and moisture available in the top soil horizon (Hocking et al. 1997) and then again it less competent than other plants including cereals such as wheat (Morillon and Lassalles 2002). It also does not tolerate intercropping even with legumes (Ahlawat and Gangaiah 2003; Zając et al. 2013). Flax genotypes however, vary greatly with regards to their tolerance to water limitations (Mostafavi 2011). Flax accessions grown under drought conditions showed a wide range of variation in traits such as grain yield, plant height and thousand seed weight (Diederichsen et al. 2006; Lisson and Mendham 2000).

Genome-wide association studies (GWAS) have become a well-established approach to investigate the genetic bases of traits (Huang and Han 2014). GWAS has been extensively performed to detect quantitative trait loci (QTL) for various traits including drought related traits in crops such as rice (Guo et al. 2018), maize (Wang et al. 2016b), wheat (Tricker et al. 2018), sorghum (Badigannavar et al. 2018) and the oilseed sesame (Dossa et al. 2019) to name a few. In flax, GWAS was used to identify genes underlying variation in traits such as oil quality (Soto-Cerda et al. 2014) , disease resistance (He et al. 2018) and early root and shoot development (see chapter 3). Here we evaluated 115 flax accessions for 11 drought-tolerance-related traits and applied GWAS for six traits and six stress indices of each based on variation among the final set of 106 accessions representing the global diversity of the crop with the objectives to: 1) identify QTL associated with these traits, 2) explore the genome and identify candidate genes linked to the associated QTL, 3) suggest QTL of potential breeding importance.

4.3. Materials and methods

4.3.1. Plant material and experimental design

A flax mini-core collection of 115 accessions comprising genotypes from more than 20 countries and representing ~95% of the genetic diversity of the flax core collection (Diederichsen et al.

2013) of 407 accessions was used. The collection was grown under irrigated and non-irrigated field conditions in a type-2 modified augmented design (MAD) (Lin and Poushnisky 1985) in Ottawa (Ontario, Canada) and Morden (Manitoba, Canada) from 2016 to 2018. In a MAD, test lines are unreplicated but variations due to soil heterogeneity are controlled through the replications of main plot and subplot control lines (Federer 1956; Lin and Poushnisky 1985). Irrigated and non-irrigated field trials were each divided into 16 main plots set up in a four row by four column design (Figure 4.1A). Each main plot contained nine subplots where the main subplot control, Canadian cultivar CDC Bethune, was located in the center plot. Canadian cultivars Prairie Thunder and Macbeth were replicated in eight randomly selected main plots as subplot controls. The 112 test lines were randomly distributed to the remaining subplots for a total of 144 subplots for each watering regime. Subplots were two 2m-rows with 15 cm between rows, 20 cm between subplots and 1 m between ranges. A 4-row border of CDC Bethune was used on both sides of the field and a full range border divided the irrigated and non-irrigated sections of the field.

An underground drip line irrigation system was used at the Ottawa location (Figure 4.1B) while the Morden location used an above-ground watering system to irrigate the fields. Drip lines were installed in the direction of the rows, 80 cm apart and 15 cm deep. Watering was regulated by timers and measured by a flow meter to provide an equivalent of an extra 2 mm of water daily from ~20 days after planting to the boll stage.

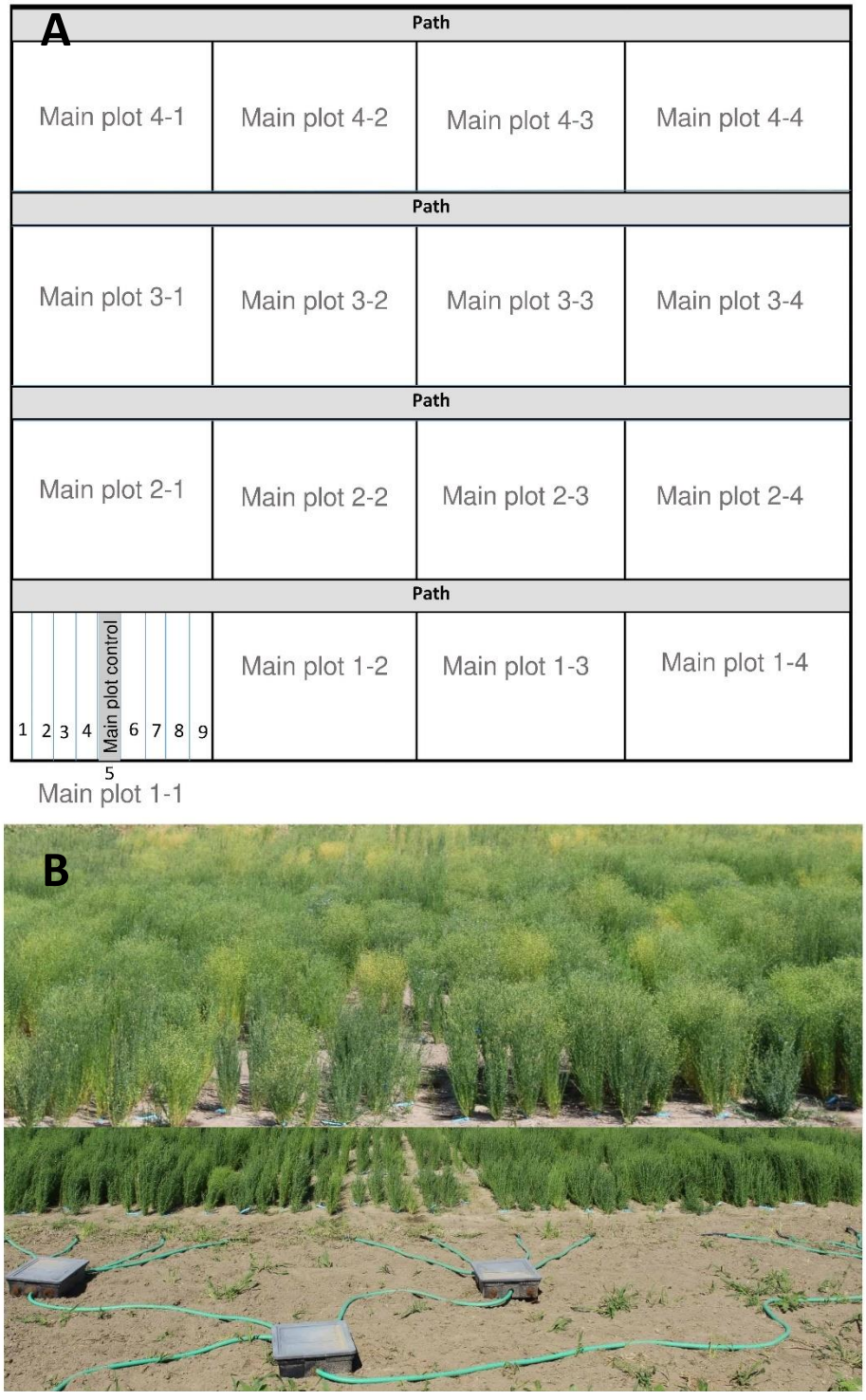


Figure 4.1 *Experiment field layout and partial view of irrigation system.*

*(**A**) Experiment field lay out. Numbers separated by hyphen following the text "Main plot" indicate the raw and columns respectively. The numbers 1-9 in Main plot 1-1 (bottom left **A**) indicate the subplots in each main plot and the middle sup-plot (5) assigned for main plot control. (**B**) Partial view of irrigated field at Ottawa.*

Phenotyping for the following six traits was performed at both locations: plant height (HT), stand (STD), days to flowering (FLD), days to maturity (MTD), thousand seed weight (TSW) and grain yield (YLD). The following additional five traits were also measured at the Ottawa location: canopy temperature (CT), bundle weight (BDLWt), boll weight (BWt), seeds per boll (SPB) and seed weight per boll (SWtPB). CT was measured twice at 5-7 days interval between flowering and ~10% green boll (seed filling) stage using an R2020 infrared thermometer (REED Instruments, 16975 Leslie St. Newmarket ON, Canada) as suggested for wheat (Pietragalla 2012). Maturity was considered when ~75 of the bolls turned brown. Seed count was performed using a seed counter (elmore C1-56354, Elmore Ltd, Mangelegg 58, Switzerland).

4.3.2. SNP data extraction and quality control

Processing of the Illumina HiSeq 100bp paired-end read of the resequencing data of each of the 407 accessions of the core collection yielded 1.7M SNPs. The 80% call rate and 5% minimum minor allele frequency (MAF) filters produced a dataset of ~570K SNPs from which the SNP data for the mini-core collection was extracted. This mini-core SNP dataset was further filtered using a 99% minimum call rate and 5% MAF in the mini-core *per se*. Then, genotypes with more than 5% missing SNPs were removed. Imputation was carried out using LinkImpute (Money et al. 2015), a tool implemented in TASSEL v5 (Bradbury et al. 2007) with default parameters and, only SNPs with no missing data points were retained.

4.3.3. Phenotypic data analysis

Precipitations during the cropping seasons were recorded from weather stations located at the two experimental sites. This information was used to calculate the total amount of water supplied to the field trials through natural rainfall for the non-irrigated trials or combined with irrigation for the irrigated trials. The high rainfall in Morden in 2016 and in Ottawa in 2017 were

non water-limiting (Appendix III-1); hence, data from these site-years were omitted in the GWAS.

Phenotypic data were adjusted based on the main plot and subplot controls using the MAD pipeline and SAS-implemented programs as previously described (You et al. 2013). Analysis of variances (ANOVA) were performed based on the adjusted data. Only traits that significantly responded to the watering regime were considered further. Broad-sense heritability (H^2) was estimated using the equation

$$H^2 = \sigma^2/(\sigma^2 + \frac{\sigma y^2}{y} + \frac{\sigma t^2}{t} + \frac{\sigma l^2}{l} + \frac{\sigma ytl^2}{ytl} + \frac{\sigma e^2}{ytl})$$

where σ^2, σy^2, σt^2, σl^2, σylt^2, and σe^2 represent the variances of the genotypes, genotype-by-year interactions, treatments (irrigated and non-irrigated), locations (environment), year-by-treatment-by-location interactions and error, respectively. *y, t*, and *l* represent year, treatment and location, respectively. Location-related variances were computed only for traits evaluated at both Morden and Ottawa locations.

Correlations among traits for each watering regime were estimated using Pearson and Spearman methods in R and visualized using the R package corrplot (Wei and Simko 2017). Only traits that significantly responded to the watering conditions were considered. To estimate the drought tolerance of the test lines, six stress indices were calculated as follows:

a. $\text{Drought susceptibility index (DSI)} = (1 - \text{TSi} / \text{TNSi}) /(\text{TSm}/\text{TNSm})$
b. $\text{Drought tolerance efficiency (DTE)} = \text{TSi} /\text{TNSi} * 100$
c. $\text{Stress susceptibility index (SSI)} = (1 - \text{TSi} / \text{TNSi}) /(1 - \text{TSm}/\text{TNSm})$
d. $\text{Schneider's stress severity index (SSSI)} = (1 - \text{TSi} / \text{TNSi}) - (1 - \text{TSm}/\text{TNSm})$
e. $\text{Stress tolerance index (STI)} = (\text{TNSi} * \text{TSi})/(\text{TNSm})^2$
f. $\text{Tolerance against stress (TOL)} = \text{TNSi} - \text{TSi}$

Where a and c indices were based on (Fischer and Maurer 1978), b on (Fischer et al. 1982), d on (Schneider et al. 1997) as modified by Singh et al. (2011a) and, e and f on (Fernandez 1992). T_{Si} and T_{NSi} represent trait values for i^{th} genotype under irrigated and non-irrigated conditions respectively while T_{Sm} and T_{NSm} are over all mean trait values under the same conditions in their order.

Pairwise correlations were calculated among the six indices, and mean values of the each traits under the two watering conditions were visualized using R. Since some indices were highly correlated (r > 90), principal component analysis (PCA) was performed for them and these indices were represented by principal component 1 (PC1) for each trait for a separate phenotype-genotype analysis.

4.3.4. Population genetic structure

A Bayesian clustering model was applied following the approach described by Pritchard et al. (2000). The appropriate number of ancestral populations (K) was estimated using the cross-validation technique described by Alexander and Lange (2011). Ancestral proportions of each genotype from K = two to 20 were computed using sparse Non-Negative Matrix Factorization sNMF (Frichot et al. 2014) in the R package LEA (Frichot and François 2015) with default parameters except for the number of runs which was increased from ten to 20. The K value with the lowest cross-validation error was selected as a suitable number of ancestral populations (Alexander and Lange 2011). In addition, a neighbor-joining (NJ) phylogenetic tree was constructed and PCA was performed using TASSEL v 5.2 (Bradbury et al. 2007) and results were visualized using Tree of Life (iTOL) v3 (Letunic and Bork 2016) and R, respectively.

4.3.5. Phenotype-genotype association

To investigate genetic variants underlying the traits that responded to the irrigation regimes, GWAS was performed using the final filtered SNP dataset for each of the six stress indices and

for the trait values from the irrigated and non-irrigated trials, i.e., a total of eight measurements per trait. For highly correlated stress indices, a separate GWAS was performed based on PC1 of each trait index. Six multi-locus models including FASTmrEMMA (Wen et al. 2016), FASTmrMLM (Zhang and Tamba 2018), ISIS EM-BLASSO (Tamba et al. 2017), mrMLM (Wang et al. 2016a), pKWmEB (Ren et al. 2018) and pLARmEB (Zhang et al. 2017) were applied using the R package multi-locus random-SNP-effect mixed linear model (mrMLM) (Wen et al. 2017). In addition, the single-locus mixed linear model implemented in TASSEL v 5.2 (Bradbury et al. 2007) was also applied. Benjamini and Hochberg (1995) false discovery rate (FDR) multiple comparison correction with *P* value threshold of 0.05 was used except for PC1 where associations were declared based on the Bonferroni correction (α=0.05/n). To further verify results, especially large-effect associated SNPs from the above methods, restricted two-stage multi-locus multi-allele GWAS (RTM-GWAS) a method that groups tightly linked SNPs into multi-allelic linkage disequilibrium (LD) blocks (He et al. 2017) was applied. Significant associations from RTM-GWAS were declared using the Bonferroni correction criterion with a threshold α=0.05/n where n was the number of LD blocks. Significant SNPs from all of the analyses were positioned onto chromosomes (You et al. 2018) as quantitative trait nucleotides (QTNs) using MapChart v2.32 (Voorrips 2002). To understand the variation due to the two alleles at large-effect QTNs (R^2 > 15%), medians of values of the associated traits between the alleles were where visualized in whisker plots and compared based on Kruskal-Wallis test.

4.3.6. Candidate gene prediction

Candidate genes for the traits were explored within 200-kb windows flanking the identified QTNs. Predicted genes were extracted from the flax reference genome (You et al. 2018b). These genes were investigated for their reported role in response to abiotic stresses, especially drought, using the Arabidopsis orthologues in the TAIR database (https://www.arabidopsis.org). Further reports about the role of the candidate genes in other plant species as they related to

drought tolerance were extracted from the literature. All genes previously reported to play a role in abiotic stress in general and drought tolerance in particular were considered.

4.4. Results

4.4.1. Phenotypic variation

Of the 11 traits under evaluation, six exhibited significant variation (P<0.05) between irrigated and non-irrigated conditions. Highly significant variation (P<0.01) between the two watering regimes was observed for bundle weight (BDLWt), canopy temperature (CT), seeds per boll (SPB) and plant height (HT) (Table4.1). Heritability (H^2) ranged between 44.7% for CT and 86.0% for days to flowering (FLD).

Table 4.1 *Summary of phenotypic variation of the traits*

Trait	Mean_NS±SE[1]	Mean_S±SE[2]	Significance [3]	CV[4]	H^2(%)[5]
BWt	0.04±0.00	0.04±0.00	NS	15.275	59.6
BDLWt	424.13±17.08	280.18±11.56	**	26.387	72.0
CT	6.29±0.24	3.15±0.13	**	28.755	44.7
SPB	4.75±0.10	5.2±0.10	**	23.678	63.4
SWtPB	0.02±0.00	0.02±0.00	NS	22.375	60.6
FLD	41.88±0.35	41.88±0.33	NS	3.876	86.0
HT	66.73±0.82	60.49±0.78	**	5.727	83.2
MTD	94.04±1.25	91.99±0.68	NS	6.281	72.3
STND	81.47±1.46	77.28±1.43	NS	9.07	85.8
TSW	5.52±0.05	5.32±0.05	*	5.117	83.7
YLD	96.39±4.00	86.24±4.31	*	39.962	67.7

[1]Mean_NS±SE=mean and standard error under Non-stressed (irrigated) condition, [2]Mean_S±SE= mean and standard error under stressed (non-irrigated) condition
[3] NS, *, and **, indicate the statistical variation between stressed and non-stressed conditions with $P > 0.05$, $P < 0.05$, and $P < 0.01$ respectively.
[4]CV=coefficient of variation
[5]H^2=broad sense heritability

Correlation coefficients (r) within traits between irrigated and non-irrigated trials ranged from 0.08 for CT to 0.88 for FLD (Figure 4.2, Appendix III-1). The within-trait correlations tended to follow heritability values for many of the traits (Figure 4.2 and Table 4.1). Correlations among stress indices that significantly responded to the irrigation were high ($|r| \geq 0.9$) for most indices.

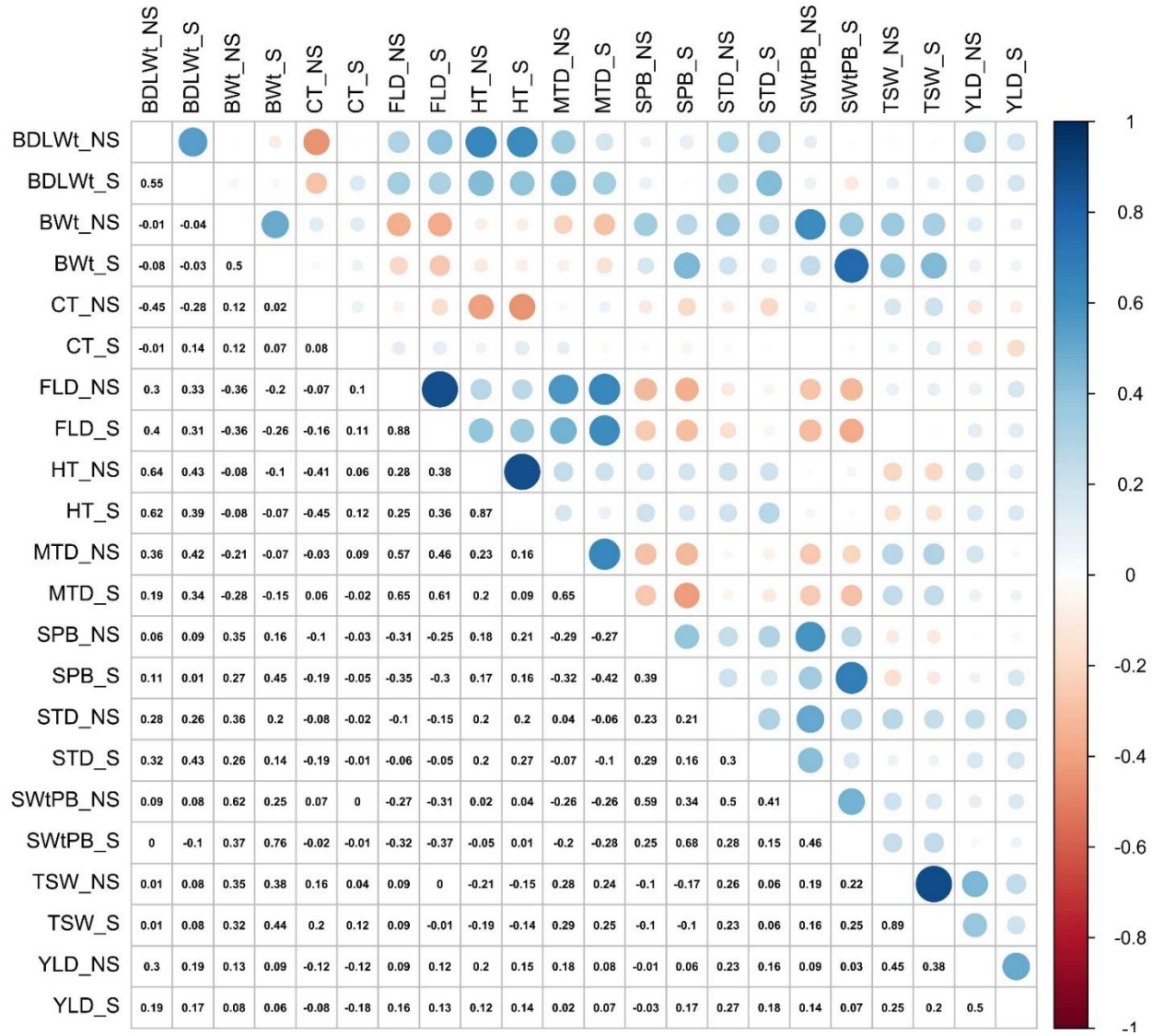


Figure 4.2 *Pearson correlation of all traits under the two watering conditions.*
***S** = stressed (non-irrigated); **NS** = non stressed (irrigated)*

A particularly high correlation ($r \sim \pm 1$) was computed among DSI, DTE, SSI and SSSI with TOL having a correlation value $r \geq \pm 0.9$ for most traits (Appendix III-2).

4.4.2. Genome-wide association

A total 12,316 SNPs with no missing data points and well distributed across all chromosomes were recovered for 106 genotypes. Given the genome size of flax of ~370 MB (Wang et al. 2012), this dataset represents a marker density of ~3.2 SNPs per 100 Kb. These SNPs could be grouped into 8,050 haplotype blocks based on their linkage disequilibrium (LD).

The genetic structure analysis clustered the genotypes into six recognizable populations that were assigned as Asian (ASIA), two Canadian (CANC1&2), two temperate (TEMP1&2) and a small cluster with two members (MININ) (Figure 4.3). The population structure and PCA yielded similar patterns where fiber and oil morphotypes were not partitioned in either PCA or admixture plot (Figure 4.3B-D). The NJ tree resulted in a slightly different clustering assignment but still distributed the genotypes into six major clades, of which, three were noticeably dominated by fiber types. The Canadian and temperate groups each were merged into a clade where the TEMP which also embraced the MININ (Figure 4.3E-F).

For the trait values under the two watering regimes and the six stress indices, a total of 144 QTNs were associated with at least one of the trait variables including 108 identified by the multi-locus models and 36 by the single locus method (Appendix III-4). In total 24 QTNs scored PVE>15% in the multi-locus, 16 of which, were detected by at least two models. From single-locus GWAS, 11 QTNs scored PVE>20% (Table 4.2). The single-locus MLM GWAS returned QTNs associated with only four BDLWt and two CT stress indices of which, three QTNs were also detected in the multi-locus GWAS (Figure 4.4). QTN Chr9:4203205 was associated with BDLWt stress indices DSI, DTE, SSI and SSSI based on three multi-locus models and on the single-locus MLM model with high PVE (R^2>12%) for all (Appendix III-4). GWAS for PC1 of highly-correlated ($|r| > 0.9$) indices of traits based on Bonferroni criterion yielded a total of nine QTNs, including seven previously detected (Figure 4.4).

Table 4.2 *QTNs with large (R^2>15%) phenotypic variance explained (PVE)*

QTN[1]	Trait[2]	Model[3]	LOD[4]	R^2(%)[5]	-LogP[6]	MAF[7]
Chr1:4520461	TSW_DSI+	5*	4.34	16.6	5.10	0.15
Chr1:7029139	TSW_NS/STI	3*,5*	3.56-6.51	18.2-20.2	4.29-7.36	0.15
Chr1:28692299	YLD_DSI+	5*	3.05	15.3	3.75-3.75	0.07
Chr2:5592445	SPB_DSI+	5*	4.38	15.8	5.15	0.10
Chr2:16676745	TSW_DTE	5*	3.165	15.8	3.87	0.12
Chr2:21621421	YLD_S	3*	5.334	15.8	6.14	0.08
Chr2:23123754	CT_DSI+	2*,3*,6	5.17-5.42	16.7	5.96-6.23	0.09
Chr3:9279281	CT_STI	3*,4*	4.2	22.6-22.9	4.96-5.14	0.09
Chr5:1375386	HT_NS/STI	3,5*,6	5.87-5.99	15.7	6.70-6.82	0.09
Chr6:2475973	BDLWt_DSI+	4*	_	20.0	5.39	0.10
Chr6:3310382	TSW_NS	2*,6*	5.02	15.3-16.6	5.82	0.12
Chr8:2514743	HT_NS/STI	2,3,5*,6	6.24-7.60	15.6-16.4	7.08-8.48	0.08
Chr8:9117708	CT_STI	4*	_	21.4	4.93	0.08
Chr8:16534117	BDLWt_STI	3,5*,6	4.83	17.0	5.62	0.08
Chr8:16606284	CT_STI	4*	_	21.0	4.85	0.08
Chr8:16698991	CT_STI	4*	_	25.7	5.70	0.09
Chr8:16722804	CT_STI	4*	_	25.3	5.64	0.10
Chr9:4203006	BDLWt_DSI+	2,3,4,5*,6	7.73	29.1	8.61	0.06
Chr9:4398116	BDLWt_DSI+	4*	_	22.9	5.99	0.06
Chr9:15446958	SPB_DSI+	2*,3*,6	6.25-6.48	18.8-21.7	7.09-7.33	0.06
Chr9:18937269	CT_STI	4*,5*	6.47	23.6-34.7	5.32-7.32	0.10
Chr9:18937872	CT_STI	4*	_	21.7	4.98	0.09
Chr9:18938885	CT_STI	4*	_	21.7	4.98	0.09
Chr9:18948993	CT_STI	4*	_	23.6	5.3224	0.09
Chr10:12528281	BDLWt_DSI+	4*	_	22.0	5.81	0.11
Chr11:3972867	YLD_DSI+	1*,2*,3*,6	5.03-6.08	15.5-19.2	5.82-6.91	0.08
Chr12:1853487	CT_TOL	4*	_	26.7	5.97	0.08
Chr12:1853750	CT_TOL	4*	_	23.9	5.46	0.08
Chr12:6352775	BDLWt_DSI+	1,2*,3*,6	8.84-9.48	25.9-28.0	9.75-10.41	0.10
Chr12:10910146	TSW_TOL	2,3,5*,6	4.05	16.6	4.81	0.18
Chr12:20557728	YLD_TOL	2,3,5*,6	4.65	16.7	5.43	0.08
Chr13:3886742	TSW_TOL	5*	5.89	21.0	6.72	0.07
Chr13:4581161	YLD_NS	2,5*	4.22	20.0	4.98	0.07
Chr14:205508	BDLWT_NS;HT_S	2*,3*;2,3,5*,6	5.11-5.64	15.5-19.8	5.91-6.46	0.08
Chr15:7035946	YLD_DSI+	5*	6.04	15.4	6.88_6.88	0.09

[1]QTN=quantitative trait nucleotide; number before and after the underscore indicate the chromosome and position of the QTN on the chromosome respectively. [2]DSI+=DSI/DTE/SSI/SSSI. [3]1=FASTmrEMMA; 2=FASTmrMLM, 3=ISIS EM-BLASSO, 4=MLM, 5=mrMLM, 6= pLARmEB; *model that detect the QTN with PVE>15% for multilocus GWAs, 4* single locus mixed linear model with PVE>20%. [4]LOD=logarithm of odds. [5]R^2=Coefficient of determination indicating PVE. [6]-logP=-log of *p* value. [7]MAF=Minor allele frequency

Figure 4.3 *Physical map of all trait associated loci and their linked genes. QTNs detected by single locus GWAS with R^2>20 in red; single locus GWAS R^2>20 & multilocus GWAS green, and single locus GWAS R^2<20% in pink. ; Blues are overlaps in LD-blocks and SNP association; prefixes Lus, L_, LDB are candidate genes, SNPs and LD-blocks respectively.*

Many QTNs were associated with multiple traits and stress indices (Appendix III-4) that were correlated (Appendix III-3). Chr15:14329974 QTN, detected by multiple models, was associated with five TSW stress tolerance indices, accounting for 4.14 to 9.68% of the phenotypic variance of these traits (Appendix III-4). The consistently detected Chr14:205605 and Chr11:3973246 QTNs had noticeably large effects with PVE exceeding 15% (Table 4.2). The former was associated with the biomass-related traits BDLWt and HT while the latter was associated with grain-yield stress indices. One of the three QTNs identified by the single- and multi-locus GWAS at Chr9:4203205 was associated with multiple BDLWt indices with PVE>15%. Significant phenotypic variations were observed between the two alleles at most of large effect (PVE>15%) QTNs detected by multiple models (Figure 4.5).

The RTM-GWAS using Bonferroni criterion captured a total of 13 LD blocks, of which eight overlapped QTN loci detected by at least one of the seven methods and the remaining were new (Table 4.3). Among the newly detected, LD block Chr12:LDB10739335-10743868 was strongly associated with five of the six SPB stress indices and with SPB per se under stress conditions (Figure 4.6 and Table 4.3). This LD-block was close to Chr12:10910146, a QTN that was associated with five of the six TSW stress indices (Appendix III-4) including TWS_TOL for which PVE > 15% was computed using multiple methods (Table 4.2).

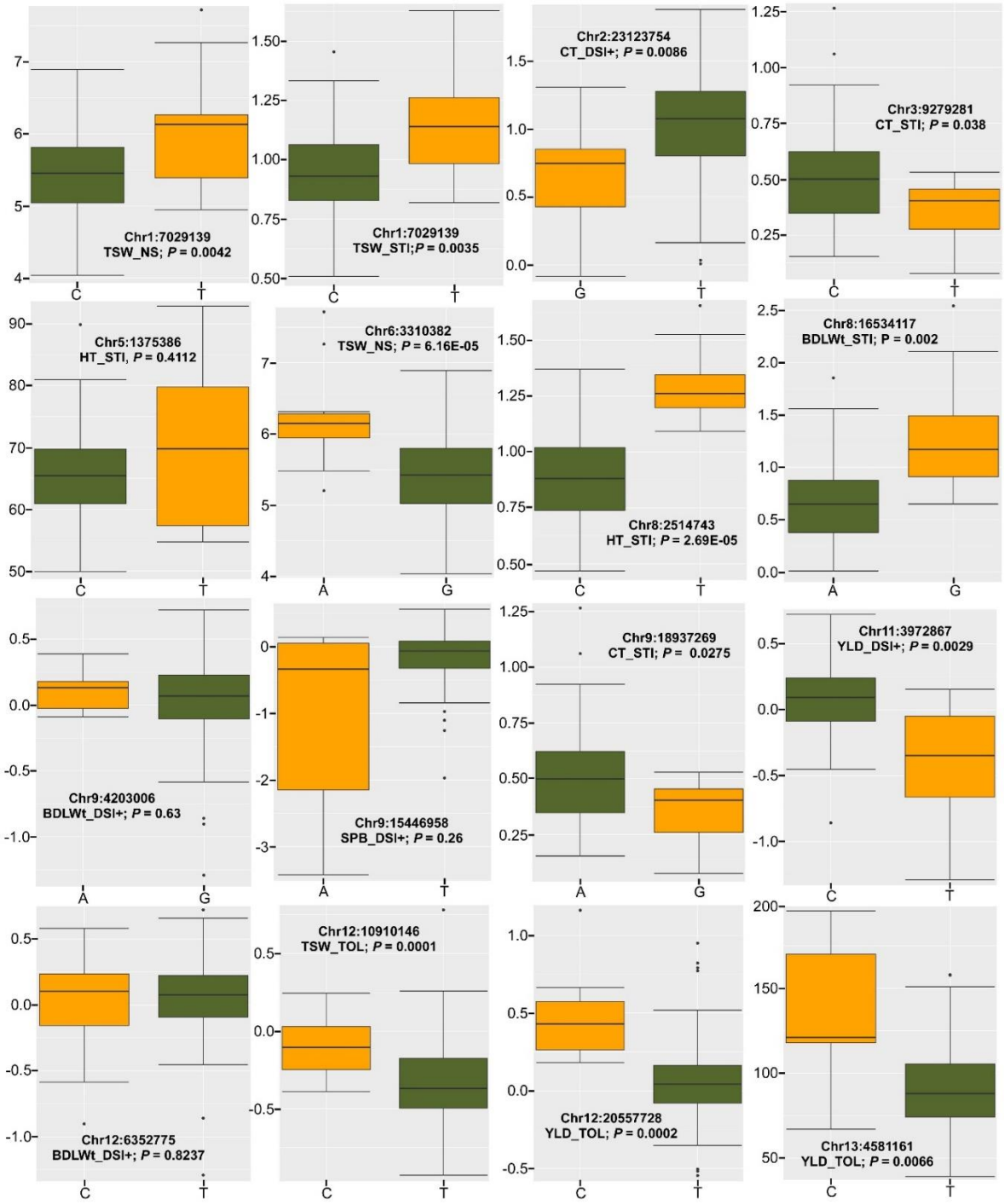


Figure 4.4 *Phenotypic contrast between alternative alleles at QTNs with large effect and detected by more than one model.*
The yellow and dark-green boxes indicate the minor and major allele respectively, DSI+ represent DSI, DTE, SSI and SSSI, P indicates the p value for difference in medians of the two allele based on Kruskal-Wallis test

Table 4.3 *LDBs detected using RTM-GWAS and predicted linked genes*

Trait	LDB[1]	*P* value	Gene	At[2] orthologue	Protein
YLD_TOL	Chr1:LDB-12677173-12686199*	8.14E-07			
SPB_DSI	Chr1:LDB-12799686-12819119	5.91E-06	Lus10020643	AT5G22380.1	NAC090
SPB_PC1	Chr1:LDB-12799686-12819119	5.92E-06			
BDLWt_DSI	Chr4:LDB-19895139-19897468	2.57E-07	Lus10000466	AT2G31110.2	TBL40
			Lus10000620	AT2G42570.1	TBL39
HT_STI	Chr5:LDB-1620825*	2.75E-06			
BDLWt_DSI	Chr6:LDB-2276653-2344083*	7.98E-09	Lus10034085	AT1G16130.1	WAKL2
			Lus10034087	AT5G15890.1	TBL21
TSW_NS	Chr6:LDB-3310405*	1.28E-06	Lus10019475/6	AT3G15354	SPA3
			Lus10019482	AT3G48430	REF6
SPB_S	Chr7:LDB-16126030	6.97E-07	Lus10038315	AT2G29960.1	CYP5
SPB_DSI	Chr7:LDB-16126030	4.72E-14			
SPB_PC1	Chr7:LDB-16126030-16126053	4.74E-14			
HT_STI	Chr8:LDB-2514780*	1.44E-06			
YLD_DSI	Chr11:LDB-3973246*	4.91E-07	Lus10042218	AT3G49400	WD40R
YLD_PC1	Chr11:LDB-3973246*	4.41E-06	Lus10042231/2	AT1G06950	TIC110
SPB_S	Chr12:LDB-10739335-10743868	2.61E-16	Lus10030150	AT4G12040.1	SAP7
SPB_DSI	Chr12:LDB-10739335-10743868	4.42E-27			
SPB_TOL	Chr12:LDB-10739335-10743868	4.11E-11			
SPB_PC1	Chr12:LDB-10739335-10743868	4.45E-27			
YLD_STI	Chr13:LDB-4209079*	2.17E-06	Lus10019778	AT2G38770	EMB2765
			Lus10019780	AT5G12380	ANNAT8
BDLWt_NS	Chr14:LDB-205605*	3.86E-06	Lus10009476	AT1G72180	CEPR2
			Lus10009481	AT4G09960	AGL11

[1]LDB=linkage disequilibrium block; * indicates LDB overlapped with loci detected by the multi-locus or single locus GWAS;
[2]At=*Arabidopsis thaliana*.

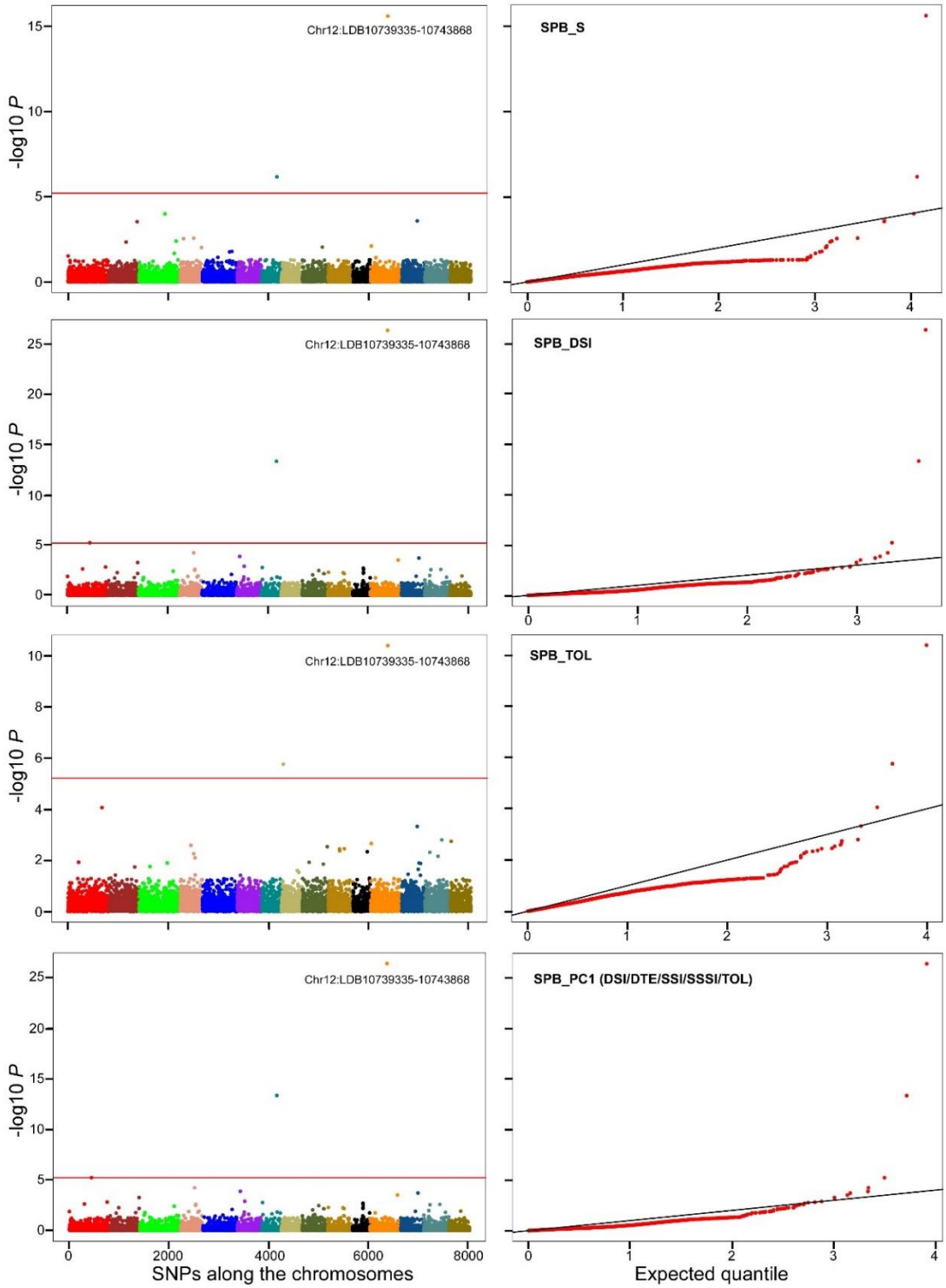


Figure 4.5 *RTM_GWAS showing a consistent high peak for Chr12: LDB 10739335-10743868 associated with seed per boll (SPB) and its stress indices.*
The locus harbors Lus10030150 which is orthologue of Arabidopsis AT4G12040, a gene predicted to encode a stress associated protein (SAP) of important role in seed development

4.4.3. Candidate genes

Loci within 100kb (Figure 1.5) up- and downstream of most QTNs harbored genes known to play roles in mediating stress responses (Appendix III-4 and Fig 4.3). Majority of genes at large effect loci were within 20kb of the associated QTN (Table 4.4). Some of the candidate genes identified were known for their involvement in drought tolerance, others were involved in other abiotic stresses such as heat and salt tolerances. Several genes at loci associated with traits measured in non-irrigated field had roles in regulating the corresponding trait with no known role in modulating stress responses. For instance, loci associated with seed related trait variables harbored genes involved in embryogenesis, pollen germination, ovule and other seed development processes including seed filling and phenological behavior such as flowering time. On the other hand, QTNs associated with stress tolerance indices and/or traits measured in non-irrigated fields were mostly linked to genes mediating responses to abiotic stresses such as drought, heat and salt (Table 4.4).

Chr15:14329974 QTN, associated with TSW stress indices, contained flax genes Lus10037856 and Lus10037864 orthologous to Arabidopsis AT4G00710 and AT2G45830 that encode a member of the MADS family AGAMOUS-Like 15 2 (AGL15 2) downstream target (DAT2) and a brassinosteroid-signaling kinase 3 (BSK3), respectively (Appendix III-4). Among the large effect (R^2>15%) loci, Chr14:205605 QTN associated with bundle weight and height harbored Lus10009480 whose Arabidopsis orthologue: AT1G15360 encodes SHINE1/WAX INDUCER1 (SHN1/WIN1) (Table 4.4). At the yield-stress indices associated Chr11: 3973246 locus, Lus10042229 and Lus10042231, orthologues of AT5G58380 and AT1G06950, respectively, encoding a CALCINEURIN B-LIKE- INTERACTING PROTEIN KINASE (CIPK) and TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS 110 (TIC110) in their order were identified. Ch13:4581423 QTN associated with yield in irrigated fields harbored genes of relevant role for grain yield status predicted to encode proteins such as

ADP/ATP CARRIER 2 (AAC2), ATP SYNTHASE 5 (ATP5), and ENHANCER OF DA1 (Big 1 in Chinese) (EOD1).

Table 4.4 *Candidate genes at loci of high PVE (R^2 > 15%) QTNs detected by multiple methods.*

		Genes		Predicted protein	Predicted function
Trait[1]	QTN[2]	Flax gene[3]	At Orth[4]		
TSW_NS/STI	Chr1:7029139	Lus10029127	AT1G23390	KFB	ovule development
		Lus10029115**	AT1G60770	RPPR4	seed development
CT_DSI+	Chr2:23123754	Lus10013240*	AT2G01850	XTH27	leaf size ; veins
CT_STI	Chr3:9279281	Lus10019365	AT2G25110	SDF2	heat
HT_NS/STI	Chr5:1375386	Lus10029690	AT1G77460	CSI3	flax fiber
		Lus10029692	AT1G01720	AF1	xylem development
TSW_NS	Chr6:3310382	Lus10019475*	AT3G15354	SPA3	seed development
		Lus10019474**	AT3G48250	PPR	seed development
		Lus10019482	AT3G48430	REF6	early flowering
HT_NS/STI	Chr8:2514743	Lus10025166	AT1G76520	PILS3	plant height
		Lus10013822	AT2G36870	XTH32	cell-elongation
BDLWt_STI	Chr8:16534117	Lus10004554**	AT2G35350	PLL1	root, shoot
BDLWt_DSI+	Chr9:4203006	Lus10040333	AT5G04530	KCS19	drought
		Lus10040335	AT3G12360	ITN1	salt
SPB_DSI+	Chr9:15446958	Lus10021766**	AT5G66850	MAPKKK5	drought
CT_STI	Chr9:18937269	Lus10024816**	AT4G37370	CYP81D8	moisture stress
YLD_DSI+	Chr11:3972867	Lus10042229*	AT5G58380	CIPK	drought
		Lus10042231	AT1G06950	TIC110	heat shock
BDLWt_DSI+	Chr12:6352775	Lus10016846	AT3G16857	RR1	drought
		Lus10016831	AT3G05640	EGR1	drought
TSW_TOL	Chr12:10910146	Lus10030137*	AT5G12840	NF-YA1	drought, seed
		Lus10030142	AT1G79280	TPR	flowering, drought
YLD_TOL	Chr12:20557728	Lus10031398*	AT3G06720	IMPA1	drought
YLD_NS	Chr13:4581161	Lus10019701**	AT5G13490	AAC2	anther dev.
		Lus10019703**	AT3G63530	EOD1	final seed size
		Lus10019706*	AT5G13450	ATP5	seed filling
BDLWT_NS; HT_S	Chr14:205508	Lus10009472**	AT1G71692	AGL12	drought, root
		Lus10009476*	AT1G72180	CEPR2	biomass
		Lus10009480*	AT1G15360	SHN1	cell wall
		Lus10009481**	AT4G09960	AGL11	plant height

[1]DSI+=DSI/DTE/SSI/SSSI; [2]QTN=Quantitative trait nucleotides; [3]*genes <20kb, **genes<10kb away from the corresponding QTN; [4]At Orth=*A. thaliana* orthologue

Chr6: 3310405 locus for TSW in irrigated field comprised four genes including, two predicted to encode suppressor phytochrome A-105 3 (SPA3) and the other two encoding a pentatricopeptide repeat (PPR) and relative of early flowering 6 (REF6) (Tables 4.2, 4.4 and Figure 4.4). The locus at the LD-block associated with the five SPB indices and SPB in non-irrigated field harbored an important gene Lus10030150 predicted to encode a STRESS-ASSOCIATED PROTEIN (SAP).

4.5. Discussion

Drought is the major challenge of crop production, reducing both yield and harvestable area (Lesk et al. 2016). Phenotype-genotype association via GWAS has become an important approach to identify small- to large-effect loci with the goal to accelerate the improvement of complex agronomic traits, most importantly yield and yield-related traits, through breeding for different environmental conditions using germplasm with no prior knowledge (Tester and Langridge 2010; Wang and Qin 2017). Using recent advances in phenotyping platforms, GWAS has been applied to dissect the complex genetic bases of drought tolerance in many crops (Josephs et al. 2017; Luo et al. 2019). Here we performed a GWAS for drought-tolerance-related traits and their corresponding indices on 106 flax genotypes representing the global distribution of the crop.

4.5.1. Phenotypic variation

Trait response to stress conditions can vary with species depending on the plasticity of the traits. The pattern of response to watering regimes of phenological traits, such as days to flowering and maturity, are inconsistent across plant species (Huang et al. 2018a). Moisture stress delayed flowering in finger millet (Mahalakshmi and Bidinger 1985) and rice (Saikumar et al. 2016). As we observed in flax, no significant response was observed in climbing bean (Ntukamazina et al. 2017). In Arabidopsis, flowering response to drought is conditioned by day

length where long days trigger and short days delay flowering (Schmalenbach et al. 2014). The significant differences in biomass and yield-related traits observed in our experiments occurred with previous reports in flax (Kariuki et al. 2016) and other plants (Saikumar et al. 2016) indicates water limitation results in less biomass and grain yields.

The heritability of most traits was higher than previously reported for the crop based on data collected from two locations and over 4-5 years (You et al. 2017), possibly as a consequence of a greater control of the variance residuals within environments in our three-year field study. However, for most traits, the magnitude of the heritability values was consistent with the previous report, i.e., traits with high heritability in You et al. (2017) were also high in our study and low heritability traits were low in both studies. The high level of correlation among the various drought indices was expected because they are strictly just various ways of expressing the traits. Such high indices' correlations were consistent with other crops such as wheat (Mollasadeghi et al. 2011), rice (Raman et al. 2012) and maize (Khayatnezhad and Gholamin 2010).

4.5.2. Phenotype-genotype association and candidate genes

The genetic structure of the population was consistent with earlier report on similar germplasm that were based on different SNP datasets (Chapter 3). It is also in agreement with the pattern of structures observed for the whole flax core collection of ~400 genotypes that was dominated by temperate flax with little discrimination between fiber and oilseed morphotypes in all analyses except the neighbor joining phylogenetic and independent of the genotyping data, i.e. ~50K SNPs (Sertse et al. 2019b) or ~400 microsatellites(Soto-Cerda et al. 2013).

Several genes located at QTN loci have previously been associated with regulation of drought tolerance or other abiotic stress tolerances. Chr15:14329974 QTN locus associated with TSW indices harbors candidate genes predicted to encode AGL15 (DAT2) and BSK3.

AGL15's role in embryogenesis and seed development processes make it a candidate for such trait (Chen et al. 2018; Perry et al. 1996), particularly in view of its responsiveness to abiotic stresses (Guo et al. 2016). Brassinosteroids, such as BSK3, are important plant hormones involved in seed development, regulating seed size and shape in Arabidopsis (Jiang et al. 2013) and, they are also responsive to abiotic stresses such as salt, cold and drought (Li et al. 2019).

SHN1/WIN1 at QTN locus Chr14:205605 associated with bundle weight and height is a strong candidate for the traits under drought conditions. The SHN1/WIN1 gene is known to regulate cuticular wax and cutin biosynthesis (Kannangara et al. 2007; Oshima et al. 2013) and leaf cellular structure including stomatal density (Yang et al. 2011). Overexpression of SHN1/WIN1 results in height suppression (Zhao et al. 2018) and improved drought tolerance (Aharoni et al. 2004; Bi et al. 2018). On the other hand, under well-watered condition, SHN1/WIN1 wheat transgenic lines were taller with more biomass than wild types (Bi et al. 2018). Despite the significant arrest of height, overexpression of SHN1 demonstrated superior biomass production under drought condition in crops such as wheat (Bi et al. 2018) and rice (Martins et al. 2018). Another candidate for biomass at this locus is Lus10009476, predicted to encode a leucine-rich receptor called C-TERMINALLY ENCODED PEPTIDE RECEPTOR-LIKE KINASE 2 (CEPR2). CEPR2 plays a regulatory function in root development (Dimitrov and Tax 2018) and its expression is dependent on nitrogen (N) availability within the root system (Tabata et al. 2014). Since biomass production can be influenced by N availability itself, performance of plant can be due to root system and response of the shoot to CEPR2 signaling that regulates growth (Tabata et al. 2014).

The KCS19 protein, predicted to be encoded by a gene positioned within Chr9:4203205 QTN locus for bundle weight stress indices (Table 4.4), is involved in cuticular biosynthesis (Suh et al. 2005), thus influencing drought tolerance (Kim et al. 2013). The other gene at this locus,

ITN1, encoding a transmembrane protein is involved in salt stress tolerance (Sakamoto et al. 2008). Drought and salt tolerance stress responses are mediated by similar biological processes involved in water efflux control and osmotic adjustments (Golldack et al. 2014). The other locus on Chromosome 12 (Chr12:6353095) harbored genes encoding a response regulator (RR) and an E growth regulator (EGR), both responding to drought stress and where the former might be involved in mediating drought response (Huang et al. 2018b) and shoot development (Xie et al. 2018b) while the latter was shown to regulate overall growth during drought (Bhaskara et al. 2017).

Candidate genes at Chr11:3973246 QTN for yield stress indices include CIPK10 and TIC110. CIPKs mediate responses to abiotic stresses including drought (Manik et al. 2015) and were reported to positively regulate drought stress (Cui et al. 2018; Luo et al. 2017). TIC110 is a component of the chloroplast heatshock protein 90 (Hsp90C) (Inoue et al. 2013) that was also hypothesized to play a role in drought tolerance (Rakhra et al. 2015).

The loci associated with trait variables in the irrigated fields can also be relevant for production when the crop is grown under more optimal water supply conditions. For example, AAC2, ATP5 and EOD1 located at Ch13:4581423 QTN locus for yield QTL are worthy candidates. AACs are ATP transporters involved in anther development and fertility, biological processes that are prerequisite for seed set (Rieder and Neuhaus 2011). This process might be more crucial for predominantly self-pollinated crops such as flax (Dillman 1938). Lus10019706 predicted to encode ATP5 is orthologous to AT5G13450 (Table 4.4) which is one of the genes associated with seed development, and especially, with seed filling in Arabidopsis (Hajduch et al. 2010). EOD1 is involved in seed size regulation in Arabidopsis (Du et al. 2014; Xia et al. 2013). This protein interacts with the ubiquitin receptor DA1 that is known to negatively regulate seed size in Arabidopsis (Li et al. 2008) and in other plants (Li and Li 2014; Wang et al. 2017).

The TSW QTN Chr6:3310405 locus identified under irrigated and non-irrigated conditions harbors genes predicted to encode SPA3 and REF6. SPAs are part of the protein complex involved in regulation of photo-morphogenesis (Laubinger et al. 2006; Shikata et al. 2014) and they modulate flowering time (Laubinger et al. 2006). Likewise, REF6 plays a regulatory role in photoperiodic responses (Noh et al. 2004), suggesting the locus as candidate QTL for flowering time. Flowering time and seed size are related traits (Du and Qi 2010). Seed size is positively related with thousand seed weight in plants (Mangini et al. 2018) including in flax (Tyson 1989) signifying the association of this QTL for thousand seed weight. This locus was previously associated with early shoot dry weight (Chapter 3)(Sertse et al. 2019a) which might reflect its effect in seed development thereby grain yield in later stage (Ellis 1992; Finch-Savage and Bassel 2015) and vice versa the effect of seed size in seedling performance (Ambika et al. 2014; Royo et al. 2006).

Most large effect QTNs showed significant phenotypic variations between contrasting alleles. For instance, at Chr8:2514743 QTN locus where alternate alleles significantly differed (P = 2.69E-5) for height stress tolerance index (Figure 4.5), the genes encoding PILS3 and XTH32 known to regulate plant height (Gallego-Giraldo et al. 2011; Mohanta et al. 2018) constitute strong candidate for QTL. PILSs play a role in auxin dependent plant growth regulation where PILS3 expression in particular results in dwarfism (Barbez et al. 2012). Similarly, XTHs are important in plant cells elongation (Irshad et al. 2008) and have been shown to mediate fiber elongation in cotton (Lee et al. 2010). Over all, several candidate genes were linked to the QTNs associated with different trait variables and predicted genes at these QTN loci largely coincide with various reports (Table 4.4).

The predicted SAP encoding gene at Chr12:LDB10739335-10743868 is orthologous to Arabidopsis gene AT4G12040 that encodes ATSAP7. This gene was reported for its positive

role under different stress conditions including drought in various plant species such as rice (Kanneganti and Gupta 2008), *Medicago* (Gimeno-Gilles et al. 2011), tobacco (Kanneganti and Gupta 2008; Saad et al. 2010), and tomato (Solanke et al. 2009). SAPs are proteins involved in seed development under stress conditions (Cooper et al. 2003). The strong association of this locus with SPB suggests its potential role under drought conditions in flax as well.

4.6. Conclusion

Drought tolerance related traits are complex and polygenic. Genome-wide association study has recently provided an opportunity to explore whole genome sequencing information to identify QTL with higher precision than the more conventional bi-parental QTL mapping. This study captured a fairly high number of genetic variants (QTNs) associated with different drought related traits, including grain yield. These variants mark potential candidate genes known to have a role in drought and other related stress tolerances such as heat, and salt. It also identified variants associated with yield and related traits under non water-limiting condition. The variants identified were often linked to genes whose function(s) were known to be associated with the underlying traits. For large effect QTN loci identified by multiple models, these candidate gene(s) are primary targets for in-depth investigation of their functional role and can be capitalized upon for marker assisted selections. Given the polygenic nature of most traits under investigation and additive effect, cross-validation of all QTNs in large populations, such as the whole core collection should be performed to entertain selection for drought tolerance by genomic selection.

5. General discussion and conclusion

5.1. Highlights of the study

Flax *(Linum usitatissimum)* with a chromosome number of 2n=2x=30 and a genome size of approximately 370 MB with over 43 thousand protein coding genes (Wang et al. 2012) and transposable elements accounting for 20-25% of the genome (González and Deyholos 2012). Flax belongs to the family *Linaceae* and the only cultivated species in the genus *Linum* that comprises over 200 species (Heywood et al. 1978). Flax is one of the eight members of the founder crops agriculture in the fertile crescent (Zohary et al. 2012) possibly domesticated ca 10000 years BP in the present Israel (Weiss and Zohary 2011) or Syria (Van Zeist and Bakker-Heeres 1975) . Although the plant was assumed to have first been domesticated for its seeds (Van Zeist and Bakker-Heeres 1975), long before this event, it had been used as source of fiber (Kvavadze et al. 2009). Since its domestication it has spread wide range from warm to cool temperate geographic regions and has been cultivated both as fiber and oilseed crop in nearly all its cultivation history (Herbig and Maier 2011). The crop is used to produce numerous products (Singh et al. 2011b) and recently gained more attention due to its nutritional and medicinal values (Goyal et al. 2014; Health-Canada 2014).

Despite its historical high values, flax production and its acreage is shrinking recently (Figure 1.3) due to its low yield that rarely exceed two tons per hectare (Wittkop et al. 2009). Flax production is constrained by different biotic and abiotic factors of which moisture availability is among the major ones. To understand available genetic diversity and assess genomic regions underlying flax drought related traits, this study performed global scale population structure in line with genetic signatures for eco-geographic adaptations and GWAS to trace loci associated with different drought tolerance related traits with further investigation of candidate genes. This study was performed based on two sets of genotypes: core (n=407) and mini-core (n=115)

collection that represent collections from over 35 and 20 major flax growing countries, respectively, both of the collections containing both fiber and oilseed types.

The flax population structure analysis based on the core collection using ~51K SNPs produced 11 distinct populations structured into four major groups: Abyssinian, Mediterranean, South Asian and Temperate. A total of eight of the 12 populations belonged to the Temperate group, two to the South Asian group and the last two were their own group. The flax population structure was attributed to the natural and anthropogenic selection pressures across the historical flax growing regions (Sertse et al. 2019b; Soto-Cerda et al. 2013). The variation between fiber and oilseed types was less pronounced than variation within morphotypes, reflecting that most of the core collection was dominated by genotypes carrying both fiber and seeds attributes (You et al. 2017). Most genotypes belonged to the Temperate group that included the Americas and Eurasia and that was dominated by a common haplotype suggesting a high rate of gene flow via human migration and germplasm exchange among breeders. The other three groups, geographically located in the hypothesized flax centers of origin (Vavilov 1951), harbored high concentration of haplotype endemism stressing the importance of these regions for further germplasm collections to increase genetic diversity in both breeding and gene bank stacks.

Two genome-wide association studies (GWAS) were performed on the flax mini-core collection: early root-shoot and drought. Early seedling vigor and root establishment is important for plant performance under stress conditions, including drought, in the entire development process including seed production (Kravić et al. 2018; Strock et al. 2019). With this understanding, two early shoot and 14 early root traits were phenotyped in a hydroponic pouch system adapted from Hund et al. (2009) (see details in chapter 3). The results indicated that the shoot traits were positively affected by the root network related traits while root depth had an

insignificant effect suggesting that the crop mainly relied on the top layer root system for nutrient and moisture (Hocking et al. 1997; Kar et al. 2007; Klimek-Kopyra et al. 2015). Genotype-phenotype association using seven different GWAS models identified 228 quantitative trait nucleotides (QTNs) associated with at least one of the 16 traits. Of those, 35 explained more than 5% of the phenotypic variance of their associated trait.

Candidate genes that might play a role in controlling the traits were identified in the 100 kb up and downstream of several of the QTNs. The large effect loci associated with root traits harbored root growth regulatory and abiotic stress (e.g. drought) responsive genes such as GRAS transcription factors (Xu et al. 2015), mitogen-activated protein kinases (MAPK) (Nakagami et al. 2005; Zhao et al. 2013), lateral organ boundary (Shuai et al. 2002; Xu et al. 2016) and other auxin related genes. Among the GRAS family genes, a predicted Scarecrow (SCR) at locus Chr5:15312783 was associated with root depth using multiple methods hinting at the root growth regulatory role of this gene (Sbabou et al. 2010) which operates in conjunction with the SHORTROOT (SHR) gene (Cui et al. 2007). The presence of other putative root growth regulatory genes such as the predicted Armadillo repeat (ARM) (Coates et al. 2006), the GATA-type zinc finger transcription factor family protein (Behringer and Schwechheimer 2015) and the drought responsive YUCCA6 (YUC6) (Atkinson et al. 2015; Woo et al. 2007) suggests the importance of this QTL.

Large-effect loci associated with shoot traits harbored genes involved in photo-morphogenesis and plant defense against various toxins. A predicted suppressor of phytochrome A-105-3 (SPA3) gene was present at the locus marked by large effect QTN Chr6:3310382 associated with shoot dry weight. SPAs are known to regulate morphogenesis and play an important role in seedling elongation (Hoecker et al. 1999; Laubinger and Hoecker 2003). SPAs interact with the repressor of light CONSTITUTIVELY PHOTOMORPHOGENIC1

(COP1) (Laubinger and Hoecker 2003) and regulate photoperiodic processes (Laubinger et al. 2006). The predicted SPA3 at this locus was reported to express in all above ground tissue with no detectable expression in roots (Zhu et al. 2008b) suggesting the importance of these gene in shoot development. Other important genes located at large effect shoot-associated loci included PPR and MATE efflux genes. PPR plays a role in plant growth, biotic and abiotic stress tolerance (Laluk et al. 2011; Wu et al. 2016; Xing et al. 2018) whereas MATE efflux genes are involved in defense against different toxins (Diener et al. 2001) including xenobiotic such as heavy metals (Li et al. 2002) and aluminum (Li et al. 2017). The presence of such genes at shoot trait associated loci may suggest the importance of shoot in toxins including heavy metals sequestration and detoxification processes (Rascio and Navari-Izzo 2011).

Analysis of variance for three consecutive years (2016-2018) of the minicore (n=115) collection at two locations under irrigated and non-irrigated field conditions for 11 drought tolerance related traits showed significant variation ($P<0.05$) between the two watering regimes for six traits (Table 3.1 Chapter 3). GWAS for these six traits from each watering condition and six stress indices of each trait using seven multi-locus and one single-locus GWAS methods detected 144 QTNs associated with at least one trait variable or one index. Further LD-block based GWAS using RTM_GWAS (He et al. 2017) detected four loci that were not detected using the aforementioned methods for a total of 148 loci associated with at least one trait. A total of 35 QTNs with large effect QTNs explaining more than 15% of the phenotypic variance were detected (Table 3.2, Chapter 3), of which, 16 were detected by more than two methods (Table 3.3, Chapter 3).

The majority of the loci flanked in 100kb up and downstream of the QTN harbored genes that have role in variation of the trait variable with which the QTN is associated. For instance QTN loci associated with seed traits including grain yield and their indices harbored genes

mediating processes such as flowering, pollen germination, and embryo and seed development. Likewise bundle weight and plant height associated loci harbored genes playing roles in fiber elongation, cell wall, stature, and biomass increment. This propensity has been seen for loci marked by large effect QTNs and their linked genes (Tables 4.1 and 4.3).

Genes playing role in stress tolerance such as SHN1/WIN1 (Bi et al. 2018; Kannangara et al. 2007), CIPK10 (Cui et al. 2018; Manik et al. 2015), TIC110 (Hsp90C) (Inoue et al. 2013; Rakhra et al. 2015), AGL15(DAT2) (Chen et al. 2018; Guo et al. 2016) and BSK3 (Jiang et al. 2013; Li et al. 2019) are some that were associated with large effect loci for stress indices of bundle weight, plant height, grain yield and thousand seed weight. These are gene candidates for drought tolerance. Loci associated with traits under non-water limiting condition harbored genes with a role in modulating the traits regardless of the stress levels. For example, large effect QTN loci Ch13:4581423 for grain yield and thousand seed weight included genes encoding AAC2 and ATP5 transporters that play a role in anther development and fertility, thence seed setting (Rieder and Neuhaus 2011), and an EOD1 gene known to control seed size (Xia et al. 2013). Large effect locus Chr6:3310405, also associated with thousand seed weight, harbored SPA3 and REF6 genes involved in flowering time regulation (Laubinger et al. 2006; Noh et al. 2004). As flowering time affect seed size (Du and Qi 2010) and consequently thousand seed weight (Mangini et al. 2018) which is typical in flax (Tyson 1989), loci containing genes controlling flowering can be considered as candidate of QTL for thousand seed weight. This locus was also one of the large effect loci associated with shoot dry weight, a trait that can be affected by the SPA genes that are involved in regulation of photo-morphogenesis (Laubinger and Hoecker 2003; Shikata et al. 2014).

5.2. Gaps and suggested future studies

This study has provided insight into the genetic architecture of drought tolerance and its related traits in flax, as well as on the genetic structure of a flax core collection, all of which have important implications in breeding and germplasm conservation. However, the research had some limitations and further validation may be required in some instances. The below section discusses some of these limitations and proposes future research directions.

5.2.1. Field based root phenotyping

The high throughput hydroponic system used to phenotype the flax roots has generated informative data, providing a basic understanding of variations in root architecture of the flax genotypes of the collection. In addition, it showed correlations between some root and shoot traits. Coupled with advances in imaging and image processing tools, such high throughput phenotyping techniques have successfully been used to determine the root architecture of several crops (Atkinson et al. 2019) and have allowed to generate phenotypic data necessary to perform GWAS of root traits (Mir et al. 2019). However, the natural soil rhizosphere in which plant roots have evolved and adapted is different and with numerous biological, physical and chemical interactions that determine the three dimensional root architecture of plants cannot be reproduced in a semi-hydroponic system (Downie et al. 2015). It would be interesting to perform a study of the same traits in soil and to compare the results. It would be challenging to phenotype so many accessions in a soil system because washing soil from roots is difficult and time consuming. However, a representative subset of accessions may suffice to perform such validation experiment.

5.2.2. Validation of QTNs and candidate genes

The systematic use of many multi-locus-GWAS models has captured more QTNs than is usually detected using the traditional single-locus approaches. The multi-locus models, not only

significantly detected a large number of QTNs, but also cross-validated one another and the single-locus model to some extent. Indeed, several QTNs were detected by more than one method. Some of these QTNs explained relatively large proportion (R^2>15%) of the traits' variance despite their polygenetic nature, i.e., the fact that they are controlled by several genes, each with small effects. However, the relatively small number of individuals (n=101-106) investigated combined to a minor allele frequency cut-off of 0.05 limit the statistical power and the confidence of QTNs with small effects (<10%). Validation of QTNs especially the large-effect ones (R^2>15%) is the next natural step and can be accomplished using bi-parental populations such as recombinant inbred line or doubled haploid populations. In such populations, the roughly 50:50 allele frequencies provides the statistical power to validate the associations. Though the large-effect QTNs should be given priorities for practical applications, judicious selection of parents for the development of bi-parental populations should enable cross-validation of many QTNs simultaneously.

Many, but especially the large-effect QTN loci, harbored gene/s known to play important role/s in mediating related-trait responses in other species. The majority of the large-effect QTN loci contained at least one candidate gene within <20kb of the QTNs, providing a certain degree of confidence for such candidate genes. However, candidate genes could be identified further up- and downstream of the QTNs as 200 kb windows were investigated. This region size was selected based on the average LD (r^2 = 0.1). LD varies significantly along chromosomes. In general, LD is higher in the centromeric and peri-centromeric regions which are rich in repetitive elements and where recombination is suppressed. Inversely, LD is lower in the telomeric and sub-telomeric regions which are highly recombinogenic and comparatively more gene-rich. The consequences of this LD landscape for GWAS is that more markers are needed to discover associations in low LD regions but QTNs in these regions are likely to be closely located to the causal gene/s or feature/s. Therefore, a smaller physical area needs to be investigated for the

identification of candidate gene/s in low LD regions. The reverse is also true, i.e. QTNs in high LD regions can be physically much farther to the causal gene/s or feature/s but a smaller number of markers are required for significant associations in these regions. Fortunately here, most of the QTNs were in the low LD regions where causal gene/s or features are likely close to the QTNs; hence, the 200-kb window may be considered more than sufficient. Nevertheless, validation of candidate genes. Transcriptomic studies using RNAseq, creating mutants using CRISPR/Cas9**,** and allelic studies using the sequence information of the individuals of the core collection are all possible ways to provide some validation of the candidate genes.

The GWAS using multiple models was efficient in capturing a number of QTNs using a ~3,000, 7,000 or 12,000 SNP data sets without missing data. The following discussion aims to address the advantages and the pitfalls associated with the extensive filtering performed herein. Given the flax genome size of 370 MB and the genome-wide LD average of r^2=0.1, these SNP data sets provided an average density of 0.8, 1.9 and 3.2 SNPs/LD, respectively. However, because LD is not uniform across the genome, the number of SNPs required to detect association varies along the chromosomes. The SNP filtering processes favored the selection of SNPs in the high-recombination telomeric and sub-telomeric regions, where more SNPs are needed, because these regions have less repetitive DNA and are more gene-rich. This is mirrored in the distribution of the QTNs on chromosomes (Figures, 3.5 and 4.4). These SNP data sets can be characterized as highly reliable and conservative. Higher SNP density would likely have identified more QTNs, possibly QTNs closer to the causal gene/s or feature/s. However, more does not always mean reliable and regardless of the size of the SNP data set, validation remains essential. Fortunately, the overall 1.7 M SNPs of the core collection will permit to test these hypotheses in the future through the application of less stringent filtering criteria with and without imputation. The phenotypic data produced in this project remains valid for such testing.

5.2.3. Phenotyping the core collection and genomic selection

GWAS remains a powerful approach in genotype-phenotype association. However, major challenges of GWAS applications lie in the complex genetic architecture of polygenetic traits and the power of detection of association with rare alleles (Korte and Farlow 2013; Tam et al. 2019). While GWAS is well suited for traits with simple genetic architecture, i.e., those controlled by few large-effect genes such as disease resistance responses (Korte and Farlow 2013; Louthan and Kay 2011), it is less effective in detecting small effect variants in complex polygenic traits (Tam et al. 2019) such as biotic stress responses (Deshmukh et al. 2014). To partially overcome this limitation, multiple multi-locus GWAS models that enabled the capture of small-effect variants have been used. However, deploying all or even some of these QTNs in breeding for selecting for drought tolerance poses a conceptual challenge. A genomic selection (GS) strategy that predict the breeding value of each marker across the whole genome may be more viable (Crossa et al. 2017). Using the phenotypic data developed herein, one could apply GS models using a subset as a training population. However, 115 individuals may be too few to obtain good prediction models. Alternatively, the core collection (~400) could be phenotyped and a subset thereof could serve as training population. Here, we showed that the breadth of genetic variability of the flax core collection could be expanded through careful sampling of genotypes from certain regions of the world (Sertse et al. 2019b). This must be considered prior to designing genomic selection populations. Once established, genomic prediction becomes feasible to hasten the breeding cycles and improve selection for multiple traits simultaneously. The genomic prediction can also be applied to estimate breeding value of the QTNs identified herein.

5.3. The practical and scientific implications of results

This study has produced major findings that can be considered for practical applications. The flax core collection is dominated by closely-related accessions from the temperate-growing

regions but it has an under-representation of accessions from the flax presumed centers of diversity regions. Conversely, these less-represented regions were shown to harbor a high concentration of unique haplotypes that indicate a need for further exploration of the flax genepools of these areas in order to capture the genetic diversity of the crop for breeding or other practical applications. This also applies to conservationists who must design sampling and conservation strategies to maintain and augment the genepools both *in situ* and in the gene banks. This is also expected to guide policy makers, trans-governmental and international institutions because it flags the potential of wrongful or incomplete interpretations of crops arising from data derived from core collections.

The genotype-phenotype association analyses have detected a number of loci underlying the various traits. All the loci and their linked candidate genes can be cross-referenced with findings of future similar studies on flax or other crops. The large-effect loci detected by multiple models and harboring plausible candidate gene/s can be readily used in marker-assisted selection upon validation. Overall, this study has generated important information for a wide-range of direct and indirect beneficiaries from policy makers to scientists.

5.4. Conclusion

Flax, an ancient fiber and oilseed crop is a versatile plant. Despite the fact that the crop has recently gained attention, especially for its nutritional and medical attributes mainly connected to its high plant based omega-3 content and other essential fatty acids, its production has been constrained by low yields which have driven growers to shift to superior-performing crops. Environmental factors such as moisture conditions are limiting flax productivity. Exploring the flax genepool for local adaptation along its eco-geographic gradients and identification of genetic variants for better performance under limiting conditions, such as moisture stress, are important. This study generated relevant information about the genetic variations and patterns

of haplotype distribution in flax across its major growing regions of the world. The genetic structure of the flax core collection could be assigned to four major groups: Temperate, South Asian, Abyssinian and Mediterranean. The Temperate group, that included Eurasia and the New Worlds (Americas and Oceania), was dominated by closely related genotypes suggesting a high gene flow between these regions via recent demographic movements. This study also revealed that the remaining three groups, which are presumed to represent the flax centers of diversity and ancient growing regions with higher genetic variation harbored a high concentration of haplotype endemism. In-depth explorations of these regions to capture further genetic variation is therefore warranted.

Genomic regions identified by the genotype-trait association analyses in this study are important outcomes with implications for further studies and/or applications upon validation. In the GWAS for early shoot-root (chapter 2) and drought tolerant related traits (chapter 3) , genes, previously identified as modulators of the large effect QTL for many traits, were identified as candidate genes for the traits, lending credence to the GWAS results. Some of these genes, especially those associated with large effect QTL, can be considered as candidates for their respective traits. Hence, further validation of these candidates is justified for their practical applications in breeding including in marker-assisted and in genomic selections.

6. References

Ackerly DD (2004) Adaptation, niche conservatism, and convergence: comparative studies of leaf evolution in the California chaparral. Am Nat 163:654-671
Aharoni A, Dixit S, Jetter R, Thoenes E, van Arkel G, Pereira A (2004) The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in Arabidopsis. Plant Cell 16:2463-2480
Ahlawat I, Gangaiah Bd (2003) Effect of land configuration and irrigation on sole and linseed (*Linum usitatissimum*) intercropped chickpea (*Cicer arietinum*). Indian J Agric Sci 80:250-253
Alessi J, Power J (1970) Influence of row spacing, irrigation, and weeds on dryland flax yield, quality, and water use. Agron J 62:635-637
Alexander DH, Lange K (2011) Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC Bioinformatics 12:246
Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655-1664
Allaby RG, Peterson GW, Merriwether DA, Fu YB (2005) Evidence of the domestication history of flax (*Linum usitatissimum* L.) from genetic diversity of the *sad2* locus. Theor Appl Genet 112:58-65
Amani I, Fischer R, Reynolds M (1996) Canopy temperature depression association with yield of irrigated spring wheat cultivars in a hot climate. J Agron Crop Sci 176:119-129
Ambika S, Manonmani V, Somasundaram G (2014) Review on effect of seed size on seedling vigour and seed yield. Res J Seed Sci 7:31-38
Araus JL, Cairns JE (2014) Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci 19:52-61
Araus JL, Serret MD, Edmeades GO (2012) Phenotyping maize for adaptation to drought. Front Physiol 3:305
Araus JL, Slafer GA, Royo C, Serret MD (2008) Breeding for yield potential and stress adaptation in cereals. Crit Rev Plant Sci 27:377-412
Ariel FD, Diet A, Crespi M, Chan RL (2010) The LOB-like transcription factor MtLBD1 controls *Medicago truncatula* root architecture under salt stress. Plant Signal Behav 5:1666-1668
Asseng S, Ewert F, Martre P, Rötter RP, Lobell DB, Cammarano D, Kimball B, Ottman MJ, Wall G, White JW (2015) Rising temperatures reduce global wheat production. Nat Clim 5:143
Atkinson JA, Pound MP, Bennett MJ, Wells DM (2019) Uncovering the hidden half of plants using new advances in root phenotyping. Curr Opin Biotechnol 55:1-8
Atkinson JA, Wingen LU, Griffiths M, Pound MP, Gaju O, Foulkes MJ, Le Gouis J, Griffiths S, Bennett MJ, King J (2015) Phenotyping pipeline reveals major seedling root growth QTL in hexaploid wheat. J Exp Bot 66:2283-2292
Augé RM (2000) Stomatal behavior of arbuscular mycorrhizal plants. In: Kapulni YDD (ed) Arbuscular mycorrhizas: physiology and function. Kluwer, Dordrecht, The Netherlands, pp 201-237
Baas S, Trujillo M, Lombardi N (2015) Impact of disasters on agriculture and food security. FAO
Badigannavar A, Teme N, de Oliveira AC, Li G, Vaksmann M, Viana VE, Ganapathi T, Sarsu F (2018) Physiological, genetic and molecular basis of drought resilience in sorghum [*Sorghum bicolor* (L.) Moench]. Indian J Plant Physiol 23:670-688
Bänziger M, Edmeades GO, Beck D, Bellon MR (2000) Breeding for drought and nitrogen stress tolerance in maize: from theory to practice. CIMMYT, DF, CIMMYT, Mexico
Barbez E, Kubeš M, Rolčík J, Béziat C, Pěnčík A, Wang B, Rosquete MR, Zhu J, Dobrev PI, Lee Y (2012) A novel putative auxin carrier family regulates intracellular auxin homeostasis in plants. Nature 485:119

References

Barnes R (2004) Textiles in Indian Ocean Societies. RoutledgeCurzon London, UK
Basu S, Ramegowda V, Kumar A, Pereira A (2016) Plant adaptation to drought stress. FResearch 5:F1000 Faculty Rev-1554
Batool A, Akram N, LV G-C, Xong J, Tian T, Wang J, Cheng Z-G, Wen Y, Xiong Y (2018) Root-to-shoot communication and its signal cross talk in plants: a physiological and agronomic prospective. Pak J Bot 50:2059-2067
Bauhus J, Messier C (1999) Soil exploitation strategies of fine roots in different tree species of the southern boreal forest of eastern Canada. Can J For Res 29:260-273
Behringer C, Schwechheimer C (2015) B-GATA transcription factors–insights into their structure, regulation, and role in plant development. Front Plant Sci 6:90
Bell AR, Nalewaja JD (1968) Competitive effects of wild oat in flax. Weed Sci 16:501-504
Benfey PN, Linstead PJ, Roberts K, Schiefelbein JW, Hauser M-T, Aeschbacher RA (1993) Root development in Arabidopsis: four mutants with dramatically altered root morphogenesis. Development 119:57-70
Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc Ser B (Stat Method) 57:289-300
Bhaskara GB, Wen T-N, Nguyen TT, Verslues PE (2017) Protein phosphatase 2Cs and microtubule-associated stress protein 1 control microtubule stability, plant growth, and drought response. Plant Cell 29:169-191
Bi H, Shi J, Kovalchuk N, Luang S, Bazanova N, Chirkova L, Zhang D, Shavrukov Y, Stepanenko A, Tricker P (2018) Overexpression of the TaSHN1 transcription factor in bread wheat leads to leaf surface modifications, improved drought tolerance, and no yield penalty under controlled growth conditions. Plant, Cell Environ 41:2549-2566
Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I, Friml J, Heidstra R, Aida M, Palme K, Scheres B (2005) The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature 433:39
Blum A (2005) Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? Aust J Agric Res 56:1159-1168
Blum A (2011a) Drought resistance–is it really a complex trait? Funct Plant Biol 38:753-757
Blum A (2011b) Plant water relations, plant stress and plant production. Springer, New York, USA
Bolle C (2016) Functional aspects of GRAS family proteins. In: Gonzalez DH (ed) Plant Transcription Factors. Elsevier, London, pp 295-311
Brachi B, Morris GP, Borevitz J (2011) Genome-wide association studies in plants: the missing heritability is in the field. Genome Biol 12:232
Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635
Brodie A, Azaria JR, Ofran Y (2016) How far from the SNP may the causative genes be? Nucleic Acids Res 44:6046-6054
Bush WS, Moore JH (2012) Genome-wide association studies. PLoS Comp Biol 8:e1002822
Canto CDLF, Kalogiros D, Ptashnyk M, George T, Waugh R, Bengough A, Russell J, Dupuy L (2018) Morphological and genetic characterisation of the root system architecture of selected barley recombinant chromosome substitution lines using an integrated phenotyping approach. J Theor Biol 447:84-97
Carter J (1993) Potential of flaxseed and flaxseed oil in baked goods and other products in human nutrition. Cereal Foods World 38:753-759
Carvalho AF, Gibaja JF, Cardoso JL (2013) Insights into the earliest agriculture of Central Portugal: sickle implements from the Early Neolithic site of Cortiçóis (Santarém). Comptes Rendus Palevol 12:31-43

References

Casa R, Russell G, Cascio BL, Rossini F (1999) Environmental effects on linseed (*Linum usitatissimum* L.) yield and growth of flax at different stand densities. Eur J Agron 11:267-278

Cattivelli L, Rizza F, Badeck F-W, Mazzucotelli E, Mastrangelo AM, Francia E, Mare C, Tondelli A, Stanca AM (2008) Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. Field Crop Res 105:1-14

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

Cha J-Y, Kim W-Y, Kang SB, Im Kim J, Baek D, Jung IJ, Kim MR, Li N, Kim H-J, Nakajima M (2015) A novel thiol-reductase activity of Arabidopsis YUC6 confers drought tolerance independently of auxin biosynthesis. Nat Commun 6:8041

Challa S, Neelapu NR (2018) Genome-wide association studies (GWAS) for abiotic stress tolerance in plants. In: Wani S (ed) Biochemical, physiological and molecular avenues for combating abiotic stress tolerance in plants. Elsevier, San Diego, USA, pp 135-150

Chandrawati, Maurya R, Singh P, Ranade S, Yadav HK (2014) Diversity analysis in Indian genotypes of linseed (*Linum usitatissimum* L.) using AFLP markers. Gene 549:171-178

Chen J, Stavro PM, Thompson LU (2002) Dietary flaxseed inhibits human breast cancer growth and metastasis and downregulates expression of insulin-like growth factor and epidermal growth factor receptor. Nutr Cancer 43:187-192

Chen N, Veerappan V, Abdelmageed H, Kang M, Allen RD (2018) HSI2/VAL1 silences AGL15 to regulate the developmental transition from seed maturation to vegetative growth in Arabidopsis. Plant Cell 30:600-619

Coates JC, Laplaze L, Haseloff J (2006) Armadillo-related proteins promote lateral root development in Arabidopsis. Proc Natl Acad Sci USA 103:1621-1626

Comas L, Becker S, Cruz VMV, Byrne PF, Dierig DA (2013) Root traits contributing to plant productivity under drought. Front Plant Sci 4:442

Cooper B, Clarke JD, Budworth P, Kreps J, Hutchison D, Park S, Guimil S, Dunn M, Luginbühl P, Ellero C (2003) A network of rice genes associated with stress response and seed development. Proc Natl Acad Sci USA 100:4945-4950

Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de los Campos G, Burgueño J, González-Camacho JM, Pérez-Elizalde S, Beyene Y (2017) Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci 22:961-975

Cui H, Levesque MP, Vernoux T, Jung JW, Paquette AJ, Gallagher KL, Wang JY, Blilou I, Scheres B, Benfey PN (2007) An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. Science 316:421-425

Cui X-Y, Du Y-T, Fu J-d, Yu T-F, Wang C-T, Chen M, Chen J, Ma Y-Z, Xu Z-S (2018) Wheat CBL-interacting protein kinase 23 positively regulates drought stress and ABA responses. BMC Plant Biol 18:93

Dana S, Herdean A, Lundin B, Spetea C (2016) Retracted: Each of the chloroplast potassium efflux antiporters affects photosynthesis and growth of fully developed Arabidopsis rosettes under short-day photoperiod. Physiol Plant 158:483-491

Dar WD, Laxmipathi Gowda CL (2013) Declining agricultural productivity and global food security. J Crop Improv 27:242-254

Davey J (2017) Ancient Egypt: Mummification and burial sites as historical resources. Agora 52:20

References

Davidson R (1969) Effect of root/leaf temperature differentials on root/shoot ratios in some pasture grasses and clover. Ann Bot 33:561-569

De Haan H (1952) Flax breeding and flax varieties in the Netherlands. Euphytica 1:212-218

Delzon S (2015) New insight into leaf drought tolerance. Funct Ecol 29:1247-1249

Dereeper A, Nicolas S, Le Cunff L, Bacilieri R, Doligez A, Peros JP, Ruiz M, This P (2011) SNiPlay: a web-based tool for detection, management and analysis of SNPs. Application to grapevine diversity projects. BMC Bioinformatics 12:134

Deshmukh R, Sonah H, Patil G, Chen W, Prince S, Mutava R, Vuong T, Valliyodan B, Nguyen HT (2014) Integrating omic approaches for abiotic stress tolerance in soybean. Front Plant Sci 5:244

Diederichsen A (2001) Comparison of genetic diversity of flax (*Linum usitatissimum* L.) between Canadian cultivars and a world collection. Plant breeding 120:360-362

Diederichsen A, Fu Y-B (2006) Phenotypic and molecular (RAPD) differentiation of four infraspecific groups of cultivated flax (*Linum usitatissimum* L. subsp. usitatissimum). Genet Resour Crop Evol 53:77-90

Diederichsen A, Kusters PM, Kessler D, Bainas Z, Gugel RK (2013) Assembling a core collection from the flax world collection maintained by Plant Gene Resources of Canada. Genet Resour Crop Evol 60:1479-1485

Diederichsen A, Richards K (2003) Cultivated flax and the genus *Linum* L.: Taxonomy and germplasm conservation. In: Muir A, Weir BS (eds) Flax: the genus *Linum*. Taylor & Francis London, UK pp 23-54

Diederichsen A, Rozhmina T, Zhuchenko A, Richards K (2006) Screening for broad adaptation in 96 flax (*Linum usitatissimum* L.) accessions under dry and warm conditions in Canada and Russia. Plant Genet Resour Newsl 146:7-14

Diederichsen A, Rozhmina TA, Kudrjavceva LP (2008) Variation patterns within 153 flax (*Linum usitatissimum* L.) genebank accessions based on evaluation for resistance to fusarium wilt, anthracnose and pasmo. Plant Genet Resour 6:22-32

Diener AC, Gaxiola RA, Fink GR (2001) Arabidopsis ALF5, a multidrug efflux transporter gene family member, confers resistance to toxins. Plant Cell 13:1625-1638

Dillman A (1938) Natural crossing in flax. Agron J 30:279-286

Dillman AC (1953) Classification of flax [linseed] varieties 1946. Tech Bull US Dep Agric. US Department of Agriculture, Washington DC, USA, p 55

Dimitrov I, Tax F (2018) Lateral root growth in Arabidopsis is controlled by short and long distance signaling through the LRR RLKs XIP1/CEPR1 and CEPR2. Plant Signal Behav 13:e1489667

Domantovich A, Koshkin V, Brutch N, Matvienko I (2012) Investigation of photoperiod sensitivity of *Linum usitatissimum* L. Lines and effect of short-day conditions on their economically valuable traits. Russ Agric Sci 38:173-177

Dossa K, Li D, Zhou R, Yu J, Wang L, Zhang Y, You J, Liu A, Mmadi MA, Fonceka D (2019) The genetic basis of drought tolerance in the high oil crop *Sesamum indicum*. Plant Biotechnol J 17:1788-1803

Downie HF, Adu M, Schmidt S, Otten W, Dupuy LX, White P, Valentine TA (2015) Challenges and opportunities for quantifying roots and rhizosphere interactions through imaging and image analysis. Plant, Cell Environ 38:1213-1232

Dracup JA, Lee KS, Paulson EG (1980) On the definition of droughts. Water Resour Res 16:297-302

Du G, Qi W (2010) Trade-offs between flowering time, plant height, and seed size within and across 11 communities of a QingHai-Tibetan flora. Plant Ecol 209:321-333

References

Du L, Li N, Chen L, Xu Y, Li Y, Zhang Y, Li C, Li Y (2014) The ubiquitin receptor DA1 regulates seed and organ size by modulating the stability of the ubiquitin-specific protease UBP15/SOD2 in Arabidopsis. Plant Cell 26:665-677

Dugassa G, Von Alten H, Schönbeck F (1996) Effects of arbuscular mycorrhiza (AM) on health of *Linum usitatissimum* L. infected by fungal pathogens. Plant Soil 185:173-182

El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MG (2014) Genotype× environment interaction QTL mapping in plants: lessons from Arabidopsis. Trends Plant Sci 19:390-398

Ellis R (1992) Seed and seedling vigour in relation to crop growth and yield. Plant Growth Regul 11:249-255

Engels JM, Hawkes JG, Worede M (1991) Plant genetic resources of Ethiopia. Cambridge University Press, New York

FAO (2009) How to feed the world in 2050: Global agriculture towards 2050; High level expert forum FAO, Rome p4

Farfan ID, De La Fuente GN, Murray SC, Isakeit T, Huang PC, Warburton M, Williams P, Windham GL, Kolomiets M (2015) Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics. PLoS One 10:e0117737

Farooq M, Hussain M, Wahid A, Siddique K (2012) Drought stress in plants: An overview. In: Aroca R (ed) Plant responses to drought stress: From morphological to molecular features. Springer, Berlin, Germany pp 1-37

Federer WT (1956) Augmented (or Hookuiaku) design. Hawaii Plr Rec 55:18

Fernandez GC (1992) Effective selection criteria for assessing plant stress tolerance. Proceeding of the International Symposium on Adaptation of Vegetables and other Food Crops in Temperature and Water Stress, Aug 13-16, Shanhua, Taiwan, 1992, pp 257-270

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol 37:4302-4315

Finch-Savage WE, Bassel GW (2015) Seed vigour and crop establishment: extending performance beyond adaptation. J Exp Bot 67:567-591

Fischer K, Johnson E, Edmeades G (1982) Breeding and selection for drought tolerance in tropical maize. In: John C (ed) Symposium on principles and methods in crop improvement for drought resistance with emphasis on rice, IRRI, Philippines. IRRI, Los Baños, Philippines, pp 378-399

Fischer R, Maurer R (1978) Drought resistance in spring wheat cultivars. I. Grain yield responses. Aust J Agric Res 29:897-912

Fischer R, Rees D, Sayre K, Lu Z-M, Condon A, Saavedra AL (1998) Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. Crop Sci 38:1467-1475

Foley JA, Ramankutty N, Brauman KA, Cassidy ES, Gerber JS, Johnston M, Mueller ND, O'Connell C, Ray DK, West PC, Balzer C, Bennett EM, Carpenter SR, Hill J, Monfreda C, Polasky S, Rockstrom J, Sheehan J, Siebert S, Tilman D, Zaks DP (2011) Solutions for a cultivated planet. Nature 478:337-342

Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic structure of populations. Genetics 174:875–891

Foster R, Pooni H, Mackay I (1998) The impact of water deprivation on the performance of *Linum usitatissimum* cultivars. J Genet Breed 52:63-71

Frichot E, François O (2015) LEA: an R package for landscape and ecological association studies. Methods Ecol Evol 6:925-929

Frichot E, Mathieu F, Trouillon T, Bouchard G, François O (2014) Fast and efficient estimation of individual ancestry coefficients. Genetics 196:973-983

References

Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. Mol Biol Evol 30:1687-1699

Friedli CN, Abiven S, Fossati D, Hund A (2019) Modern wheat semi-dwarfs root deep on demand: response of rooting depth to drought in a set of Swiss era wheats covering 100 years of breeding. Euphytica 215:85

Fu Y-B (2005) Geographic patterns of RAPD variation in cultivated flax. Crop Sci 45:1084-1091

Gabiana CP (2005) Response of linseed (*Linum usitatissimum* L.) to irrigation, nitrogen and plant population. Lincoln University, Christchurch, New Zealand

Galkovskyi T, Mileyko Y, Bucksch A, Moore B, Symonova O, Price CA, Topp CN, Iyer-Pascuzzi AS, Zurek PR, Fang S, Harer J, Benfey PN, Weitz JS (2012) GiA Roots: software for the high throughput analysis of plant root system architecture. BMC Plant Biol 12:116

Gallego-Giraldo L, Escamilla-Trevino L, Jackson LA, Dixon RA (2011) Salicylic acid mediates the reduced growth of lignin down-regulated plants. Proc Natl Acad Sci USA 108:20814-20819

Gao M-J, Parkin I, Lydiate D, Hannoufa A (2004) An auxin-responsive SCARECROW-like transcriptional activator interacts with histone deacetylase. Plant Mol Biol 55:417-431

Garnier E (1991) Resource capture, biomass allocation and growth in herbaceous plants. Trends Ecol Evol 6:126-131

Gebauer SK, Psota TL, Harris WS, Kris-Etherton PM (2006) n-3 fatty acid dietary recommendations and food sources to achieve essentiality and cardiovascular benefits. Am J Clin Nutr 83:1526S-1535S

Geleta M, Asfaw Z, Bekele E, Teshome A (2002) Edible oil crops and their integration with the major cereals in North Shewa and South Welo, Central Highlands of Ethiopia: an ethnobotanical perspective. Hereditas 137:29-40

Gervers M (1990) Cotton and cotton weaving in Meroitic Nubia and medieval Ethiopia. Text Hist 21:13-30

Ghelli R, Brunetti P, Napoli N, De Paolis A, Cecchetti V, Tsuge T, Serino G, Matsui M, Mele G, Rinaldi G (2018) A newly identified flower-specific splice variant of AUXIN RESPONSE FACTOR8 regulates stamen elongation and endothecium lignification in Arabidopsis. Plant Cell 30:620-637

Gimeno-Gilles C, Gervais M-L, Planchet E, Satour P, Limami AM, Lelievre E (2011) A stress-associated protein containing A20/AN1 zing-finger domains expressed in *Medicago truncatula* seeds. Plant Physiol Biochem 49:303-310

Golldack D, Li C, Mohan H, Probst N (2014) Tolerance to drought and salt stress in plants: unraveling the signaling networks. Front Plant Sci 5:151

Gomez C, Terrier N, Torregrosa L, Vialet S, Fournier-Level A, Verries C, Souquet J-M, Mazauric J-P, Klein M, Cheynier V (2009) Grapevine MATE-type proteins act as vacuolar H+-dependent acylated anthocyanin transporters. Plant Physiol 150:402-415

González LG, Deyholos MK (2012) Identification, characterization and distribution of transposable elements in the flax (*Linum usitatissimum* L.) genome. BMC Genomics 13:644

Gowda VR, Henry A, Yamauchi A, Shashidhar H, Serraj R (2011) Root biology and genetic improvement for drought avoidance in rice. Field Crops Res 122:1-13

Goyal A, Sharma V, Upadhyay N, Gill S, Sihag M (2014) Flax and flaxseed oil: an ancient medicine & modern functional food. J Food Sci Technol 51:1633-1653

Gray SF (1988) Linaceae. In: Hickey M, King CJ (eds) 100 Families of Flowering Plants. Cambridge University Press New York pp 305-310

Green A, Marshall D (1984) Isolation of induced mutants in linseed (*Linum usitatissimum*) having reduced linolenic acid content. Euphytica 33:321-328

Gruenhagen R, Nalewaja JD (1969) Competition between flax and wild buckwheat. Weed Sci 17:380-384

References

Guo R, Hao W, Gong D (2012) Effects of water stress on germination and growth of linseed seedlings (*Linum usitatissimum* L), photosynthetic efficiency and accumulation of metabolites. J Agr Sci 4:253-265

Guo X, Chen G, Cui B, Gao Q, Guo J-E, Li A, Zhang L, Hu Z (2016) *Solanum lycopersicum* agamous-like MADS-box protein AGL15-like gene, SlMBP11, confers salt stress tolerance. Mol Breed 36:125

Guo Z, Yang W, Chang Y, Ma X, Tu H, Xiong F, Jiang N, Feng H, Huang C, Yang P (2018) Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. Mol Plant 11:789-805

Gupta PK, Kulwal PL, Jaiswal V (2019) Association mapping in plants in the post-GWAS genomics era. Adv Genet 104:75-154

Gutaker RM, Zaidem M, Fu Y-B, Diederichsen A, Smith O, Ware R, Allaby RG (2019) Flax latitudinal adaptation at LuTFL1 altered architecture and promoted fiber production. Sci Rep 9:976

Hajduch M, Hearne LB, Miernyk JA, Casteel JE, Joshi T, Agrawal GK, Song Z, Zhou M, Xu D, Thelen JJ (2010) Systems analysis of seed filling in Arabidopsis: using general linear modeling to assess concordance of transcript and protein expression. Plant Physiol 152:2078-2087

Hall LM, Booker H, Siloto RM, Jhala AJ, Weselake RJ (2016) Flax (*Linum usitatissimum* L.). Industrial Oil Crops. Elsevier, pp 157-194

Han S, Wang C-W, Jiang J (2014) Mitogen-activated protein kinase 6 controls root growth in Arabidopsis by modulating Ca2+-based Na+ flux in root cell under salt stress. J Plant Physiol 171:26-34

Hao Z, AghaKouchak A, Nakhjiri N, Farahmand A (2014) Global integrated drought monitoring and prediction system. Sci Data 1:140001

Harvey JJ, Strauss SH (2009) Towards physiological sculpture of plants. New Phytol 181:8-12

He J, Meng S, Zhao T, Xing G, Yang S, Li Y, Guan R, Lu J, Wang Y, Xia Q (2017) An innovative procedure of genome-wide association analysis fits studies on germplasm population and plant breeding. Theor Appl Genet 130:2327-2343

He L, Xiao J, Rashid KY, Yao Z, Li P, Jia G, Wang X, Cloutier S, You FM (2018) Genome-wide association studies for pasmo resistance in flax (*Linum usitatissimum* L.). Front Plant Sci 9:1982

Health-Canada (2014) Summary of health Canada's assessment of a health claim about ground whole flaxseed and blood cholesterol lowering. Bureau of Nutritional Sciences; Food Directorate, Health Products and Food Branch, Health Canada, p 5

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59:1633-1638

Henry A, Gowda VR, Torres RO, McNally KL, Serraj R (2011) Variation in root system architecture and drought response in rice (*Oryza sativa*): phenotyping of the OryzaSNP panel in rainfed lowland fields. Field Crops Res 120:205-214

Herbig C, Maier U (2011) Flax for oil or fibre? Morphometric analysis of flax seeds and new aspects of flax cultivation in Late Neolithic wetland settlements in southwest Germany. Veg Hist Archaeobot 20:527

Heywood VH, Moore D, Dunkley J, King C (1978) Flowering plants of the world. Oxford University Press, Oxford, UK

Hijmans RJ, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. Plant Genet Resour Newsl 127:15-19

Hill W, Robertson A (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38:226-231

Hochholdinger F, Yu P, Marcon C (2018) Genetic control of root system development in maize. Trends Plant Sci 23:79-88

Hocking P (1995) Effects of nitrogen supply on the growth, yield components, and distribution of nitrogen in Linola. J Plant Nutr 18:257-275

Hocking P, Kirkegaard J, Angus J, Gibson A, Koetz E (1997) Comparison of canola, Indian mustard and Linola in two contrasting environments. I. Effects of nitrogen fertilizer on dry-matter production, seed yield and seed quality. Field Crops Res 49:107-125
Hodge A (2004) The plastic plant: root responses to heterogeneous supplies of nutrients. New Phytol 162:9-24
Hoecker U, Tepperman JM, Quail PH (1999) SPA1, a WD-repeat protein specific to phytochrome A signal transduction. Science 284:496-499
Huang G, Li Ch, Li Y (2018a) Phenological responses to nitrogen and water addition are linked to plant growth patterns in a desert herbaceous community. Ecol Evol 8:5139-5152
Huang X, Bao Y, Wang B, Liu L, Chen J, Dai L, Peng D (2016) Identification and expression of Aux/IAA, ARF, and LBD family transcription factors in Boehmeria nivea. Biol Plant 60:244-250
Huang X, Han B (2014) Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol 65:531-551
Huang X, Hou L, Meng J, You H, Li Z, Gong Z, Yang S, Shi Y (2018b) The antagonistic action of abscisic acid and cytokinin signaling mediates drought stress response in Arabidopsis. Mol Plant 11:970-982
Hund A, Trachsel S, Stamp P (2009) Growth of axile and lateral roots of maize: I development of a phenotying platform. Plant Soil 325:335-349
Hung HY, Shannon LM, Tian F, Bradbury PJ, Chen C, Flint-Garcia SA, McMullen MD, Ware D, Buckler ES, Doebley JF, Holland JB (2012) ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc Natl Acad Sci USA 109:E1913-1921
Hunter MC, Smith RG, Schipanski ME, Atwood LW, Mortensen DA (2017) Agriculture in 2050: Recalibrating targets for sustainable intensification. Bioscience 67:386-391
Inoue H, Li M, Schnell DJ (2013) An essential role for chloroplast heat shock protein 90 (Hsp90C) in protein import into chloroplasts. Proc Natl Acad Sci USA 110:3173-3178
Irshad M, Canut H, Borderies G, Pont-Lezica R, Jamet E (2008) A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. BMC Plant Biol 8:94
Itoh H, Nonoue Y, Yano M, Izawa T (2010) A pair of floral regulators sets critical day length for Hd3a florigen expression in rice. Nat Genet 42:635-638
Jackson R, Manwaring J, Caldwell M (1990) Rapid physiological adjustment of roots to localized soil enrichment. Nature 344:58
Jain M, Khurana JP (2009) Transcript profiling reveals diverse roles of auxin-responsive genes during reproductive development and abiotic stress in rice. FEBS J 276:3148-3162
Jenkins K, Warren R (2015) Drought-damage functions for the estimation of drought costs under future projections of climate change. J Extreme Events 2:1550001
Jeon E, Young Kang N, Cho C, Joon Seo P, Chung Suh M, Kim J (2017) LBD14/ASL17 positively regulates lateral root formation and is involved in ABA response for root architecture in Arabidopsis. Plant Cell Physiol 58:2190-2201
Jia H, Hao L, Guo X, Liu S, Yan Y, Guo X (2016) A Raf-like MAPKKK gene, GhRaf19, negatively regulates tolerance to drought and salt and positively regulates resistance to cold stress by modulating reactive oxygen species in cotton. Plant Sci 252:267-281
Jiang W-B, Huang H-Y, Hu Y-W, Zhu S-W, Wang Z-Y, Lin W-H (2013) Brassinosteroid regulates seed size and shape in Arabidopsis. Plant Physiol 162:1965-1977
Josephs EB, Stinchcombe JR, Wright SI (2017) What can genome-wide association studies tell us about the evolutionary forces maintaining genetic variation for quantitative traits? New Phytol 214:21-33
Jost L (2008) GST and its relatives do not measure differentiation. Mol Ecol 17:4015-4026

Jung H, Lee D-K, Do Choi Y, Kim J-K (2015) OsIAA6, a member of the rice Aux/IAA gene family, is involved in drought tolerance and tiller outgrowth. Plant Sci 236:304-312
Kamiya N, Itoh JI, Morikami A, Nagato Y, Matsuoka M (2003) The SCARECROW gene's role in asymmetric cell divisions in rice plants. Plant J 36:45-54
Kannangara R, Branigan C, Liu Y, Penfield T, Rao V, Mouille G, Höfte H, Pauly M, Riechmann JL, Broun P (2007) The transcription factor WIN1/SHN1 regulates cutin biosynthesis in *Arabidopsis thaliana*. Plant Cell 19:1278-1294
Kanneganti V, Gupta AK (2008) Overexpression of OsiSAP8, a member of stress associated protein (SAP) gene family of rice confers tolerance to salt, drought and cold stress in transgenic tobacco and rice. Plant Mol Biol 66:445-462
Kar G, Kumar A, Martha M (2007) Water use efficiency and crop coefficients of dry season oilseed crops. Agric Water Manage 87:73-82
Kariuki LW, Masinde P, Githiri S, Onyango AN (2016) Effect of water stress on growth of three linseed (*Linum usitatissimum* L.) varieties. SpringerPlus 5:759
Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 4:782-788
Kell DB (2011) Breeding crop plants with deep roots: their role in sustainable carbon, nutrient and water sequestration. Ann Bot 108:407-418
Khayatnezhad M, Gholamin R (2010) Study of drought tolerance of maize genotypes using the stress tolerance index. Am Eur J Agric Environ Sci 9:359-363
Kim J, Jung JH, Lee SB, Go YS, Kim HJ, Cahoon R, Markham JE, Cahoon EB, Suh MC (2013) Arabidopsis 3-ketoacyl-coenzyme a synthase9 is involved in the synthesis of tetracosanoic acids as precursors of cuticular waxes, suberins, sphingolipids, and phospholipids. Plant Physiol 162:567-580
Klimek-Kopyra A, Glab T, Zajac T, Stoklosa A, Kulig B (2015) Vertical distribution of the root system of linseed (*Linum usitatissimum* L.) and legumes in pure and mixed sowing. Acta Agrobot 68:43-52
Kogan F, Adamenko T, Guo WJRSL (2013) Global and regional drought dynamics in the climate warming era. Remote Sens Lett 4:364-372
Kooyers N (2015) The evolution of drought escape and avoidance in natural herbaceous populations. Plant Sci 234:155-162
Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour 15:1179-1191
Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant methods 9:29
Kramer-Walter KR, Bellingham PJ, Millar TR, Smissen RD, Richardson SJ, Laughlin DC (2016) Root traits are multidimensional: specific root length is independent from root tissue density and the plant economic spectrum. J Ecol 104:1299-1310
Kravić N, Šukalović VH-T, Babić V, Srdić J, Mesarović J, Anđelković V (2018) Maize seedling performance as a potential index for drought tolerance. Arch Biol Sci 70:167-177
Kumar S, You FM, Cloutier S (2012) Genome wide SNP discovery in flax through next generation sequencing of reduced representation libraries. BMC Genomics 13:684
Kunz HH, Gierth M, Herdean A, Satoh-Cruz M, Kramer DM, Spetea C, Schroeder JI (2014) Plastidial transporters KEA1, -2, and -3 are essential for chloroplast osmoregulation, integrity, and pH regulation in Arabidopsis. Proc Natl Acad Sci USA 111:7480-7485
Kvavadze E, Bar-Yosef O, Belfer-Cohen A, Boaretto E, Jakeli N, Matskevich Z, Meshveliani T (2009) 30,000-year-old wild flax fibers. Science 325:1359-1359

References

Laluk K, AbuQamar S, Mengiste T (2011) The Arabidopsis mitochondria-localized pentatricopeptide repeat protein PGN functions in defense against necrotrophic fungi and abiotic stress tolerance. Plant Physiol 156:2053-2068

Lara MV, Andreo CS, adaptations (2011) C4 plants adaptation to high levels of CO2 and to drought environments. In: Shanker A, Venkateswarlu B (eds) Abiotic stress in plants - Mechanisms and Adaptations. InTech, University Campus STeP Ri , Rijeka, Croatia pp 415-428

Laubinger S, Hoecker U (2003) The SPA1-like proteins SPA3 and SPA4 repress photomorphogenesis in the light. Plant J 35:373-385

Laubinger S, Marchal V, Gentilhomme J, Wenkel S, Adrian J, Jang S, Kulajta C, Braun H, Coupland G, Hoecker U (2006) Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. Development 133:3213-3222

Lee HW, Cho C, Pandey SK, Park Y, Kim M-J, Kim J (2019) LBD16 and LBD18 acting downstream of ARF7 and ARF19 are involved in adventitious root formation in Arabidopsis. BMC Plant Biol 19:46

Lee J-H, Terzaghi W, Gusmaroli G, Charron J-BF, Yoon H-J, Chen H, He YJ, Xiong Y, Deng XW (2008) Characterization of Arabidopsis and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. Plant Cell 20:152-167

Lee J, Burns TH, Light G, Sun Y, Fokar M, Kasukabe Y, Fujisawa K, Maekawa Y, Allen RD (2010) Xyloglucan endotransglycosylase/hydrolase genes in cotton and their role in fiber elongation. Planta 232:1191-1205

Lesk C, Rowhani P, Ramankutty N (2016) Influence of extreme weather disasters on global crop production. Nature 529:84

Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242-245

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

Li L, He Z, Pandey GK, Tsuchiya T, Luan S (2002) Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. J Biol Chem 277:5360-5368

Li N, Li Y (2014) Ubiquitin-mediated control of seed size in plants. Front Plant Sci 5:332

Li N, Meng H, Xing H, Liang L, Zhao X, Luo K (2017) Genome-wide analysis of MATE transporters and molecular characterization of aluminum resistance in Populus. J Exp Bot 68:5669-5683

Li Y, Zheng L, Corke F, Smith C, Bevan MW (2008) Control of final seed and organ size by the DA1 gene family in *Arabidopsis thaliana*. Genes Dev 22:1331-1336

Li Z, Shen J, Liang J (2019) Genome-Wide Identification, Expression Profile, and Alternative Splicing Analysis of the Brassinosteroid-Signaling Kinase (BSK) Family Genes in Arabidopsis. Int J Mol Sci 20:1138

Lin CS, Poushnisky G (1985) A modified augmented design (type 2) for rectangular plots. Can J Plant Sci 65:743-749

Lisson S, Mendham N (2000) Agronomic studies of flax (*Linum usitatissimum* L.) in south-eastern Australia. Aust J Exp Agric 40:1101-1112

Liu F-H, Chen X, Long B, Shuai R-Y, Long C-L (2011) Historical and botanical evidence of distribution, cultivation and utilization of *Linum usitatissimum* L.(flax) in China. Veg Hist Archaeobot 20:561

Liu H, Wang S, Yu X, Yu J, He X, Zhang S, Shou H, Wu P (2005) ARL1, a LOB-domain protein required for adventitious root formation in rice. Plant J 43:47-56

Liu J, Zhang F, Zhou J, Chen F, Wang B, Xie X (2012) Phytochrome B control of total leaf area and stomatal density affects drought tolerance in rice. Plant Mol Biol 78:289-300

Liu X (2010) The Silk Road in world history. Oxford University Press, OXford, UK

References

Liu Z, Zhu Y, Gao J, Yu F, Dong A, Shen WH (2009) Molecular and reverse genetic characterization of NUCLEOSOME ASSEMBLY PROTEIN1 (NAP1) genes unravels their function in transcription and nucleotide excision repair in *Arabidopsis thaliana*. Plant J 59:27-38
Lloyd-Hughes B (2013) The impracticality of a universal drought definition. Theor Appl Climatol 117:607-611
Lopes MS, Reynolds MP (2010) Partitioning of assimilates to deeper roots is associated with cooler canopies and increased yield under drought in wheat. Funct Plant Biol 37:147-156
Louthan AM, Kay KM (2011) Comparing the adaptive landscape across trait types: larger QTL effect size in traits under biotic selection. BMC Evol Biol 11:60
Lu X, Clarke RC (1995) The cultivation and use of hemp (*Cannabis sativ*a L.) in ancient China. J Intl Hemp Assoc 2:26-30
Lüdecke D (2017) Data visualization for statistics in social science. R package https://doi org/105281/zenodo. R, https://CRAN.R-project.org/package=sjPlot.
Luo L, Xia H, Lu B (2019) Crop Breeding for Drought Resistance. Front Plant Sci 10:314
Luo Q, Wei Q, Wang R, Zhang Y, Zhang F, He Y, Zhou S, Feng J, Yang G, He G (2017) BdCIPK31, a calcineurin B-like protein-interacting protein kinase, regulates plant response to drought and salt stress. Front Plant Sci 8:1184
Ma H-S, Liang D, Shuai P, Xia X-L, Yin W-L (2010) The salt-and drought-inducible poplar GRAS protein SCL7 confers salt and drought tolerance in *Arabidopsis thaliana*. J Exp Bot 61:4011-4019
Mahalakshmi V, Bidinger F (1985) Flowering response of pearl millet to water stress during panicle development. Ann Appl Biol 106:571-578
Mangini G, Gadaleta A, Colasuonno P, Marcotuli I, Signorile AM, Simeone R, De Vita P, Mastrangelo AM, Laidò G, Pecchioni N (2018) Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. PloS one 13:e0190162
Manik S, Shi S, Mao J, Dong L, Su Y, Wang Q, Liu H (2015) The calcium sensor CBL-CIPK is involved in plant's response to abiotic stresses. Int J Genomics 2015:10
Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209-220
Marchenkov A, Rozhmina T, Uschapovsky I, Muir AD (2003) Cultivation of flax. In: Muir A, Westcott N (eds) Flax: the genus *Linum*, Taylor & Francis London, pp 74-91
Marques G, Rencoret J, Gutiérrez Suárez A, Río Andrade JCd (2010) Evaluation of the chemical composition of different non-woody plant fibers used for pulp and paper manufacturing. Open Agr J 4:93-101
Martins APB, dos Santos Brito M, Mayer JLS, Llerena JPP, Oliveira JF, Takahashi NG, Carlin SD, Borges DNAF, Andrade LM, Peixoto-Júnior RF (2018) Ectopic expression of sugarcane SHINE changes cell wall and improves biomass in rice. Biomass Bioenergy 119:322-334
Mason S, Galusha T, Kmail Z (2018) Planting Date Influence on Yield of Drought-Tolerant Maize with Different Maturity Classifications. Agron J 110:293-299
Matiu M, Ankerst DP, Menzel A (2017) Interactions between temperature and drought in global and regional crop yield variability during 1961-2014. PloS one 12:e0178339
McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proc Natl Acad Sci USA 106:12273-12278
Medail F, Quezel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. Ann Missouri Bot Gard 84:112-127
Meirmans PG, Hedrick PW (2011) Assessing population structure: FST and related measures. Molecular Ecology Resources 11:5-18

References

Michael TP, VanBuren R (2015) Progress, challenges and the future of crop genomes. Curr Opin Plant Biol 24:71-81

Mills MC, Rahal C (2019) A scientometric review of genome-wide association studies. Commun Biol 2:1-11

Mir RR, Reynolds M, Pinto F, Khan MA, Bhat MA (2019) High-throughput phenotyping for crop improvement in the genomics era. Plant Sci 282:60-72

Mirshekari M, Amiri R, Nezhad HI, Noori SAS, Zandvakili OR (2012) Effects of planting date and water deficit on quantitative and qualitative traits of flax seed. Am Eurasian J Agr Environ Sci 12:901-913

Mishra AK, Singh VP (2010) A review of drought concepts. J Hydrol 391:202-216

Mohanta TK, Bashir T, Hashem A, Abd_Allah EF, Khan AL, Al-Harrasi AS (2018) Molecular players of auxin transport systems: advances in genomic and molecular events. J Plant Interact 13:483-495

Molinier J (2017) Genome and epigenome surveillance processes underlying UV exposure in plants. Genes (Basel) 8:316

Mollasadeghi V, Valizadeh M, Shahryari R, Imani AA (2011) Evaluation of end drought tolerance of 12 wheat genotypes by stress indices. World Appl Sci J 13:545-551

Morillon R, Lassalles J-P (2002) Water deficit during root development: effects on the growth of roots and osmotic water permeability of isolated root protoplasts. Planta 214:392-399

Moro MF, Silva IA, de Araujo FS, Nic Lughadha E, Meagher TR, Martins FR (2015) The role of edaphic environment and climate in structuring phylogenetic pattern in seasonally dry tropical plant communities. PLoS One 10:e0119166

Morris DH (2007) Description and composition of flax. In: Morris DH (ed) Flax: a health and nutrition primer. Flax Council of Canada, Winnipeg, Canada

Mostafavi K (2011) A study effects of drought stress on germination and early seedling growth of flax (*Linum usitatissimum* L.) cultivars. Adv Environ Biol 5:3307-3312

Mueller JC (2004) Linkage disequilibrium for different scales and applications. Brief Bioinform 5:355-364

Muthukumar C, Subathra T, Aiswarya J, Gayathri V, Babu RC (2015) Comparative genome-wide association studies for plant production traits under drought in diverse rice (*Oryza sativa* L.) lines using SNP and SSR markers. Curr Sci India 109:139-147

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21:2194-2202

Nagpal P, Ellis CM, Weber H, Ploense SE, Barkawi LS, Guilfoyle TJ, Hagen G, Alonso JM, Cohen JD, Farmer EE (2005) Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. Development 132:4107-4118

Nakagami H, Pitzschke A, Hirt H (2005) Emerging MAP kinase pathways in plant stress signalling. Trends Plant Sci 10:339-346

Ndour A, Vadez V, Pradal C, Lucas M (2017) Virtual plants need water too: Functional-structural root system models in the context of drought tolerance breeding. Front Plant Sci 8:1577

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann Hum Genet 47:253-259

Nishihama R, Banno H, Shibata W, Hirano K, Nakashima M, Usami S, Machida Y (1995) Plant homologues of components of MAPK (mitogen-activated protein kinase) signal pathways in yeast and animal cells. Plant Cell Physiol 36:749-757

Nobel P, North G (1996) Features of roots of CAM plants. In: Smith J, Winter K (eds) Crassulacean acid metabolism. Springer, Berlin Heidelberg, Germany, pp 266-280

Noh B, Lee S-H, Kim H-J, Yi G, Shin E-A, Lee M, Jung K-J, Doyle MR, Amasino RM, Noh Y-S (2004) Divergent roles of a pair of homologous Jumonji/Zinc-Finger–class transcription factor proteins in the regulation of Arabidopsis flowering time. Plant Cell 16:2601-2613
Ntukamazina N, Onwonga RN, Sommer R, Mukankusi CM, Mburu J, Rubyogo JC (2017) Effect of excessive and minimal soil moisture stress on agronomic performance of bush and climbing bean (Phaseolus vulgaris L.). Cogent Food Agric 3:1373414
Ogburn RM, Edwards EJ (2010) The ecological water-use strategies of succulent plants. Adv Bot Res. Elsevier, pp 179-225
Olivares-Villegas JJ, Reynolds MP, McDonald GK (2007) Drought-adaptive attributes in the Seri/Babax hexaploid wheat population. Funct Plant Biol 34:189-203
Onda Y, Mochida K (2016) Exploring genetic diversity in plants using high-throughput sequencing techniques. Curr Genomics 17:358-367
Oshima Y, Shikata M, Koyama T, Ohtsubo N, Mitsuda N, Ohme-Takagi M (2013) MIXTA-like transcription factors and WAX INDUCER1/SHINE1 coordinately regulate cuticle development in Arabidopsis and Torenia fournieri. Plant Cell 25:1609-1624
Pagani L, Kivisild T, Tarekegn A, Ekong R, Plaster C, Gallego Romero I, Ayub Q, Mehdi SQ, Thomas MG, Luiselli D, Bekele E, Bradman N, Balding DJ, Tyler-Smith C (2012) Ethiopian genetic diversity reveals linguistic stratification and complex influences on the Ethiopian gene pool. Am J Hum Genet 91:83-96
Pask A, Pietragalla J (2012) Leaf area, green crop area and senescence. In: Pask A, Pietragalla J, Mullan D, Reynolds M (eds) Physiological breeding II: a field guide to wheat phenotyping. CIMMYT, CIMMYT, DF, Mexico, pp 58-62
Pastor A, Palazzo A, Havlik P, Biemans H, Wada Y, Obersteiner M, Kabat P, Ludwig F (2019) The global nexus of food–trade–water sustaining environmental flows by 2050. Nat Sustain 2:499
Pavelek M, Tejklová E, Bjelková M (2014) Flax and Linseed. In: Von Mark VC, Dierig DA (eds) Industrial crops:- Breeding for bioenergy and bioproducts. Springer, New York pp 233-262
Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537-2539
Perrino P (1988) The diversity in Vavilov's Mediterranean Gene Center. Genet Resour Crop Evol 36:85-105
Perry SE, Nichols KW, Fernandez DE (1996) The MADS domain protein AGL15 localizes to the nucleus during early stages of seed development. Plant Cell 8:1977-1989
Phillipson DW (1993) The antiquity of cultivation and herding in Ethiopia. In: Shaw T, Sinclair P, Andah B, Okpoko A (eds) The Archaeology of Africa: Food, Metals and Towns. Routledge, London, pp 344-357
Pietragalla J (2012) Canopy temperature. In: Pask A, Pietragalla J, Mullan D, Reynolds M (eds) Physiological breeding II: A field guide to wheat phenotyping. CIMMYT, CIMMYT, DF, Mexico, pp 10-14
Pirooznia M, Goes FS, Zandi PP (2015) Whole-genome CNV analysis: advances in computational approaches. Front Genet 6:138
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959
Prudhomme C, Giuntoli I, Robinson EL, Clark DB, Arnell NW, Dankers R, Fekete BM, Franssen W, Gerten D, Gosling SN (2014) Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. P Nat Acad Sci 111:3262-3267
Przybylski R (2005) Flax oil and high linolenic oils. Bailey's industrial oil and fat products In: Shahidi F (ed) Edible oil and fat products. John Wiley & Sons, pp 281-292

References

Qian H, Lu H, Ding H, Lavoie M, Li Y, Liu W, Fu Z (2015) Analyzing *Arabidopsis thaliana* root proteome provides insights into the molecular bases of enantioselective imazethapyr toxicity. Sci Rep 5:11975

Quarrie S, Jones H (1979) Genotypic variation in leaf water potential, stomatal conductance and abscisic acid concentration in spring wheat subjected to artificial drought stress. Ann Bot 44:323-332

Rabetafika HN, Van Remoortel V, Danthine S, Paquot M, Blecker C (2011) Flaxseed proteins: food uses and health benefits. Int J Food Sci Tech 46:221-228

Rachinskaya OA, Lemesh VA, Muravenko OV, Yurkevich OYu, Guzenko EV, Bol'sheva NL, Bogdanova MV, Samatadze TE, Popov KV, Malyshev SV (2011) Genetic polymorphism of flax *Linum usitatissimum* based on the use of molecular cytogenetic markers. Russ J Genet 47:56-65

Rajwade AV, Arora RS, Kadoo NY, Harsulkar AM, Ghorpade PB, Gupta VS (2010) Relatedness of Indian flax genotypes (*Linum usitatissimum* L.): an inter-simple sequence repeat (ISSR) primer assay. Mol Biotechnol 45:161-170

Rakhra G, Sharma AD, Singh J (2015) Drought-induced changes in the accumulation of boiling-soluble proteins (p40, GST, HSP90) in the grains of drought-tolerant and drought-sensitive cultivars of *Triticum aestivum*. Crop Pasture Sci 66:904-911

Raman A, Verulkar S, Mandal N, Variar M, Shukla V, Dwivedi J, Singh B, Singh O, Swain P, Mall A (2012) Drought yield index to select high yielding rice lines under different drought stress severities. Rice 5:31

Rascio N, Navari-Izzo F (2011) Heavy metal hyperaccumulating plants: how and why do they do it? And what makes them so interesting? Plant Sci 180:169-181

Reinert S, Kortz A, Léon J, Naz AA (2016) Genome-wide association mapping in the global diversity set reveals new QTL controlling root system and related shoot variation in barley. Front Plant Sci 7:1061

Ren W-L, Wen Y-J, Dunwell JM, Zhang Y-M (2018) pKWmEB: integration of Kruskal–Wallis test with empirical Bayes under polygenic background control for multi-locus genome-wide association study. Heredity 120:208

Reynolds M, MUJEEB-KAZI A, Sawkins M (2005) Prospects for utilising plant-adaptive mechanisms to improve wheat and other crops in drought and salinity-prone environments. Ann Appl Biol 146:239-259

Reynolds M, Pask A, Mullan D (2012) Physiological breeding I: interdisciplinary approaches to improve crop adaptation. CIMMYT, CIMMYT, DF, Mexico

Richards K, Diederichsen A (2003) Cultivated flax and the genus *Linum* L.: Taxonomy and germplasm conservation. In: Muir A, Westcott N (eds) Flax. Taylor & Francis Group London, UK, pp 34-66

Richards R (1996) Defining selection criteria to improve yield under drought. Plant Growth Regul 20:157-166

Rieder B, Neuhaus HE (2011) Identification of an Arabidopsis plasma membrane–located ATP transporter important for anther development. Plant Cell 23:1932-1944

Robinson SI, McLaughlin OB, Marteinsdottir B, O'Gorman EJ (2018) Soil temperature effects on the structure and diversity of plant and invertebrate communities in a natural warming experiment. J Anim Ecol 87:634-646

Rossini F, Casa R (2003) Influence of sowing and harvest time on fibre flax (*Linum usitatissimum* L.) in the Mediterranean environment. J Agron Crop Sci 189:191-196

Rowland G (1998) Growing flax, production, management & diagnostic guide. In: Commission FCoCaSFD (ed), Winnipeg, MB, Canada, p 56

Royo C, Ramdani A, Moragues M, Villegas D (2006) Durum wheat under Mediterranean conditions as affected by seed size. J Agron Crop Sci 192:257-266

Rusinko J, McPartlon M (2017) Species tree estimation using Neighbor Joining. J Theor Biol 414:5-7
Saad RB, Zouari N, Ramdhan WB, Azaza J, Meynard D, Guiderdoni E, Hassairi A (2010) Improved drought and salt stress tolerance in transgenic tobacco overexpressing a novel A20/AN1 zinc-finger "AlSAP" gene isolated from the halophyte grass *Aeluropus littoralis*. Plant Mol Biol 72:171
Saikumar S, Varma C, Saiharini A, Kalmeshwer G, Nagendra K, Lavanya K, Ayyappa D (2016) Grain yield responses to varied level of moisture stress at reproductive stage in an interspecific population derived from Swarna/O. glaberrima introgression line. NJAS-Wagen J Life Sci 78:111-122
Sakamoto H, Matsuda O, Iba K (2008) ITN1, a novel gene encoding an ankyrin-repeat protein that affects the ABA-mediated production of reactive oxygen species and is involved in salt-stress tolerance in *Arabidopsis thaliana*. Plant J 56:411-422
Salekdeh GH, Reynolds M, Bennett J, Boyer J (2009) Conceptual framework for drought phenotyping during molecular breeding. Trends Plant Sci 14:488-496
Samberg LH, Shennan C, Zavaleta ES (2010) Human and environmental factors affect patterns of crop diversity in an Ethiopian highland agroecosystem. Prof Geogr 62:395-408
Sbabou L, Bucciarelli B, Miller S, Liu J, Berhada F, Filali-Maltouf A, Allan D, Vance C (2010) Molecular analysis of SCARECROW genes expressed in white lupin cluster roots. J Exp Bot 61:1351-1363
Schmalenbach I, Zhang L, Reymond M, Jiménez-Gómez JM (2014) The relationship between flowering time and growth responses to drought in the Arabidopsis Landsberg erecta x Antwerp-1 population. Front Plant Sci 5:609
Schneider KA, Rosales-Serna R, Ibarra-Perez F, Cazares-Enriquez B, Acosta-Gallegos JA, Ramirez-Vallejo P, Wassimi N, Kelly JD (1997) Improving common bean performance under drought stress. Crop Sci 37:43-50
Sertse D, You FM, Ravichandran S, Cloutier S (2019a) The complex genetic architecture of early root and shoot traits in flax revealed by genome-wide association analyses. Front Plant Sci 10:1483
Sertse D, You FM, Ravichandran S, Cloutier S (2019b) The genetic structure of flax illustrates environmental and anthropogenic selections that gave rise to its eco-geographical adaptation. Mol Phylogenetics Evol 137:22-32
Shavrukov Y, Kurishbayev A, Jatayev S, Shvidchenko V, Zotova L, Koekemoer F, de Groot S, Soole K, Langridge P (2017) Early flowering as a drought escape mechanism in plants: How can it aid wheat production? Front Plant Sci 8:1950
Sheng P, Tan J, Jin M, Wu F, Zhou K, Ma W, Heng Y, Wang J, Guo X, Zhang X, Cheng Z, Liu L, Wang C, Liu X, Wan J (2014) Albino midrib 1, encoding a putative potassium efflux antiporter, affects chloroplast development and drought tolerance in rice. Plant Cell Rep 33:1581-1594
Shikata H, Hanada K, Ushijima T, Nakashima M, Suzuki Y, Matsushita T (2014) Phytochrome controls alternative splicing to mediate light responses in Arabidopsis. Proc Natl Acad Sci USA 111:18781-18786
Shuai B, Reynaga-Pena CG, Springer PS (2002) The lateral organ boundaries gene defines a novel, plant-specific gene family. Plant Physiol 129:747-761
Singh B, Rao K, Sharma H (2011a) Comparison of selection indices to identify sorghum genotypes resistant to the spotted stemborer Chilo partellus (Lepidoptera: Noctuidae). Int J Trop Insect Sci 31:38-51
Singh KK, Mridula D, Rehal J, Barnwal P (2011b) Flaxseed: a potential source of food, feed and fiber. Crit Rev Food Sci Nutr 51:210-222
Slatkin M (2008) Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nat Rev Genet 9:477

References

Solanke AU, Sharma MK, Tyagi AK, Sharma AK (2009) Characterization and phylogenetic analysis of environmental stress-responsive SAP gene family encoding A20/AN1 zinc finger proteins in tomato. Mol Genet Genomics 282:153-164

Solh M, van Ginkel M (2014) Drought preparedness and drought mitigation in the developing world′ s drylands. Weather Climate Extremes 3:62-66

Sonah H, O'Donoughue L, Cober E, Rajcan I, Belzile F (2015) Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. Plant Biotechnol J 13:211-221

Soto-Cerda B, Cloutier S, Quian R, Gajardo H, Olivos M, You F (2018) Genome-wide association analysis of mucilage and hull content in flax (*Linum usitatissimum* L.) seeds. Int J Mol Sci 19:2870

Soto-Cerda BJ, Cloutier S (2012) Association mapping in plant genomes. In: Caliskan M (ed) Genetic Diversity in Plants. InTech, Rijecta, Croatia, pp 29-54

Soto-Cerda BJ, Cloutier S, Gajardo HA, Aravena G, Quian R (2019) Identifying drought-resilient flax genotypes and related-candidate genes based on stress indices, root traits and selective sweep. Euphytica 215:41

Soto-Cerda BJ, Diederichsen A, Ragupathy R, Cloutier S (2013) Genetic characterization of a core collection of flax (*Linum usitatissimum* L.) suitable for association mapping studies and evidence of divergent selection between fiber and linseed types. BMC Plant Biol 13:78

Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S (2014) Association mapping of seed quality traits using the Canadian flax (*Linum usitatissimum* L.) core collection. Theor Appl Genet 127:881-896

Stella JC, Rodríguez-González PM, Dufour S, Bendix J (2013) Riparian vegetation research in Mediterranean-climate regions: common patterns, ecological processes, and considerations for management. Hydrobiologia 719:291-315

Stirnberg P, van De Sande K, Leyser HO (2002) MAX1 and MAX2 control shoot lateral branching in Arabidopsis. Development 129:1131-1141

Strock CF, Burridge J, Massas AS, Beaver J, Beebe S, Camilo SA, Fourie D, Jochua C, Miguel M, Miklas PN (2019) Seedling root architecture and its relationship with seed yield across diverse environments in Phaseolus vulgaris. Field Crops Res 237:53-64

Sturm A, Šebková V, Lorenz K, Hardegger M, Lienhard S, Unger C (1995) Development-and organ-specific expression of the genes for sucrose synthase and three isoenzymes of acid β-fructofuranosidase in carrot. Planta 195:601-610

Suh MC, Samuels AL, Jetter R, Kunst L, Pollard M, Ohlrogge J, Beisson F (2005) Cuticular lipid composition, surface structure, and gene expression in Arabidopsis stem epidermis. Plant Physiol 139:1649-1665

Sukumaran S, Dreisigacker S, Lopes M, Chavez P, Reynolds MP (2015) Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. Theor Appl Genet 128:353-363

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P (2018) STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47:D607-D613

Tabata R, Sumida K, Yoshii T, Ohyama K, Shinohara H, Matsubayashi Y (2014) Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. Science 346:343-346

Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D (2019) Benefits and limitations of genome-wide association studies. Nat Rev Genet 20:467-484

References

Tamba CL, Ni Y-L, Zhang Y-M (2017) Iterative sure independence screening EM-Bayesian LASSO algorithm for multi-locus genome-wide association studies. PLoS Comp Biol 13:e1005357

Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Science 327:818-822

The 1001 Genomes Consortium (2016) 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. Cell 166:481-491

Thomas CL, Graham N, Hayden R, Meacham MC, Neugebauer K, Nightingale M, Dupuy LX, Hammond JP, White PJ, Broadley MR (2016) High-throughput phenotyping (HTP) identifies seedling root traits linked to variation in seed yield and nutrient capture in field-grown oilseed rape (Brassica napus L.). Ann Bot 118:655-665

Tian Ce, Muto H, Higuchi K, Matamura T, Tatematsu K, Koshiba T, Yamamoto KT (2004) Disruption and overexpression of auxin response factor 8 gene of Arabidopsis affect hypocotyl elongation and root growth habit, indicating its possible involvement in auxin homeostasis in light condition. Plant J 40:333-343

Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. Proc Natl Acad Sci USA 108:20260-20264

Tolkachev O, Zhuchenko A (2004) Biologically active substances of flax: medicinal and nutritional properties (a review). Pharm Chem J 34:360-367

Townshend J, Bolyend J (2008) Evaluation of two new linseed cultivars for oilseed production in Canterbury. In: McGill C, Rowarth J (eds) Seed symposium: Seeds for futures Agronomy Society of New Zealand and New Zealand Grassland Association Joint Symposium Agronomy Society of New Zealand, Massaey University, Palmerston North, New Zealand pp 105-112

Tricker PJ, ElHabti A, Schmidt J, Fleury D (2018) The physiological and genetic basis of combined drought and heat tolerance in wheat. J Exp Bot 69:3195-3210

Tsai YH (2011) PhyloGeoViz: a web-based program that visualizes genetic data on maps. Molecular Ecology Resources 11:557-561

Tuberosa R (2012) Phenotyping for drought tolerance of crops in the genomics era. Frontiers Physiol 3:1-26

Tyson H (1989) Genetic control of seed weight in flax (*Linum usitatissimum*) and possible implications. Theor Appl Genet 77:260-270

Uga Y, Sugimoto K, Ogawa S, Rane J, Ishitani M, Hara N, Kitomi Y, Inukai Y, Ono K, Kanno N (2013) Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. Nat Genet 45:1097

UN (2019) World Population Prospects 2019 United Nations Department of Economic and Social Affairs, Population Division, New York, p 91

Vaisey-Genser M, Morris D (2003) History of the cultivation and uses of flaxseed. In: Muir A, Westcott N (eds) Flax: the genus *Linum*. Taylor & Francis, London, pp 1-21

Van Zeist W, Bakker-Heeres J (1975) Evidence for linseed cultivation before 6000 BC. J Archaeol Sci 2:215-219

VanLiere JM, Rosenberg NA (2008) Mathematical properties of the r2 measure of linkage disequilibrium. Theor Popul Biol 74:130-137

Vavilov NI (1951) Phytogeorgraphic basis of plant breeding. In: Verdoorn J (ed) The origin, variation, immunity and breeding of cultivated plants: Selected writings of NI Vavilov The Chronica Botanica Co., New York p387

Von Reichenbach H, Schönbeck F (1995) Influence of VA-mycorrhiza on drought tolerance of flax (*Linum usitatissimum* L.). II: Effect of VA-mycorrhiza on stomatal gas exchange, shoot water potential, phosphorus nutrition and the accumulation of stress metabolites. Angew Bot 69:183-188

References

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77-78

Vos PG, Paulo MJ, Voorrips RE, Visser RG, van Eck HJ, van Eeuwijk FA (2017) Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. Theor Appl Genet 130:123-135

Wang H, Qin F (2017) Genome-wide association study reveals natural variations contributing to drought resistance in crops. Front Plant Sci 8:1110

Wang JL, Tang MQ, Chen S, Zheng XF, Mo HX, Li SJ, Wang Z, Zhu KM, Ding LN, Liu SY (2017) Down-regulation of BnDA1, whose gene locus is associated with the seeds weight, improves the seeds weight and organ size in Brassica napus. Plant Biotechnol J 15:1024-1033

Wang S-B, Feng J-Y, Ren W-L, Huang B, Zhou L, Wen Y-J, Zhang J, Dunwell JM, Xu S, Zhang Y-M (2016a) Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. Sci Rep 6:19444

Wang S, Bai Y, Shen C, Wu Y, Zhang S, Jiang D, Guilfoyle TJ, Chen M, Qi Y (2010) Auxin-related gene families in abiotic stress response in *Sorghum bicolor*. Funct Integr Genomics 10:533-546

Wang X, Wang H, Liu S, Ferjani A, Li J, Yan J, Yang X, Qin F (2016b) Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings. Nat Genet 48:1233

Wang Y-F, Jankauskiene Z, Qiu C-S, Gruzdeviene E, Long S-H, Alexopoulou E, Guo Y, Szopa J, Hao D-M, Fernando AL, Wang H (2018) Fiber flax breeding in China and Europe. J Nat Fibers 15:309-324

Wang Y, Li K, Chen L, Zou Y, Liu H, Tian Y, Li D, Wang R, Zhao F, Ferguson BJ (2015) MicroRNA167-directed regulation of the auxin response factors GmARF8a and GmARF8b is required for soybean nodulation and lateral root development. Plant Physiol 168:984-999

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R, Lambert G, Galbraith DW, Grassa CJ, Geraldes A, Cronk QC, Cullis C, Dash PK, Kumar PA, Cloutier S, Sharpe AG, Wong GK, Wang J, Deyholos MK (2012) The genome of flax (*Linum usitatissimum*) assembled de novo from short shotgun sequence reads. Plant J 72:461-473

Wasson A, Rebetzke G, Kirkegaard J, Christopher J, Richards R, Watt M (2014) Soil coring at multiple field environments can directly quantify variation in deep root traits to select wheat genotypes for breeding. J Exp Bot 65:6231-6249

Wei T, Simko V (2017) R package “corrplot”: Visualization of a correlation matrix. URL https://github com/taiyun/corrplot,(Version 084)

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. Evolution 38:1358-1370

Weiss E, Zohary D (2011) The Neolithic Southwest Asian founder crops: their biology and archaeobotany. Curr Anthropol 52:S237-S254

Wen Y-J, Zhang H, Ni Y-L, Huang B, Zhang J, Feng J-Y, Wang S-B, Dunwell JM, Zhang Y-M, Wu R (2017) Methodological implementation of mixed linear models in multi-locus genome-wide association studies. Brief Bioinform 19:700-712

Wen Y-J, Zhang H, Zhang J, Feng J-Y, Huang B, Dunwell JM, Zhang Y-M, Wu R (2016) A fast multi-locus random-SNP-effect EMMA for genome-wide association studies. bioRxiv:077404

Wilhite DA, Glantz MH (1985) Understanding: the drought phenomenon: the role of definitions. Water Int 10:111-120

Wittkop B, Snowdon R, Friedt W (2009) Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. Euphytica 170:131

Woo Y-M, Park H-J, Su’udi M, Yang J-I, Park J-J, Back K, Park Y-M, An G (2007) Constitutively wilted 1, a member of the rice YUCCA gene family, is required for maintaining water homeostasis and an appropriate root to shoot ratio. Plant Mol Biol 65:125-136

References

World-Bank (2015) World development indicators: gross domestic product 2014. December 16, 2015 edn. World Bank, http://wdi.worldbank.org

Wu L, Wu J, Liu Y, Gong X, Xu J, Lin D, Dong Y (2016) The rice pentatricopeptide repeat gene TCD10 is needed for chloroplast development under cold stress. Rice 9:67

Xia T, Li N, Dumenil J, Li J, Kamenski A, Bevan MW, Gao F, Li Y (2013) The ubiquitin receptor DA1 interacts with the E3 ubiquitin ligase DA2 to regulate seed and organ size in Arabidopsis. Plant Cell 25:3347-3359

Xie D, Dai Z, Yang Z, Tang Q, Sun J, Yang X, Song X, Lu Y, Zhao D, Zhang L (2018a) Genomic variations and association study of agronomic traits in flax. BMC Genomics 19:512

Xie M, Chen H, Huang L, O'Neil RC, Shokhirev MN, Ecker JR (2018b) A B-ARR-mediated cytokinin transcriptional network directs hormone cross-regulation and shoot development. Nat Commun 9:1604

Xing H, Fu X, Yang C, Tang X, Guo L, Li C, Xu C, Luo K (2018) Genome-wide investigation of pentatricopeptide repeat gene family in poplar and their expression analysis in response to biotic and abiotic stresses. Sci Rep 8:2817

Xu C, Luo F, Hochholdinger F (2016) LOB domain proteins: beyond lateral organ boundaries. Trends Plant Sci 21:159-167

Xu K, Chen S, Li T, Ma X, Liang X, Ding X, Liu H, Luo L (2015) OsGRAS23, a rice GRAS transcription factor gene, is involved in drought stress response through regulating expression of stress-responsive genes. BMC Plant Biol 15:141

Yang J, Ordiz MI, Jaworski JG, Beachy RN (2011) Induced accumulation of cuticular waxes enhances drought tolerance in Arabidopsis by changes in development of stomata. Plant Physiol Biochem 49:1448-1455

Yang P, Smalle J, Lee S, Yan N, Emborg TJ, Vierstra RD (2007) Ubiquitin C-terminal hydrolases 1 and 2 affect shoot architecture in Arabidopsis. Plant J 51:441-457

You F, Xiao J, Li P, Yao Z, Jia G, He L, Kumar S, Soto-Cerda B, Duguid S, Booker H (2018a) Genome-wide association study and selection signatures detect genomic regions associated with seed yield and oil quality in flax. Int J Mol Sci 19:2303

You FM, Duguid SD, Lam I, Cloutier S, Rashid KY, Booker HM (2016) Pedigrees and genetic base of flax cultivars registered in Canada. Can J Plant Sci 96:837-852

You FM, Duguid SD, Thambugala D, Cloutier S (2013) Statistical analysis and field evaluation of the type 2 modified augmented design (MAD) in phenotyping of flax (*Linum usitatissimum*) germplasms in multiple environments. Aust J Crop Sci 7:1789-1800

You FM, Huo N, Deal KR, Gu YQ, Luo M-C, McGuire PE, Dvorak J, Anderson OD (2011) Annotation-based genome-wide SNP discovery in the large and complex *Aegilops tauschii* genome using next-generation sequencing without a reference genome sequence. BMC Genomics 12:59

You FM, Jia G, Xiao J, Duguid SD, Rashid KY, Booker HM, Cloutier S (2017) Genetic variability of 27 traits in a core collection of flax (*Linum usitatissimum* L.). Front Plant Sci 8:1636

You FM, Xiao J, Li P, Yao Z, Jia G, He L, Zhu T, Luo MC, Wang X, Deyholos MK, Cloutier S (2018b) Chromosome-scale pseudomolecules refined by optical, physical and genetic maps in flax. Plant J 95:371-384

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203-208

Yue B, Xue W, Xiong L, Yu X, Luo L, Cui K, Jin D, Xing Y, Zhang Q (2006) Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. Genetics 172:1213-1228

Zajac T, Oleksy A, Klimek-Kopyra A, Kulig B (2012) Biological determinants of plant and crop productivity of flax (*Linum usitatissimum* L.). Acta Agrobot 65:3-14

References

Zając T, Oleksy A, Stokłosa A, Klimek-Kopyra A, Kulig B (2013) The development competition and productivity of linseed and pea-cultivars grown in a pure sowing or in a mixture. Eur J Agron 44:22-31

Zapata L, Peña-Chocarro L, Pérez-Jordá G, Stika H-P (2004) Early neolithic agriculture in the Iberian Peninsula. J World Prehist 18:283-325

Zeng Y-W, Du J, Pu X-Y, Yang J-Z, Yang T, Yang S-M, Yang X-M (2015) Coevolution between human's anticancer activities and functional foods from crop origin center in the world. Asian Pac J Cancer Prev 16:2119-2128

Zhang J, Feng J, Ni Y, Wen Y, Niu Y, Tamba C, Yue C, Song Q, Zhang Y (2017) pLARmEB: integration of least angle regression with empirical Bayes for multilocus genome-wide association studies. Heredity 118:517

Zhang J, Nguyen HT, Blum A (1999) Genetic analysis of osmotic adjustment in crop plants. J Exp Bot 50:291-302

Zhang T (2005) Influence of the seasonal snow cover on the ground thermal regime: An overview. Rev Geophys 43

Zhang Y-M, Tamba CL (2018) A fast mrMLM algorithm for multi-locus genome-wide association studies. bioRxiv:341784

Zhao FY, Hu F, Zhang SY, Wang K, Zhang CR, Liu T (2013) MAPKs regulate root growth by influencing auxin signaling and cell cycle-related gene expression in cadmium-stressed rice. Environ Sci Pollut Res 20:5449-5460

Zhao T, Dai A (2015) The magnitude and causes of global drought changes in the twenty-first century under a low–moderate emissions scenario. J Climate 28:4490-4512

Zhao Y, Cheng X, Liu X, Wu H, Bi H, Xu H (2018) The wheat MYB transcription factor TaMYB31 is involved in drought stress responses in Arabidopsis. Front Plant Sci 9

Zhao Y, Jiang C-H, Rehman RMA, Zhang H-L, Li J, Li Z-C (2019) Genetic analysis of roots and shoots in rice seedling by association mapping. Genes & Genomics 41:95-105

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

Zhu D, Maier A, Lee J-H, Laubinger S, Saijo Y, Wang H, Qu L-J, Hoecker U, Deng XW (2008b) Biochemical characterization of Arabidopsis complexes containing CONSTITUTIVELY PHOTOMORPHOGENIC1 and SUPPRESSOR OF PHYA proteins in light control of plant development. Plant Cell 20:2307-2323

Zhu J-K (2001) Cell signaling under salt, water and cold stresses. Curr Opin Plant Biol 4:401-406

Zhu J, Lynch JP (2004) The contribution of lateral rooting to phosphorus acquisition efficiency in maize (*Zea mays*) seedlings. Funct Plant Biol 31:949-958

Zohary D, Hopf M (2000) Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley. Oxford University Press, New York

Zohary D, Hopf M, Weiss E (2012) Domestication of plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. Oxford University Press

7. Appendices

I Supplementary materials for Chapter 2

I-1 Table I-1 Historical flax production reference locations.

Population	Country	Province	Reference for historical flax production	Locality	Latitude	Longitude
ABYS	Ethiopia	Amhara	Central statistical agency of Ethiopia, 2013. Area and production of major crops, Addis Ababa, Ethiopia	Lalibela	12.0309	39.0476
CANC	Canada	Saskatchewan	MacCracken J.A., 1916. A review of the status and possibilities of flax production and manipulation in Canada; Dominion of Canada, Department of Agriculture Hon. Martin Burrell, Minister of Agriculture, Ont. Ottawa	Saskatoon	52.133	-106.67
FIB1	Netherland	Arnhem	Eyre V.J., 1913. Projected revival of flax industry. In: Sir Ross, R.(ed) Science progress in twentieth century, Vol. VII Jhon Murray, London pp 596-629	The province	51.9692	5.6654
FIB2	Ukraine	Kyiv	Prylmachuk, T., Shtanko, T., Kovallov, V., 2017. Development of flax production in Ukraine	The province	50.4501	30.5234
MEDT	Portugal	Minho	Neto M.S., 2016. Conflicts and decline, 1620-1703. In: Freire, D., Lains, P.(Eds.), an agrarian history of Portugal, 1000-2000 : economic development on the European frontier, Brill, Boston pp 101-131	Braga	41.5511	-8.4283
NWAD	Argentina	Buenos Aires	Duval, L., 1916. The production and handling of grain in Argentina. Year book of the department of agriculture pp 281-298.	North Buenos Aires	-33.6037	-58.3816
OWAD	Turkey	Istanbul	Ertug, F., 2000. Linseed oil and oil mills in central Turkey Flax/*Linum* and Eruca, important oil plants of Anatolia	Istanbul	41.0082	28.9784
SAIN	India	Uttar-Pradesh	Gill, K.S., 1987. Linseed. Indian council of agricultural research. New Delhi, India. 386 pp.	Allahabad	25.4458	81.8463

Population	Country	Province	Reference for historical flax production	Locality	Latitude	Longitude
SAPK	Pakistan	Punjab	Amjad, M., 2014. Oilseeds crops of Pakistan. Pakistan agricultural research council, plant sciences division. Islamabad, Pakistan.	Faisalabad	31.4167	73.0911
TEM1	Russia	Vologda	Eyre V.J., 1913. The conditions of Russian agriculture. In: Sir Ross, R.(ed) Science progress in twentieth century, Vol. VII Jhon Murray, London, pp 175-193	Vologda	59.2667	39.9
TEM2	Georgia	Javakheti	Akhalkatsi, M., Ekhvaia, J., Asanidze, Z., 2012. Diversity and Genetic Erosion of Ancient Crops and Wild Relatives of Agricultural Cultivars for Food: Implications for Nature conservation in Georgia (Caucasus). In: Tiefenbacher, J., (ed). Perspectives on Nature Conservation – Patterns, Pressures and Prospects. InTech, Rijeka, Croatia, pp 51-92	The province	41.5479	43.7381
WEUR	France	Normandy	Eyre V.J., 1913. Projected revival of flax industry. In: Sir Ross, R.(ed) Science progress in twentieth century, Vol. VII Jhon Murray, London, pp 596-629	The province	48.8799	0.1713

The province indicate the coordinates were marked based on the region mentioned while the rest with specific locality were based on the locality.

I-2 Table I-2 Admixture ancestral estimate of genotypes and their putative populations at k=11

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_CAN_B_CN101472	ABYS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
O_CAN_B_CN101482	ABYS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
O_ETH_C_CN96988	ABYS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
O_ETH_C_CN96991	ABYS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
O_ETH_C_CN97004	ABYS	0.000	0.000	0.000	0.000	0.059	0.000	0.000	0.940	0.000	0.000	0.000
O_RUS_C_CN96845	ABYS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
O_USA_C_CN97393	ABYS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
F_RUS_C_CN32542	CANC	0.000	0.536	0.000	0.000	0.464	0.000	0.000	0.000	0.000	0.000	0.000
F_TUR_U_CN101382	CANC	0.000	0.392*	0.000	0.000	0.215	0.000	0.200	0.000	0.000	0.193	0.000
O_CAN_B_CN101560	CANC	0.259	0.468*	0.000	0.157	0.083	0.000	0.033	0.000	0.000	0.000	0.000
O_CAN_C_CDCSorrel	CANC	0.000	0.839	0.000	0.000	0.041	0.000	0.000	0.000	0.000	0.000	0.120
O_CAN_C_CN18973	CANC	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN18980	CANC	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN18981	CANC	0.000	0.654	0.000	0.000	0.280	0.000	0.000	0.000	0.000	0.000	0.067
O_CAN_C_CN19005	CANC	0.000	0.835	0.000	0.000	0.165	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN19017	CANC	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN33385	CANC	0.000	0.566	0.000	0.000	0.433	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN33388	CANC	0.000	0.655	0.000	0.000	0.345	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN33389	CANC	0.000	0.799	0.000	0.000	0.200	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN33397	CANC	0.000	0.523	0.000	0.000	0.477	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN37286	CANC	0.000	0.648	0.000	0.000	0.352	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN52732	CANC	0.000	0.580	0.000	0.000	0.420	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_E1747	CANC	0.000	0.601	0.000	0.000	0.399	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_FP2214	CANC	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_Linola989	CANC	0.000	0.489*	0.000	0.000	0.477	0.000	0.000	0.000	0.034	0.000	0.000
O_CAN_C_M5791	CANC	0.000	0.904	0.000	0.000	0.037	0.000	0.000	0.000	0.059	0.000	0.000
O_CAN_C_Macbeth	CANC	0.000	0.939	0.000	0.000	0.061	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_PrairieBlue	CANC	0.000	0.598	0.000	0.000	0.401	0.000	0.000	0.000	0.000	0.000	0.000

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_CAN_C_PrairieThunder	CANC	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_Shape	CANC	0.000	0.781	0.000	0.000	0.219	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_UGG5_5	CANC	0.000	0.539	0.000	0.000	0.349	0.000	0.000	0.000	0.000	0.112	0.000
O_USA_B_CN100785	CANC	0.023	0.716	0.027	0.000	0.106	0.024	0.000	0.000	0.000	0.103	0.000
O_USA_C_CN18994	CANC	0.000	0.598	0.000	0.000	0.402	0.000	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN33399	CANC	0.000	0.685	0.000	0.000	0.066	0.000	0.000	0.000	0.000	0.249	0.000
O_USA_C_CN33400	CANC	0.203	0.356	0.000	0.000	0.441	0.000	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN33992	CANC	0.000	0.663	0.000	0.000	0.337	0.000	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN97873	CANC	0.000	0.848	0.000	0.000	0.152	0.000	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN97881	CANC	0.000	0.634	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.361	0.000
O_USA_C_CN97921	CANC	0.000	0.600	0.000	0.000	0.078	0.000	0.100	0.025	0.137	0.060	0.000
F_BLR_C_CN101038	FIB1	0.000	0.000	0.000	0.487*	0.356	0.000	0.000	0.000	0.000	0.157	0.000
F_CHN_B_CN101053	FIB1	0.000	0.000	0.000	0.926	0.074	0.000	0.000	0.000	0.000	0.000	0.000
F_CHN_B_CN101230	FIB1	0.000	0.000	0.000	0.531	0.469	0.000	0.000	0.000	0.000	0.000	0.000
F_EU_C_Viking	FIB1	0.000	0.000	0.000	0.962	0.038	0.000	0.000	0.000	0.000	0.000	0.000
F_FRA_C_CN18982	FIB1	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
F_FRA_C_CN18988	FIB1	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN18983	FIB1	0.000	0.000	0.000	0.515	0.485	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN18987	FIB1	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN33390	FIB1	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN40081	FIB1	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_B_CN101055	FIB1	0.034	0.000	0.000	0.495*	0.258	0.000	0.000	0.000	0.000	0.213	0.000
F_RUS_B_CN101114	FIB1	0.000	0.000	0.000	0.555	0.445	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_B_CN101127	FIB1	0.000	0.000	0.000	0.527	0.472	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_C_CN101096	FIB1	0.000	0.000	0.000	0.513	0.487	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_C_CN101099	FIB1	0.000	0.000	0.000	0.668	0.332	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_U_CN101395	FIB1	0.000	0.000	0.000	0.764	0.236	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_U_CN101396	FIB1	0.000	0.000	0.000	0.627	0.373	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_U_CN101401	FIB1	0.000	0.000	0.000	0.615	0.384	0.000	0.000	0.000	0.000	0.000	0.000

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
F_UNK_C_CN33393	FIB1	0.000	0.000	0.000	0.772	0.227	0.000	0.000	0.000	0.000	0.000	0.000
O_LTU_C_CN101237	FIB1	0.000	0.000	0.000	0.996	0.004	0.000	0.000	0.000	0.000	0.000	0.000
O_RUS_B_CN101241	FIB1	0.000	0.000	0.000	0.598	0.402	0.000	0.000	0.000	0.000	0.000	0.000
O_RUS_B_CN101296	FIB1	0.000	0.000	0.000	0.817	0.183	0.000	0.000	0.000	0.000	0.000	0.000
O_RUS_B_CN101298	FIB1	0.000	0.000	0.000	0.607	0.393	0.000	0.000	0.000	0.000	0.000	0.000
O_RUS_B_CN101301	FIB1	0.000	0.000	0.000	0.523	0.243	0.000	0.000	0.000	0.000	0.233	0.000
F_AFG_U_CN100952	FIB2	0.000	0.000	0.000	0.000	0.246	0.000	0.000	0.004	0.000	0.634	0.116
F_CAN_B_CN101559	FIB2	0.296	0.000	0.000	0.000	0.186	0.000	0.000	0.000	0.000	0.517	0.000
F_EGY_C_CN98826	FIB2	0.000	0.000	0.000	0.000	0.306	0.000	0.000	0.000	0.000	0.693	0.000
F_HUN_C_CN98286	FIB2	0.000	0.000	0.000	0.075	0.353	0.000	0.000	0.000	0.000	0.571	0.000
F_NLD_C_CN97610	FIB2	0.000	0.000	0.000	0.023	0.238	0.000	0.000	0.000	0.000	0.739	0.000
F_NLD_U_CN101407	FIB2	0.000	0.000	0.000	0.000	0.067	0.000	0.000	0.000	0.000	0.933	0.000
F_RUS_L_CN97533	FIB2	0.000	0.000	0.000	0.000	0.457	0.000	0.000	0.000	0.000	0.543	0.000
F_RUS_U_CN101348	FIB2	0.000	0.000	0.000	0.000	0.065	0.000	0.000	0.000	0.000	0.934	0.000
F_UKR_U_CN101378	FIB2	0.000	0.000	0.000	0.000	0.062	0.000	0.000	0.000	0.000	0.937	0.000
F_UKR_U_CN101379	FIB2	0.000	0.000	0.000	0.000	0.312	0.000	0.000	0.000	0.000	0.688	0.000
F_USA_B_CN97665	FIB2	0.000	0.000	0.000	0.000	0.263	0.000	0.000	0.000	0.000	0.737	0.000
F_USA_B_CN98903	FIB2	0.000	0.000	0.000	0.096	0.088	0.000	0.000	0.000	0.000	0.816	0.000
O_ARG_C_CN98014	FIB2	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.030	0.021	0.947	0.000
O_CAN_C_CN19157	FIB2	0.049	0.000	0.000	0.000	0.248	0.000	0.000	0.000	0.000	0.703	0.000
O_CAN_C_CN33386	FIB2	0.000	0.260	0.000	0.000	0.225	0.000	0.026	0.000	0.000	0.488	0.000
O_FRA_C_CN97350	FIB2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.308	0.000	0.691	0.000
O_JPN_C_CN97470	FIB2	0.000	0.000	0.000	0.000	0.286	0.000	0.000	0.000	0.000	0.714	0.000
O_NLD_C_CN97613	FIB2	0.010	0.000	0.000	0.000	0.314	0.000	0.018	0.025	0.021	0.613	0.000
O_RUS_C_CN97487	FIB2	0.000	0.000	0.000	0.000	0.409	0.000	0.000	0.000	0.000	0.591	0.000
O_RUS_C_CN97604	FIB2	0.000	0.000	0.000	0.000	0.104	0.000	0.000	0.000	0.000	0.896	0.000
O_UNK_C_CN100547	FIB2	0.000	0.059	0.000	0.000	0.171	0.000	0.000	0.000	0.000	0.770	0.000
O_USA_B_CN101286	FIB2	0.000	0.199	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.801	0.000
O_USA_B_CN97404	FIB2	0.000	0.000	0.000	0.000	0.355	0.000	0.000	0.000	0.000	0.645	0.000

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_USA_B_CN97404B	FIB2	0.000	0.000	0.000	0.000	0.455	0.000	0.000	0.000	0.000	0.545	0.000
O_USA_C_CN97377	FIB2	0.000	0.061	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.884	0.055
O_USA_C_CN97444	FIB2	0.000	0.000	0.000	0.000	0.383	0.000	0.000	0.000	0.000	0.617	0.000
O_USA_C_CN97463	FIB2	0.000	0.000	0.000	0.000	0.264	0.000	0.000	0.000	0.000	0.735	0.000
O_USA_C_CN97642	FIB2	0.000	0.389	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.611	0.000
O_USA_C_CN97649	FIB2	0.000	0.358	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.641	0.000
U_ESP_U_CN101327	FIB2	0.000	0.000	0.000	0.000	0.182	0.000	0.000	0.000	0.000	0.534	0.284
U_ETH_L_CN100895B	FIB2	0.000	0.169	0.000	0.000	0.071	0.000	0.033	0.000	0.000	0.522	0.204
U_FRA_L_CN98742	FIB2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
U_RUS_L_CN97483	FIB2	0.000	0.000	0.000	0.000	0.294	0.000	0.000	0.000	0.000	0.706	0.000
U_USA_B_CN97402	FIB2	0.000	0.000	0.000	0.000	0.321	0.000	0.000	0.000	0.000	0.679	0.000
U_USA_B_CN97406	FIB2	0.000	0.000	0.000	0.000	0.327	0.000	0.094	0.000	0.000	0.579	0.000
U_USA_C_CN97397	FIB2	0.000	0.000	0.000	0.000	0.472	0.000	0.000	0.000	0.000	0.528	0.000
U_USA_C_CN97453	FIB2	0.000	0.000	0.000	0.000	0.082	0.000	0.000	0.000	0.000	0.918	0.000
O_TUR_U_CN100837	MEDT	0.009	0.000	0.104	0.000	0.000	0.036	0.520**	0.036	0.000	0.000	0.296
U_GRC_U_CN100885	MEDT	0.007	0.000	0.055	0.000	0.000	0.085	0.514**	0.054	0.000	0.000	0.285
U_PRT_C_CN100852	MEDT	0.000	0.000	0.064	0.000	0.000	0.077	0.529**	0.050	0.000	0.000	0.280
F_FRA_C_CN97351	NWAD	0.000	0.000	0.024	0.192	0.217	0.000	0.082	0.000	0.226	0.149	0.110
O_ARG_C_CN97334	NWAD	0.104	0.206	0.066	0.000	0.361	0.000	0.211	0.007	0.032	0.000	0.012
O_ARG_C_CN97953	NWAD	0.000	0.000	0.000	0.000	0.475	0.000	0.525	0.000	0.000	0.000	0.000
O_ARG_C_CN97967	NWAD	0.028	0.004	0.000	0.055	0.311	0.000	0.153	0.013	0.345	0.000	0.092
O_ARG_C_CN98012	NWAD	0.000	0.000	0.000	0.000	0.428	0.000	0.572	0.000	0.000	0.000	0.000
O_ARG_C_CN98027	NWAD	0.000	0.000	0.000	0.000	0.462	0.000	0.538	0.000	0.000	0.000	0.000
O_ARG_C_CN98039	NWAD	0.000	0.000	0.000	0.000	0.463	0.000	0.536	0.000	0.000	0.000	0.000
O_ARG_C_CN98634	NWAD	0.000	0.000	0.000	0.000	0.314	0.000	0.686	0.000	0.000	0.000	0.000
O_CAN_B_CN101463	NWAD	0.000	0.383	0.000	0.000	0.185	0.000	0.432	0.000	0.000	0.000	0.000
O_CAN_B_CN101554	NWAD	0.000	0.000	0.213	0.067	0.183	0.043	0.066	0.000	0.428	0.000	0.000
O_CAN_B_CN101580	NWAD	0.000	0.000	0.000	0.000	0.394	0.317	0.279	0.000	0.000	0.010	0.000
O_CAN_B_CN101598	NWAD	0.043	0.000	0.000	0.043	0.308	0.000	0.278	0.018	0.290	0.000	0.021

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_CAN_C_CN97633	NWAD	0.000	0.000	0.000	0.000	0.349	0.000	0.651	0.000	0.000	0.000	0.000
O_GEO_U_CN101366	NWAD	0.000	0.013	0.000	0.122	0.412	0.000	0.021	0.023	0.030	0.158	0.222
O_IND_L_CN96974	NWAD	0.000	0.000	0.000	0.000	0.467	0.000	0.533	0.000	0.000	0.000	0.000
O_NLD_C_CN97458	NWAD	0.000	0.046	0.000	0.000	0.398	0.000	0.253	0.000	0.000	0.303	0.000
O_PAK_C_CN97103	SAPK	0.000	0.000	0.000	0.000	0.190	0.447	0.000	0.104	0.000	0.149	0.110
O_RUS_B_CN101307	NWAD	0.015	0.091	0.000	0.020	0.350	0.000	0.007	0.010	0.081	0.243	0.182
O_USA_B_CN97679	NWAD	0.000	0.162	0.000	0.054	0.102	0.000	0.415	0.000	0.000	0.267	0.000
O_USA_B_CN97679B	NWAD	0.000	0.148	0.000	0.043	0.131	0.000	0.409	0.000	0.000	0.268	0.000
O_USA_C_CN97587	NWAD	0.000	0.000	0.000	0.000	0.444	0.000	0.242	0.000	0.086	0.104	0.124
O_USA_C_CN97907	NWAD	0.112	0.314	0.000	0.000	0.174	0.000	0.166	0.021	0.151	0.000	0.061
O_USA_C_CN98639	NWAD	0.023	0.000	0.000	0.007	0.324	0.000	0.076	0.029	0.436	0.105	0.000
F_CAN_C_CrepitamTabor	OWAD	0.000	0.000	0.000	0.000	0.424	0.000	0.000	0.000	0.000	0.000	0.576
F_HUN_C_CN98303	OWAD	0.000	0.000	0.000	0.000	0.446	0.000	0.000	0.000	0.000	0.000	0.554
F_IRN_C_CN97180	OWAD	0.173	0.000	0.000	0.142	0.472	0.000	0.000	0.012	0.021	0.035	0.145
F_RUS_B_CN101119	OWAD	0.036	0.000	0.000	0.422	0.457	0.000	0.000	0.000	0.000	0.085	0.000
F_RUS_C_CN101136	OWAD	0.000	0.071	0.000	0.339	0.015	0.000	0.000	0.000	0.073	0.477	0.025
F_TUR_U_CN101385	OWAD	0.000	0.000	0.000	0.000	0.284	0.000	0.000	0.000	0.000	0.435	0.281
F_USA_B_CN98923	OWAD	0.000	0.000	0.232	0.000	0.372	0.110	0.000	0.000	0.000	0.000	0.286
F_USA_B_CN98926	OWAD	0.154	0.000	0.000	0.023	0.364	0.000	0.000	0.000	0.000	0.460	0.000
F_USA_C_CN98954	OWAD	0.392	0.000	0.000	0.080	0.091	0.000	0.000	0.000	0.000	0.437	0.000
O_AFG_U_CN100807	OWAD	0.000	0.003	0.000	0.043	0.185	0.000	0.097	0.024	0.204	0.443	0.001
O_AFG_U_CN101338	OWAD	0.000	0.000	0.000	0.000	0.453	0.000	0.000	0.035	0.000	0.000	0.512
O_ARM_U_CN101373	OWAD	0.000	0.000	0.000	0.000	0.486	0.000	0.000	0.000	0.000	0.000	0.514
O_CAN_B_CN101448	OWAD	0.222	0.000	0.000	0.000	0.381	0.000	0.000	0.000	0.000	0.397	0.000
O_CAN_B_CN101565	OWAD	0.013	0.000	0.000	0.000	0.473	0.000	0.000	0.000	0.249	0.000	0.265
O_CAN_C_CN19159	OWAD	0.033	0.027	0.000	0.000	0.389	0.000	0.106	0.000	0.146	0.240	0.059
O_CZE_C_CN98683	OWAD	0.352	0.000	0.000	0.242	0.226	0.000	0.000	0.000	0.000	0.181	0.000
O_ETH_B_CN19007	OWAD	0.164	0.000	0.000	0.000	0.185	0.000	0.111	0.218	0.000	0.000	0.321
O_FRA_B_CN98741	OWAD	0.006	0.000	0.000	0.000	0.387	0.000	0.214	0.000	0.061	0.000	0.332

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_FRA_B_CN98806	OWAD	0.019	0.000	0.043	0.055	0.352	0.000	0.048	0.000	0.134	0.274	0.075
O_FRA_C_CN98734	OWAD	0.000	0.000	0.000	0.000	0.390	0.000	0.000	0.000	0.000	0.000	0.610
O_GBR_C_CN101265	OWAD	0.000	0.000	0.000	0.000	0.491	0.000	0.000	0.000	0.000	0.000	0.509
O_HUN_C_CN97238	OWAD	0.016	0.000	0.000	0.000	0.238	0.000	0.077	0.000	0.379	0.000	0.290
O_IND_C_CN101208	OWAD	0.000	0.000	0.208	0.000	0.372	0.031	0.000	0.000	0.175	0.000	0.214
O_IND_U_CN101310	OWAD	0.000	0.000	0.000	0.000	0.336	0.000	0.000	0.164	0.000	0.000	0.500
O_IRN_L_CN97129	OWAD	0.000	0.000	0.000	0.000	0.454	0.000	0.000	0.000	0.000	0.000	0.546
O_LTU_B_CN101240	OWAD	0.000	0.000	0.000	0.234	0.415	0.000	0.000	0.014	0.000	0.209	0.128
O_MAR_B_CN101026	OWAD	0.108	0.086	0.000	0.000	0.093	0.000	0.069	0.000	0.340	0.000	0.305
O_NLD_C_CN98056B	OWAD	0.013	0.000	0.000	0.351	0.447	0.000	0.000	0.000	0.000	0.189	0.000
O_RUS_B_CN101132	OWAD	0.000	0.057	0.000	0.000	0.271	0.000	0.078	0.000	0.000	0.287	0.307
O_RUS_B_CN101279	OWAD	0.116	0.000	0.000	0.000	0.312	0.000	0.105	0.000	0.104	0.061	0.303
O_RUS_C_CN98505	OWAD	0.000	0.000	0.000	0.000	0.336	0.000	0.000	0.007	0.000	0.000	0.658
O_SUN_C_CN100827	OWAD	0.000	0.000	0.000	0.000	0.405	0.000	0.000	0.000	0.063	0.000	0.532
O_TUR_C_CN96911	OWAD	0.000	0.000	0.000	0.000	0.419	0.000	0.000	0.005	0.000	0.000	0.576
O_TUR_C_CN97147	OWAD	0.054	0.000	0.031	0.027	0.267	0.001	0.097	0.037	0.116	0.000	0.370
O_TUR_C_CN97153	OWAD	0.000	0.000	0.000	0.000	0.475	0.000	0.000	0.000	0.000	0.000	0.525
O_TUR_C_CN98869	OWAD	0.000	0.000	0.000	0.000	0.313	0.000	0.000	0.000	0.000	0.000	0.687
O_TUR_U_CN101331	OWAD	0.000	0.000	0.000	0.021	0.480	0.000	0.000	0.000	0.000	0.221	0.278
O_UNK_C_CN30861	OWAD	0.000	0.147	0.000	0.002	0.216	0.000	0.199	0.000	0.000	0.076	0.359
O_UNK_U_CN100841	OWAD	0.000	0.000	0.000	0.000	0.444	0.000	0.000	0.000	0.000	0.000	0.556
O_USA_C_CN97403	OWAD	0.000	0.000	0.235	0.000	0.330	0.059	0.000	0.000	0.000	0.376	0.000
O_USA_C_CN97586	OWAD	0.083	0.000	0.000	0.000	0.386	0.000	0.067	0.000	0.094	0.370	0.000
O_USA_C_CN98231	OWAD	0.019	0.000	0.025	0.000	0.407	0.000	0.147	0.019	0.030	0.000	0.352
U_CHN_C_CN101016	OWAD	0.000	0.000	0.067	0.000	0.142	0.145	0.000	0.000	0.000	0.348	0.298
U_PAK_C_CN100790	OWAD	0.000	0.000	0.000	0.000	0.286	0.324	0.000	0.000	0.000	0.000	0.389
U_USA_C_CN97452	OWAD	0.000	0.000	0.000	0.000	0.401	0.000	0.000	0.000	0.000	0.362	0.237
O_CAN_B_CN101536	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101539	SAIN	0.000	0.000	0.567	0.000	0.252	0.095	0.000	0.000	0.000	0.000	0.086

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_CZE_C_CN97176	SAIN	0.000	0.000	0.909	0.000	0.000	0.091	0.000	0.000	0.000	0.000	0.000
O_HUN_C_CN98854	SAIN	0.000	0.000	0.801	0.000	0.000	0.199	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN97308	SAIN	0.000	0.000	0.458*	0.000	0.105	0.054	0.000	0.000	0.382	0.000	0.000
O_IND_C_CN97312	SAIN	0.070	0.000	0.483*	0.000	0.326	0.000	0.000	0.000	0.121	0.000	0.000
O_IND_C_CN98109	SAIN	0.000	0.000	0.945	0.000	0.000	0.055	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98135	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98157	SAIN	0.000	0.000	0.452*	0.000	0.000	0.072	0.000	0.000	0.476	0.000	0.000
O_IND_C_CN98254	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98364	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98370	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98415	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98440	SAIN	0.000	0.000	0.708	0.000	0.292	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98467	SAIN	0.000	0.000	0.699	0.000	0.167	0.000	0.000	0.000	0.134	0.000	0.000
O_IND_C_CN98468	SAIN	0.000	0.000	0.431*	0.000	0.235	0.333	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98969	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98973	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98982	SAIN	0.000	0.000	0.494*	0.028	0.348	0.000	0.000	0.000	0.130	0.000	0.000
O_IND_L_CN98240B	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_L_CN98242	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_PAK_C_CN98237	SAIN	0.000	0.000	0.548	0.000	0.000	0.452	0.000	0.000	0.000	0.000	0.000
O_PAK_L_CN97072	SAIN	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN97396	SAIN	0.000	0.000	0.506	0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.000
U_ROM_C_CN100678	SAIN	0.000	0.000	0.873	0.000	0.000	0.127	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101510	SAPK	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101511	SAPK	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101596	SAPK	0.000	0.000	0.000	0.000	0.423	0.576	0.000	0.000	0.000	0.000	0.000
O_HUN_C_CN98263B	SAPK	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
O_IRN_C_CN97050	SAPK	0.000	0.000	0.000	0.000	0.482	0.518	0.000	0.000	0.000	0.000	0.000
O_PAK_C_CN100629	SAPK	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_PAK_C_CN97064	SAPK	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
O_PAK_C_CN97092	SAPK	0.000	0.000	0.164	0.000	0.299	0.476	0.000	0.000	0.000	0.000	0.061
O_PAK_C_CN97096	SAPK	0.000	0.000	0.094	0.000	0.439	0.467	0.000	0.000	0.000	0.000	0.000
O_PAK_C_CN98239	SAPK	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
O_PAK_L_CN97083	SAPK	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN98541	SAPK	0.000	0.085	0.000	0.000	0.129	0.786	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN98613	SAPK	0.108	0.000	0.000	0.000	0.006	0.885	0.000	0.000	0.000	0.000	0.000
O_CAN_CDCBeth	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
F_CAN_B_CN101486	TEM1	0.082	0.000	0.005	0.000	0.506	0.000	0.042	0.008	0.000	0.216	0.140
F_CAN_C_G1186_94	TEM1	0.003	0.130	0.000	0.000	0.792	0.000	0.000	0.000	0.010	0.000	0.065
F_CHN_B_CN101052	TEM1	0.000	0.000	0.000	0.432	0.568	0.000	0.000	0.000	0.000	0.000	0.000
F_CHN_B_CN101417	TEM1	0.000	0.000	0.000	0.500	0.500	0.000	0.000	0.000	0.000	0.000	0.000
F_CHN_B_CN101419	TEM1	0.000	0.000	0.000	0.095	0.905	0.000	0.000	0.000	0.000	0.000	0.000
F_CHN_B_CN101421	TEM1	0.000	0.000	0.000	0.317	0.662	0.000	0.000	0.000	0.000	0.021	0.000
F_CZE_C_CN98704	TEM1	0.000	0.000	0.000	0.000	0.578	0.000	0.368	0.000	0.054	0.000	0.000
F_FRA_C_CN101392	TEM1	0.000	0.000	0.000	0.222	0.738	0.000	0.000	0.000	0.000	0.040	0.000
F_FRA_L_CN98710	TEM1	0.000	0.000	0.000	0.000	0.609	0.000	0.000	0.039	0.000	0.353	0.000
F_HUN_C_CN100864	TEM1	0.128	0.000	0.000	0.108	0.631	0.000	0.000	0.004	0.018	0.007	0.104
F_LTU_B_CN101118	TEM1	0.000	0.000	0.000	0.068	0.932	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN100795	TEM1	0.000	0.000	0.000	0.000	0.656	0.000	0.000	0.000	0.000	0.344	0.000
F_NLD_C_CN100929	TEM1	0.000	0.000	0.000	0.291	0.709	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN18998	TEM1	0.000	0.000	0.000	0.224	0.776	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN19001	TEM1	0.000	0.000	0.000	0.454	0.546	0.000	0.000	0.000	0.000	0.000	0.000
F_NLD_C_CN97424	TEM1	0.087	0.000	0.000	0.000	0.779	0.000	0.000	0.000	0.000	0.133	0.000
F_NLD_C_CN97616	TEM1	0.000	0.020	0.000	0.000	0.746	0.000	0.000	0.000	0.000	0.234	0.000
F_NLD_U_CN98150	TEM1	0.030	0.000	0.000	0.208	0.534	0.000	0.000	0.000	0.000	0.229	0.000
F_ROM_U_CN101403	TEM1	0.000	0.000	0.000	0.235	0.555	0.000	0.000	0.000	0.000	0.184	0.026
F_ROM_U_CN101405	TEM1	0.117	0.000	0.000	0.259	0.536	0.000	0.000	0.000	0.088	0.000	0.000
F_RUS_B_CN101039	TEM1	0.000	0.000	0.000	0.338	0.662	0.000	0.000	0.000	0.000	0.000	0.000

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
F_RUS_B_CN101115	TEM1	0.000	0.000	0.000	0.253	0.747	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_B_CN101116	TEM1	0.000	0.000	0.000	0.262	0.738	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_C_CN101154	TEM1	0.000	0.000	0.000	0.077	0.888	0.000	0.000	0.000	0.000	0.035	0.000
F_RUS_C_CN35791	TEM1	0.000	0.000	0.000	0.358	0.642	0.000	0.000	0.000	0.000	0.000	0.000
F_RUS_L_CN97503	TEM1	0.000	0.000	0.000	0.000	0.561	0.000	0.000	0.000	0.000	0.439	0.000
F_RUS_L_CN97530	TEM1	0.000	0.000	0.000	0.041	0.664	0.000	0.000	0.000	0.000	0.295	0.000
F_RUS_L_CN97531	TEM1	0.000	0.000	0.000	0.000	0.521	0.000	0.000	0.000	0.000	0.479	0.000
F_RUS_U_CN101364	TEM1	0.000	0.000	0.000	0.000	0.641	0.000	0.000	0.000	0.000	0.359	0.000
F_RUS_U_CN101394	TEM1	0.000	0.000	0.000	0.303	0.560	0.000	0.000	0.000	0.000	0.137	0.000
F_RUS_U_CN101402	TEM1	0.000	0.131	0.000	0.000	0.590	0.000	0.041	0.000	0.030	0.000	0.209
F_RUS_U_CN101406	TEM1	0.000	0.000	0.000	0.139	0.861	0.000	0.000	0.000	0.000	0.000	0.000
F_SWE_C_Atlas	TEM1	0.000	0.000	0.000	0.000	0.762	0.120	0.000	0.000	0.000	0.118	0.000
F_SWE_C_CN97871	TEM1	0.000	0.017	0.000	0.000	0.802	0.000	0.000	0.000	0.051	0.130	0.000
F_TUR_U_CN101386	TEM1	0.000	0.332	0.000	0.023	0.545	0.000	0.000	0.000	0.000	0.038	0.062
F_UKR_C_CN32546	TEM1	0.000	0.000	0.000	0.047	0.897	0.000	0.000	0.000	0.000	0.056	0.000
F_UKR_U_CN101397	TEM1	0.000	0.000	0.000	0.000	0.810	0.000	0.000	0.000	0.000	0.190	0.000
F_USA_C_CN98934	TEM1	0.000	0.000	0.000	0.000	0.655	0.000	0.010	0.009	0.048	0.277	0.000
O_AFG_C_CN98176	TEM1	0.000	0.000	0.000	0.000	0.601	0.000	0.000	0.000	0.000	0.000	0.399
O_ARG_B_CN98037	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
O_ARG_B_CN98037B	TEM1	0.000	0.000	0.000	0.000	0.708	0.000	0.184	0.000	0.059	0.000	0.048
O_ARG_C_CN97214	TEM1	0.000	0.000	0.000	0.000	0.774	0.000	0.226	0.000	0.000	0.000	0.000
O_ARG_C_CN97958	TEM1	0.022	0.000	0.000	0.000	0.620	0.000	0.170	0.000	0.150	0.000	0.037
O_ARG_C_CN97961	TEM1	0.016	0.000	0.000	0.000	0.607	0.000	0.263	0.000	0.114	0.000	0.000
O_ARG_C_CN98007	TEM1	0.066	0.000	0.000	0.000	0.618	0.000	0.153	0.018	0.115	0.013	0.017
O_ARG_C_CN98279	TEM1	0.027	0.000	0.000	0.000	0.880	0.000	0.000	0.000	0.000	0.000	0.093
O_AUS_C_Avantgarde	TEM1	0.000	0.000	0.000	0.000	0.642	0.000	0.000	0.000	0.062	0.000	0.296
O_AUS_C_CN98984	TEM1	0.000	0.000	0.000	0.000	0.849	0.000	0.000	0.000	0.151	0.000	0.000
O_CAN_B_CN101466	TEM1	0.000	0.000	0.314	0.000	0.652	0.033	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101469	TEM1	0.000	0.000	0.000	0.000	0.801	0.000	0.000	0.199	0.000	0.000	0.000

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_CAN_B_CN101471	TEM1	0.000	0.000	0.000	0.000	0.635	0.000	0.000	0.365	0.000	0.000	0.000
O_CAN_B_CN101493	TEM1	0.000	0.000	0.156	0.000	0.577	0.266	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101600	TEM1	0.000	0.000	0.000	0.000	0.658	0.000	0.000	0.000	0.000	0.000	0.342
O_CAN_C_CDCGold	TEM1	0.000	0.272	0.000	0.000	0.618	0.000	0.000	0.000	0.111	0.000	0.000
O_CAN_C_CDCMons	TEM1	0.000	0.193	0.000	0.000	0.807	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN101413	TEM1	0.000	0.335	0.000	0.000	0.495*	0.000	0.000	0.000	0.000	0.000	0.170
O_CAN_C_CN18979	TEM1	0.000	0.122	0.000	0.000	0.877	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN19003	TEM1	0.000	0.194	0.000	0.000	0.806	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_CN19004	TEM1	0.000	0.233	0.000	0.000	0.749	0.000	0.018	0.000	0.000	0.000	0.000
O_CAN_C_CN97392	TEM1	0.000	0.000	0.000	0.000	0.584	0.000	0.000	0.000	0.000	0.319	0.096
O_CAN_C_CN97571	TEM1	0.027	0.000	0.000	0.000	0.734	0.000	0.027	0.000	0.198	0.000	0.014
O_CAN_C_CN97671	TEM1	0.000	0.000	0.000	0.000	0.618	0.000	0.000	0.000	0.000	0.382	0.000
O_CAN_C_DoubleLow	TEM1	0.000	0.191	0.000	0.000	0.602	0.000	0.200	0.000	0.007	0.000	0.000
O_CAN_C_FP2270	TEM1	0.008	0.114	0.000	0.000	0.752	0.000	0.000	0.000	0.011	0.000	0.116
O_CAN_C_Hanley	TEM1	0.000	0.109	0.000	0.000	0.891	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_Lirina	TEM1	0.000	0.194	0.000	0.000	0.700	0.000	0.000	0.000	0.000	0.000	0.106
O_CAN_C_M96006	TEM1	0.000	0.397	0.000	0.000	0.603	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_PrairieGrande	TEM1	0.000	0.368	0.000	0.000	0.632	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_S95407	TEM1	0.000	0.322	0.000	0.000	0.539	0.000	0.019	0.000	0.000	0.080	0.041
O_CAN_C_SP2047	TEM1	0.000	0.279	0.000	0.000	0.721	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_C_UGG146_1	TEM1	0.000	0.037	0.000	0.000	0.893	0.000	0.070	0.000	0.000	0.000	0.000
O_CAN_C_YSED18	TEM1	0.000	0.168	0.000	0.000	0.619	0.000	0.000	0.000	0.213	0.000	0.000
O_CAN_c_Norman	TEM1	0.000	0.263	0.000	0.000	0.737	0.000	0.000	0.000	0.000	0.000	0.000
O_CSK_C_CN100884	TEM1	0.000	0.000	0.033	0.028	0.734	0.000	0.000	0.005	0.074	0.127	0.000
O_CYP_U_CN100838	TEM1	0.000	0.000	0.000	0.000	0.615	0.000	0.000	0.000	0.322	0.000	0.062
O_CZE_C_CN100805	TEM1	0.000	0.000	0.000	0.000	0.621	0.000	0.000	0.000	0.000	0.000	0.379
O_DEU_B_CN97430	TEM1	0.000	0.000	0.000	0.000	0.826	0.000	0.174	0.000	0.000	0.000	0.000
O_DEU_B_CN97430B	TEM1	0.000	0.000	0.000	0.000	0.800	0.000	0.200	0.000	0.000	0.000	0.000
O_DEU_C_CN100881	TEM1	0.000	0.000	0.000	0.000	0.515	0.000	0.000	0.000	0.485	0.000	0.000

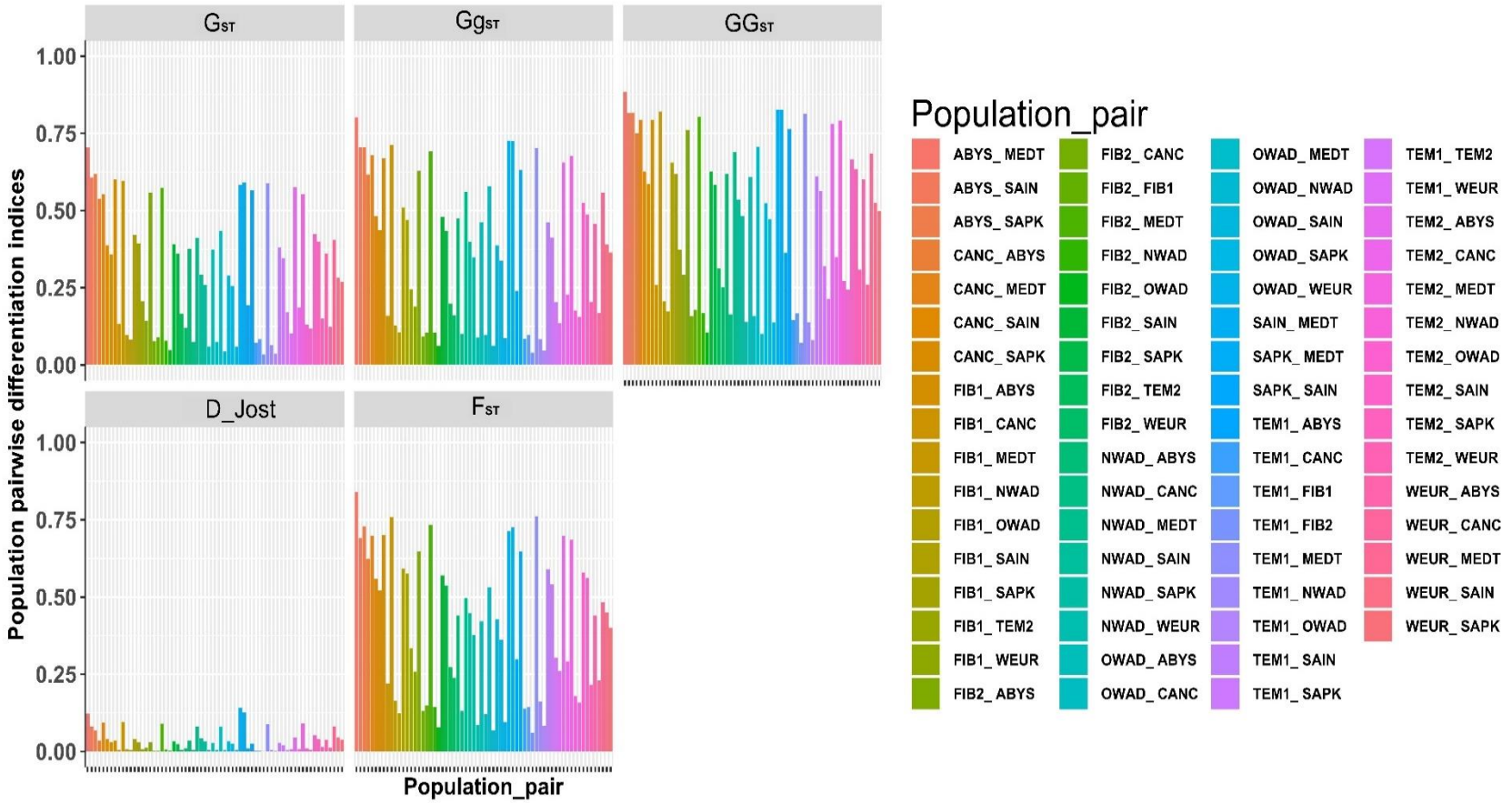
Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_DEU_C_CN97886	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
O_DEU_C_CN98475	TEM1	0.000	0.000	0.000	0.000	0.660	0.000	0.000	0.000	0.031	0.000	0.309
O_EGY_U_CN101329	TEM1	0.000	0.000	0.000	0.112	0.819	0.000	0.000	0.000	0.021	0.048	0.000
O_FRA_B_CN100863	TEM1	0.000	0.000	0.000	0.000	0.565	0.000	0.000	0.000	0.000	0.435	0.000
O_FRA_C_CN18989	TEM1	0.000	0.267	0.000	0.000	0.733	0.000	0.000	0.000	0.000	0.000	0.000
O_FRA_C_CN98712	TEM1	0.000	0.000	0.000	0.000	0.732	0.000	0.000	0.000	0.268	0.000	0.000
O_FRA_C_CN98752	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
O_FRA_C_CN98794	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
O_FRA_C_CN98807	TEM1	0.027	0.000	0.000	0.000	0.580	0.000	0.000	0.000	0.227	0.122	0.045
O_GRC_U_CN101325	TEM1	0.000	0.000	0.001	0.007	0.817	0.003	0.000	0.003	0.011	0.100	0.058
O_HUN_C_CN100883	TEM1	0.038	0.000	0.000	0.000	0.497*	0.000	0.000	0.000	0.465	0.000	0.000
O_HUN_C_CN97287	TEM1	0.000	0.000	0.000	0.000	0.858	0.000	0.000	0.000	0.000	0.000	0.142
O_HUN_C_CN97300	TEM1	0.001	0.147	0.000	0.000	0.705	0.000	0.067	0.003	0.040	0.037	0.000
O_HUN_C_CN98275	TEM1	0.000	0.000	0.000	0.000	0.562	0.000	0.000	0.000	0.438	0.000	0.000
O_HUN_C_CN98278	TEM1	0.000	0.000	0.000	0.000	0.966	0.000	0.000	0.000	0.000	0.000	0.034
O_IND_C_CN98250	TEM1	0.000	0.000	0.000	0.000	0.614	0.386	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98363	TEM1	0.000	0.000	0.000	0.000	0.853	0.147	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98397	TEM1	0.000	0.000	0.222	0.000	0.778	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98398	TEM1	0.000	0.000	0.120	0.000	0.880	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98961	TEM1	0.000	0.000	0.008	0.000	0.992	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_C_CN98974	TEM1	0.000	0.000	0.147	0.000	0.853	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_L_CN98240	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
O_IND_U_CN98569	TEM1	0.000	0.000	0.084	0.000	0.760	0.000	0.000	0.000	0.156	0.000	0.000
O_IRN_L_CN97129B	TEM1	0.000	0.000	0.000	0.000	0.717	0.000	0.000	0.000	0.000	0.000	0.283
O_NLD_C_CN18993	TEM1	0.000	0.000	0.000	0.000	0.699	0.000	0.300	0.000	0.000	0.000	0.000
O_NLD_C_CN98056	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
O_PAK_C_CN97056	TEM1	0.000	0.000	0.000	0.000	0.615	0.385	0.000	0.000	0.000	0.000	0.000
O_POL_B_CN98733	TEM1	0.000	0.000	0.000	0.000	0.597	0.000	0.000	0.000	0.000	0.000	0.403
O_PRT_L_CN100910	TEM1	0.000	0.000	0.004	0.000	0.635	0.033	0.161	0.007	0.014	0.057	0.089

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_ROM_C_CN100674	TEM1	0.039	0.000	0.000	0.000	0.673	0.000	0.000	0.000	0.000	0.000	0.288
O_ROM_C_CN97321	TEM1	0.000	0.000	0.000	0.000	0.708	0.000	0.026	0.000	0.028	0.000	0.237
O_RUS_B_CN101289	TEM1	0.000	0.000	0.000	0.354	0.646	0.000	0.000	0.000	0.000	0.000	0.000
O_RUS_B_CN101299	TEM1	0.000	0.000	0.000	0.281	0.609	0.000	0.000	0.000	0.000	0.110	0.000
O_RUS_C_CN100939	TEM1	0.385	0.000	0.000	0.000	0.615	0.000	0.000	0.000	0.000	0.000	0.000
O_RUS_C_CN97475	TEM1	0.000	0.000	0.000	0.000	0.528	0.000	0.000	0.000	0.000	0.471	0.000
O_RUS_C_CN97489	TEM1	0.000	0.000	0.000	0.000	0.648	0.000	0.000	0.000	0.000	0.351	0.000
O_RUS_C_CN97520	TEM1	0.010	0.000	0.000	0.000	0.691	0.000	0.000	0.000	0.000	0.299	0.000
O_RUS_C_CN97529	TEM1	0.000	0.000	0.000	0.000	0.737	0.000	0.000	0.000	0.000	0.263	0.000
O_RUS_L_CN97605	TEM1	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
O_RUS_U_CN101375	TEM1	0.000	0.000	0.000	0.000	0.614	0.000	0.010	0.023	0.000	0.278	0.075
O_TUR_L_CN96958	TEM1	0.000	0.000	0.177	0.000	0.499*	0.324	0.000	0.000	0.000	0.000	0.000
O_TUR_U_CN101332	TEM1	0.000	0.000	0.000	0.000	0.557	0.000	0.000	0.000	0.000	0.222	0.220
O_UKR_C_CN30860	TEM1	0.000	0.000	0.000	0.000	0.559	0.000	0.000	0.000	0.000	0.000	0.441
O_URY_C_CN98100	TEM1	0.000	0.000	0.000	0.000	0.600	0.000	0.400	0.000	0.000	0.000	0.000
O_USA_B_CN98566	TEM1	0.000	0.000	0.021	0.000	0.927	0.000	0.000	0.000	0.052	0.000	0.000
O_USA_C_CN19160	TEM1	0.000	0.000	0.000	0.000	0.695	0.000	0.305	0.000	0.000	0.000	0.000
O_USA_C_CN97366	TEM1	0.000	0.000	0.000	0.000	0.938	0.000	0.054	0.000	0.007	0.000	0.000
O_USA_C_CN97407	TEM1	0.000	0.037	0.000	0.000	0.600	0.000	0.339	0.000	0.000	0.023	0.000
O_USA_C_CN97639	TEM1	0.000	0.000	0.000	0.000	0.765	0.000	0.014	0.000	0.019	0.201	0.000
O_USA_C_CN97639B	TEM1	0.000	0.000	0.000	0.000	0.819	0.000	0.000	0.000	0.000	0.181	0.000
O_USA_C_CN97718	TEM1	0.000	0.032	0.000	0.000	0.832	0.000	0.093	0.043	0.000	0.000	0.000
O_USA_C_CN97728	TEM1	0.000	0.177	0.000	0.000	0.621	0.000	0.105	0.000	0.000	0.097	0.000
O_USA_C_CN97740	TEM1	0.000	0.223	0.000	0.000	0.745	0.000	0.000	0.000	0.000	0.032	0.000
O_USA_C_CN97749	TEM1	0.113	0.223	0.000	0.000	0.502	0.000	0.000	0.000	0.000	0.162	0.000
O_USA_C_CN97890	TEM1	0.000	0.216	0.000	0.000	0.579	0.000	0.083	0.001	0.062	0.059	0.000
O_USA_C_CN98535	TEM1	0.034	0.000	0.000	0.000	0.549	0.000	0.000	0.014	0.312	0.000	0.092
O_USA_C_CN98542	TEM1	0.000	0.034	0.000	0.000	0.841	0.000	0.000	0.000	0.124	0.000	0.000
O_USA_C_CN98821	TEM1	0.000	0.365	0.000	0.000	0.634	0.000	0.000	0.000	0.000	0.000	0.000

Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
U_ARG_C_CN97341	TEM1	0.000	0.001	0.000	0.031	0.870	0.000	0.063	0.000	0.000	0.035	0.000
U_CAN_B_CN101610	TEM1	0.000	0.000	0.000	0.000	0.606	0.208	0.186	0.000	0.000	0.000	0.000
U_CZE_C_CN98689	TEM1	0.016	0.000	0.000	0.000	0.852	0.000	0.077	0.000	0.055	0.000	0.000
U_IND_C_CN100799	TEM1	0.000	0.000	0.000	0.000	0.680	0.000	0.000	0.000	0.000	0.000	0.320
U_MAR_C_CN98193	TEM1	0.000	0.000	0.000	0.000	0.808	0.000	0.000	0.000	0.192	0.000	0.000
U_NZL_B_CN100797	TEM1	0.000	0.000	0.000	0.000	0.652	0.000	0.348	0.000	0.000	0.000	0.000
U_NZL_B_CN100797B	TEM1	0.000	0.000	0.000	0.000	0.582	0.000	0.418	0.000	0.000	0.000	0.000
U_PRT_C_CN97768	TEM1	0.017	0.006	0.006	0.083	0.562	0.000	0.071	0.012	0.088	0.128	0.028
U_TUR_U_CN100828	TEM1	0.000	0.000	0.000	0.000	0.517	0.000	0.000	0.000	0.000	0.000	0.483
U_USA_B_CN97670	TEM1	0.031	0.129	0.000	0.000	0.743	0.000	0.070	0.000	0.024	0.000	0.004
U_USA_B_CN98644	TEM1	0.000	0.000	0.011	0.000	0.884	0.000	0.000	0.000	0.105	0.000	0.000
U_USA_C_CN98610	TEM1	0.000	0.025	0.000	0.000	0.799	0.176	0.000	0.000	0.000	0.000	0.000
F_CAN_C_CN100848	TEM2	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
F_JPN_C_CN98072	TEM2	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
F_USA_C_CN98946	TEM2	0.931	0.000	0.000	0.000	0.069	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101454	TEM2	0.664	0.000	0.000	0.000	0.336	0.000	0.000	0.000	0.000	0.000	0.000
O_CAN_B_CN101461	TEM2	0.567	0.000	0.000	0.006	0.000	0.000	0.263	0.057	0.014	0.000	0.092
O_CAN_C_CN19158	TEM2	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O_GEO_U_CN101367	TEM2	0.924	0.000	0.000	0.000	0.076	0.000	0.000	0.000	0.000	0.000	0.000
O_USA_C_CN98812	TEM2	0.447*	0.215	0.000	0.000	0.337	0.000	0.000	0.000	0.000	0.000	0.000
U_USA_B_CN97584	TEM2	0.730	0.000	0.000	0.000	0.270	0.000	0.000	0.000	0.000	0.000	0.000
U_USA_B_CN97584B	TEM2	0.548	0.000	0.000	0.000	0.452	0.000	0.000	0.000	0.000	0.000	0.000
F_FRA_C_CN98708	WEUR	0.189	0.000	0.000	0.000	0.153	0.000	0.000	0.000	0.658	0.000	0.000
O_CAN_B_CN101496	WEUR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.046	0.953	0.000	0.000
O_CAN_B_CN101594	WEUR	0.083	0.000	0.000	0.000	0.154	0.000	0.139	0.007	0.619	0.000	0.000
O_CAN_B_CN101595	WEUR	0.000	0.000	0.000	0.000	0.209	0.000	0.000	0.000	0.791	0.000	0.000
O_CZE_C_CN100851	WEUR	0.000	0.000	0.000	0.054	0.014	0.000	0.000	0.000	0.603	0.302	0.027
O_FRA_C_CN100928	WEUR	0.000	0.000	0.000	0.000	0.138	0.000	0.073	0.045	0.654	0.000	0.089
O_FRA_C_CN98753	WEUR	0.000	0.000	0.000	0.000	0.334	0.000	0.000	0.000	0.666	0.000	0.000

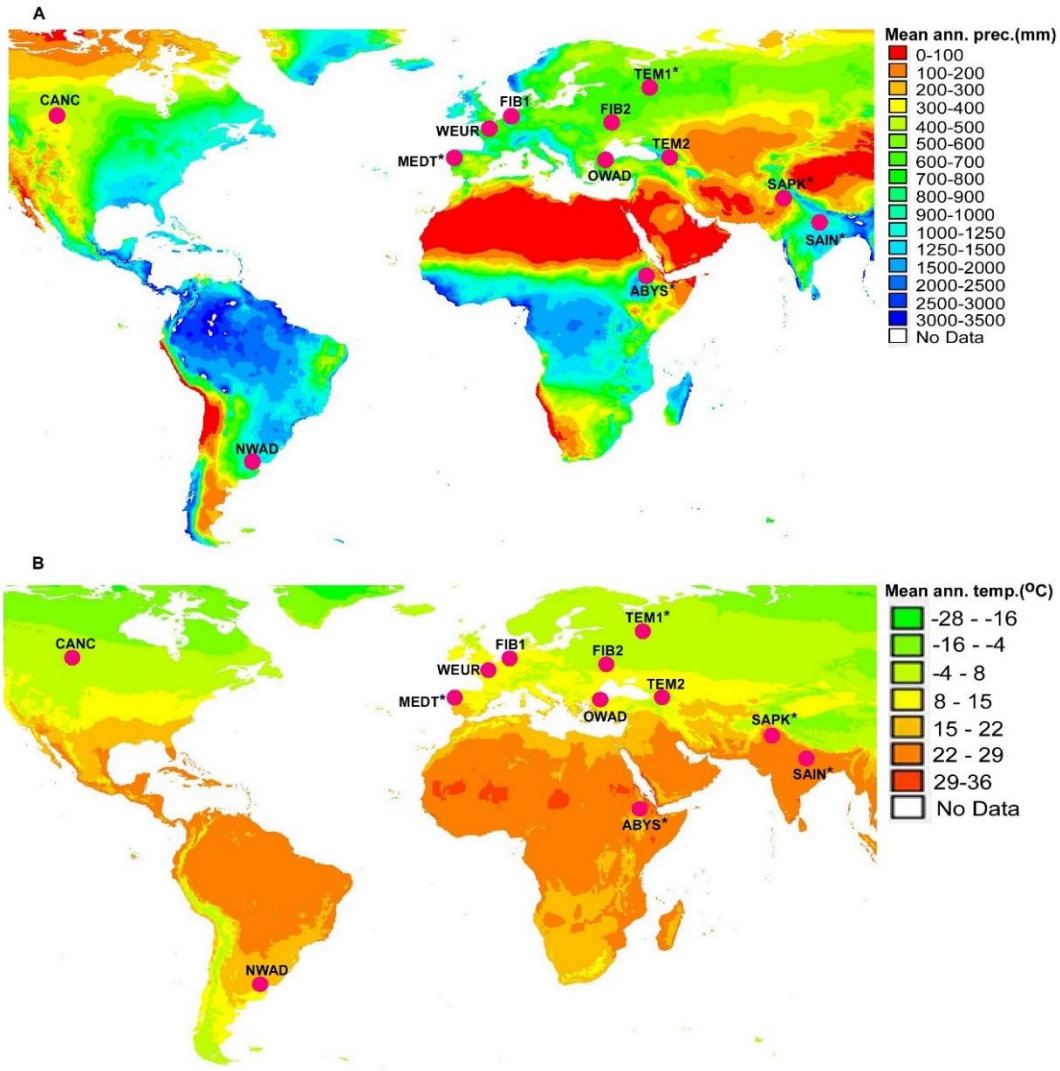
Genotype	Population	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
O_FRA_C_CN98767	WEUR	0.000	0.000	0.000	0.001	0.286	0.000	0.000	0.000	0.655	0.000	0.058
O_HUN_C_CN98276	WEUR	0.000	0.000	0.000	0.000	0.280	0.000	0.091	0.006	0.622	0.000	0.000
O_IND_C_CN97306	WEUR	0.000	0.000	0.246	0.000	0.078	0.000	0.000	0.000	0.676	0.000	0.000
O_IRL_C_CN98192	WEUR	0.000	0.000	0.000	0.000	0.283	0.000	0.000	0.000	0.717	0.000	0.000
O_USA_B_CN98566B	WEUR	0.000	0.000	0.298	0.000	0.000	0.000	0.000	0.000	0.702	0.000	0.000
O_USA_B_CN98566C	WEUR	0.000	0.000	0.301	0.000	0.000	0.000	0.000	0.000	0.699	0.000	0.000
U_CAN_B_CN101572	WEUR	0.173	0.000	0.000	0.000	0.259	0.000	0.000	0.000	0.516	0.000	0.051

Genotype names: The first letter indicate the type of the genotype F=fiber, O=oil and U=unknown (undifferentiated), the following three letters indicate internationally standard abbreviation (code) of country's name from where the genotype was originally collected; a single letter next indicate the breeding status C=cultivar, B=under breeding scheme, L=landrace, the last number preceded by CN is accession number at PGRC (Plant genetic resource Canada) but exceptions for with known cultivars where names are presented instead. The colors shaded on genotype and their population corresponds with the colors of their cluster (Q) of high ancestral value they belong to. Admixed populations (NWAD and OWAD) have more than one shaded Qs. Individuals with ancestral value higher than 50% can be assigned as one of the admixed population if they do not correspond in cluster that has at least one individuals having Q>90%.* Individuals clustered in a particular population while still have ancestral values less than 50% of that cluster are exceptionally assigned since they shared high accentual value with members in the cluster they are assigned at Ks 9, 10, 12 and 13. **MEDT in this grouping is ~50% in Q5 but over 90% in k=12 and above.



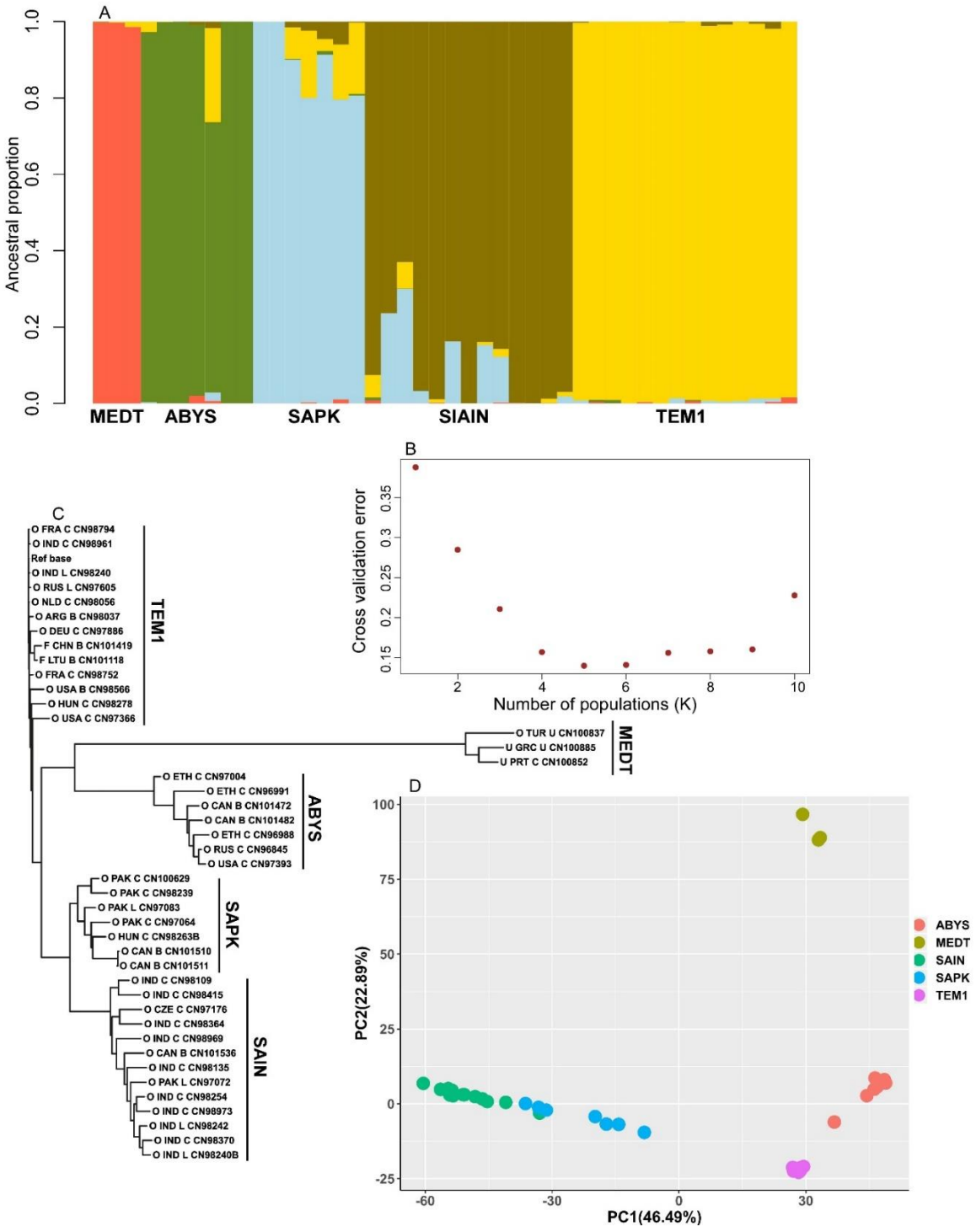
I-3 Fig. I-1 Pattern of population pairwise differentiation based on different algorisms:

GST (Nei & Chesser, 1983); GgST (Hedrick, 2005); GGST (Meirmans and Hedrick 2011); D_Jost (Jost, 2008); FST (Weir and Cockerham, 1984).



I-4 Fig. I-2 Coordinates of the populations overlaid to world climate data maps.

(**A**) mean annual total precipitation (mm); (**B**) mean annual daily temperature (°C); * representative populations of major groups. Note that the SA populations had differences in precipitation.

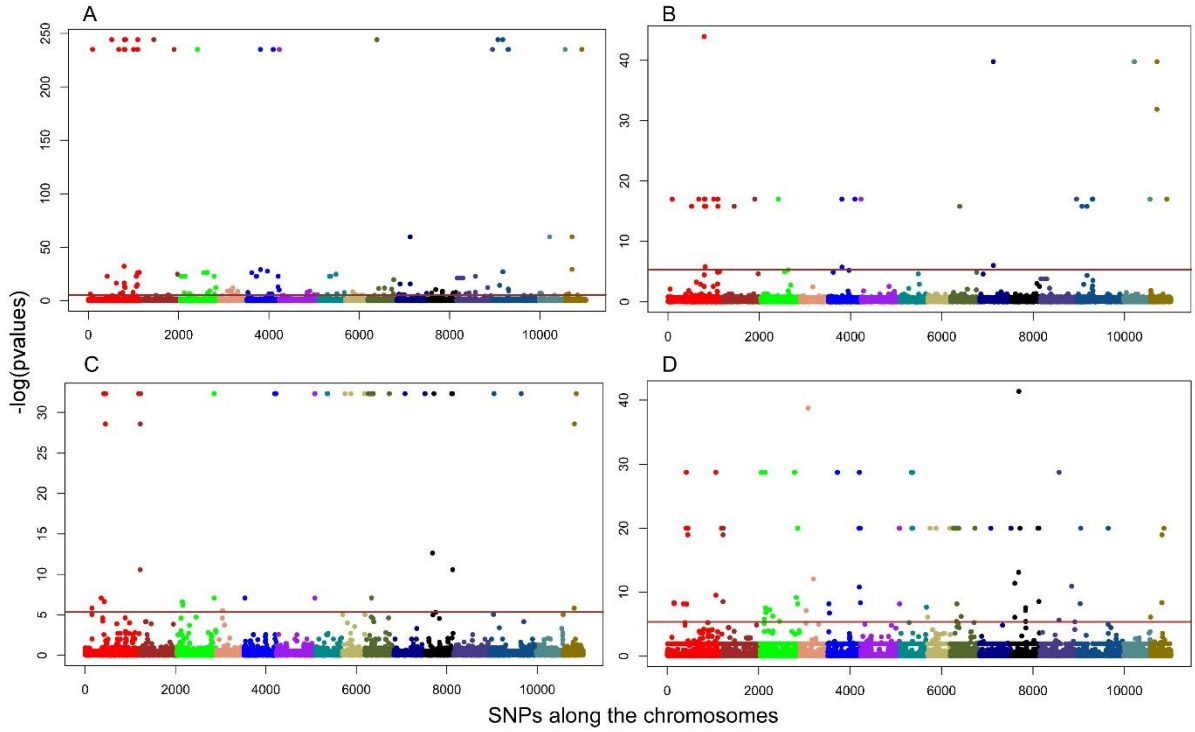


I-5 Fig. I-3 Population structure of genotypes used for genome-environment association.

(**A**) structure plot; (**B**) plot indicating appropriate number of populations; (**C**) phylogenetic trees of the selected genotypes; d) plot showing the first two principal components (PCs) of the genotypes.

I-6 Table I-3 Total number of loci associated to targeted environmental factors by chromosome

	Number of loci associated with environmental factors					
Chromosome	Crop_Mon_Temp	Crop_Mon_Prec	Ann_Temp	Ann_Prec	Day-length	Latitude
1	28	9	15	13	1	42
2	3	4	2	4	0	24
3	13	5	1	14	0	16
4	6	1	1	3	0	5
5	13	3	3	9	0	10
6	6	4	1	4	0	14
7	8	4	0	8	0	11
8	2	4	0	5	0	3
9	6	14	1	16	0	20
10	7	3	2	3	2	19
11	6	5	0	11	0	19
12	14	0	1	5	0	32
13	9	2	5	3	0	39
14	5	0	3	0	2	8
15	15	5	8	6	2	9
Total	141	63	43	104	7	271



I-7 Fig. I-4 Manhattan plots of genome-environment association for temperature and precipitation variables.

(**A**) Mean cropping month temperature; (**B**) Mean annual temperature; (**C**) Mean total cropping month precipitation; (**D**) Mean total annual precipitation.

I-8 Table I-4 Differentiation values of the five loci strongly associated with day length and latitude in original dataset (~51K SNPs, 12 populations)

Chromosome	Locus (position)	Nei and Chesser (1983)	Weir and Cockerham (1984)	Jost (2008)	Hedrick (2005)	Meirmans and Hedrick (2011)
10	7642230	0.5931 (0.4703 - 0.6964)	0.4872 (0.4046-0.5889)	0.4003 (0.3654-0.4422)	0.756 (0.6845 - 0.8277)	0.7685 (0.6981-0.8381)
14	6543734	0.7545 (0.6442 - 0.8465)	0.5324 (0.4454-0.6549)	0.3675 (0.3373-0.4431)	0.8447 (0.7914 - 0.9055)	0.8547 (0.8033-0.9124)
	6543744	0.7576 (0.6395 - 0.8384)	0.679 (0.5715-0.784)	0.3533 (0.2984-0.4287)	0.8432 (0.7621 - 0.8982)	0.8533 (0.7757-0.9056)
15	5662122	0.6239 (0.4903 - 0.7143)	0.55 (0.4585-0.6294)	0.4092 (0.3668-0.4486)	0.7778 (0.6907 - 0.8387)	0.7898 (0.704-0.8485)
	5662687	0.62 (0.5091 - 0.7072)	0.5511 (0.4599-0.6308)	0.3652 (0.3073-0.4477)	0.7588 (0.6786 - 0.8206)	0.7717 (0.693-0.8314)
%loci>min.Diff		28	16	1.8	12.1	12.3

%loc>min.Diff: percentage of loci with higher differentiation index than the least differentiated locus from the five loci above under each algorithm using the 51K dataset. Based on Jost (2008), only 1.8% of the 51575 SNPs were with higher differentiation than the least differentiated (D_jost=0.3533) locus 6543744 on chromosome 14. Numbers in parenthesis below differentiation value indicate confidence interval.

II Supplementary materials for chapter 3

II-1 Table II-1 List of superior genotypes for each traits that showed significant variations

Trait[1]	Population Mean±SD[2]	Superior genotype		Genotype Mean±SD[5]	Superiority over[6]
		TMP name[3]	Passport name[4]		
MaxR	8.94±3.37	TMP-2530	U_MAR_C_CN98193	19±7.07	16
		TMP-8216-5	U_CZE_C_CN98689	17.5±3.54	7
		TMP-1922	F_CHN_B_CN101419	15.67±2.31	5
MedR	3.99±1.65	TMP-2530	U_MAR_C_CN98193	10±4.24	15
		TMP-7590	O_RUS_C_CN97520	9.5±4.95	9
NWPer	510.2±163.8	TMP-2530	U_MAR_C_CN98193	1061.73±170.8	8
		TMP-1155	F_FRA_C_CN18986	820.83±212.21	1
		TMP-10020	O_CAN_B_CN101595	796.39±253.18	1
NWA	7.27±2.15	TMP-2530	U_MAR_C_CN98193	14.67±2.52	4
		TMP-1155	F_FRA_C_CN18986	11.86±0.61	1
NWL	245.3±77.5	TMP-2530	U_MAR_C_CN98193	513.66±73.18	7
		TMP-10020	O_CAN_B_CN101595	393.71±144.81	1
		TMP-1155	F_FRA_C_CN18986	391.42±52.7	1
NWSA	26.49±7.94	TMP-2530	U_MAR_C_CN98193	53.93±9.26	5
		TMP-1155	F_FRA_C_CN18986	43.62±2.41	1
NWV	0.252±0.074	TMP-2530	U_MAR_C_CN98193	0.48±0.1	1
SDWt	0.028±0.008	TMP-2530	U_MAR_C_CN98193	0.062±0.022	8
SL	7.91±1.30	TMP-8152-12	U_ARG_C_CN97341	13.27±3.1	9
SRL	976.6±116.4	TMP-1833	F_RUS_C_CN101094	1240.50±159.89	5
		TMP-1855	F_RUS_B_CN101116	1207.78±82.5	3
ARD	0.035±0.0001	TMP-8284	O_FRA_C_CN98752	0.041±0.002	3
		TMP-8168-3	U_USA_B_CN98644	0.04±0.003	3
		TMP-2650-14	O_TUR_L_CN96958	0.04±0.001	2
NWW_Depth	0.632±0.159	TMP-2530	U_MAR_C_CN98193	1.32±0.001	9
NWW	12.43±3.12	TMP-2530	U_MAR_C_CN98193	21.33±0.31	1

[1]Abrevation of the traits (see Table 1 for details and units)
[2]Overall mean and standard deviation (SD)
[3]TMP=temporary followed by temporary accession number at plant gene resource of Canada (PGRC), names used during experiment.
[4]Name in passport; naming convention indicates the type (O=oil type, F=fiber type, U= unknown), the originating country, the breeding status of the genotype (C=cultivar; B=breeding material; L=landrace) and accession number at PGRC.
[5]Mean and SD of the superior genotype
[6]Number of genotype over which the superior genotype has significantly higher value.

II-2 Table II-2 Pearson pairwise correlation of two shoot and 14 root traits.

Traits	SL	SDWt	RDWt	ARD	NWDep	NWDis	MaxR	NWW	MedR	NWA	NWPer	SRL	NWSA	NWL	NWV
SDWt	0.593***														
RDWt	0.513***	0.602***													
ARD	0.146*	0.169*	0.069												
NWDep	0.064	0.033	0.390***	-0.208**											
NWDis	0.136*	0.074	0.344***	0.231***	0.321***										
MaxR	0.324***	0.426***	0.498***	-0.322**	0.082	-0.069									
NWW	0.450***	0.471***	0.667***	-0.112	0.329***	0.246***	0.572***								
MedR	0.131*	0.275***	0.147*	-0.315**	-0.229***	-0.385***	0.693***	0.264***							
NWA	0.459***	0.572***	0.776***	-0.11	0.445***	0.114	0.707***	0.774***	0.413***						
NWPer	0.404***	0.523***	0.721***	-0.304**	0.452***	0.07	0.773***	0.762***	0.478***	0.974***					
SRL	-0.147*	-0.163*	-0.089	-0.981**	0.155*	-0.225***	0.302***	0.068	0.319***	0.07	0.269***				
NWSA	0.457***	0.571***	0.775***	-0.107	0.442***	0.118	0.697***	0.774***	0.405***	0.999***	0.972***	0.065			
NWL	0.409***	0.523***	0.733***	-0.279**	0.460***	0.092	0.736***	0.767***	0.441***	0.981***	0.996***	0.239***	0.981***		
NWV	0.475***	0.580***	0.780***	0.07	0.414***	0.145*	0.631***	0.745***	0.350***	0.978***	0.909***	-0.115	0.980***	0.925***	
NWW_Dep	0.409***	0.455***	0.434***	-0.016	-0.213**	0.04	0.533***	0.831***	0.445***	0.535***	0.520***	-0.001	0.536***	0.520***	0.521***

*P<0.05; **P<0.01, ***P<0.001

II-3 Table II-3 QTN detected in both 3 and 7K datasets using seven models

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
ARD	Chr1:250405	pLARmEB	3	23.2	0.57	21.78	Inf	Lus10036136	AT3G45640	MAPK3
								Lus10036135	AT3G07630	ADT2
NWW	Chr1:513047	LFMM	7				2.22E-05			
RDWt	Chr1:741582	pLARmEB	7	10.5	0.05	7.92	3.82E-12	Lus10036052	AT2G26780	ARM
RDWt	Chr1:756854	pKWmEB	3	3.1	6.32	6.06	1.59E-04	Lus10036050	AT5G19360	CPK34
NWW	Chr1:756854	ISIS EM-BLASSO	7	6.0	13.07	5.94	1.55E-07	Lus10036049	AT1G62350	PPR
NWW_Dep	Chr1:756854	ISIS EM-BLASSO	7	4.7	11.04	5.94	3.27E-06			
NWW	Chr1:756854	LFMM	7				6.65E-05			
NWW	Chr1:756854	pLARmEB	7	3.1	13.07	5.94	1.43E-04			
RDWt	Chr1:1122450	pLARmEB	7	4.8	0.00	1.98	2.42E-06	Lus10035975	AT2G17270	PHT3;3
RDWt	Chr1:3031443	pLARmEB	7	18.3	0.02	1.98	Inf	Lus10042414	AT4G31820	NPY1
RDWt	Chr1:4908649	pKWmEB	3	4.8	8.50	7.07	2.39E-06	Lus10002819	AT5G02910	RIN_FB
								Lus10002820	AT3G58930	RIN_FB
								Lus10002822	AT2G15320	LRR
								Lus10002825	AT4G34150	CaLB
								Lus10002828	AT3G19320	LRR
NWW_Dep	Chr1:7339896	LFMM	7				5.39E-07	Lus10029174	AT3G53710	AGD6
NWL	Chr1:9613871	LFMM	3				1.29E-05	Lus10004206	AT3G44350	NAC061
NWPer	Chr1:9613871	LFMM	3				7.43E-06	Lus10004214	AT3G11620	ABH
NWSA	Chr1:9613871	LFMM	3				8.16E-06	Lus10004225	AT3G11570	TBL8
NWV	Chr1:9613871	LFMM	3				8.57E-06			
NWW	Chr1:9613871	LFMM	3				1.52E-05			
RDWt	Chr1:9613871	LFMM	3				2.90E-05			
NWSA	Chr1:9613871	LFMM	7				1.46E-05			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
NWA	Chr1:9613871	LFMM	7				6.94E-06			
NWW	Chr1:9613871	LFMM	7				2.29E-07			
RDWt	Chr1:9613871	LFMM	7				8.91E-06			
MedR	Chr1:11064283	mrMLM	3	3.9	7.82	7.07	2.09E-05	Lus10016129	AT2G28085	SAUR
RDWt	Chr1:11839052	pKWmEB	3	3.5	3.66	29.29	5.29E-05	Lus10009383	AT5G22760	PHD
RDWt	Chr1:12047870	pLARmEB	7	12.0	0.01	1.98	1.12E-13	Lus10009398	AT2G29940	PDR3
RDWt	Chr1:18212245	pKWmEB	3	4.0	0.36	9.09	1.77E-05			
ARD	Chr1:18434560	pLARmEB	3	10.5	0.33	12.87	3.48E-12			
ARD	Chr1:18434560	ISIS EM-BLASSO	7	3.7	0.06	12.87	3.49E-05			
SDWt	Chr1:18858986	FASTmrMLM	3	4.3	0.01	6.93	7.65E-06			
SDWt	Chr1:18970469	mrMLM	3	3.2	10.65	8.08	1.37E-04	Lus10018180	AT1G77600	ARM
SDWt	Chr1:18970469	FASTmrMLM	7	6.0	0.00	7.92	1.46E-07			
RDWt	Chr1:19577662	pLARmEB	7	4.0	0.07	7.92	1.93E-05	Lus10020829	AT2G01420	PIN4
SDWt	Chr1:20356976	pKWmEB	3	6.7	8.04	6.06	2.60E-08	Lus10015898	AT2G13600	PPR
SDWt	Chr1:20356976	FASTmrMLM	3	19.8	1.74	5.94	Inf	Lus10015894	AT5G65450	UBP17
RDWt	Chr1:20356976	FASTmrMLM	3	4.5	0.44	5.94	5.41E-06	Lus10015893	AT4G24560	UBP16
RDWt	Chr1:20356976	ISIS EM-BLASSO	3	3.4	0.39	5.94	8.16E-05			
SDWt	Chr1:20356976	FASTmrMLM	7	22.8	0.62	5.94	Inf			
RDWt	Chr1:20356976	pLARmEB	7	4.8	0.26	5.94	2.31E-06			
SDWt	Chr1:20356976	pLARmEB	7	3.2	0.00	5.94	1.13E-04			
ARD	Chr1:21006827	pLARmEB	3	23.4	0.03	9.9	Inf			
SDWt	Chr1:22813857	FASTmrMLM	7	3.4	0.01	6.93	7.26E-05	Lus10014632	AT2G14820	NPY2
SDWt	Chr1:25703627	FASTmrMLM	7	4.3	0.01	14.85	7.97E-06	Lus10003106/7	AT3G04290	LTL1
NWDis	Chr2:4513304	ISIS EM-BLASSO	3	3.4	11.49	5.94	7.22E-05	Lus10020193	AT5G16530	PIN5
NWDis	Chr2:4513304	FASTmrMLM	7	3.2	0.00	5.94	1.29E-04	Lus10013610	AT1G65620	AS2

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
ARD	Chr2:5139403	pLARmEB	3	4.9	0.09	6.93	2.16E-06	Lus10038128	AT5G20690	LRR
ARD	Chr2:5452843	pLARmEB	3	4.4	0.17	17.82	7.16E-06	Lus10038110	AT5G20380	PHT4;5
ARD	Chr2:5529143	pLARmEB	3	27.0	0.06	5.94	Inf			
NWW	Chr2:5554181	LFMM	7				3.57E-05			
NWW	Chr2:5554185	LFMM	7				3.57E-05			
NWW_Dep	Chr2:5963452	FASTmrMLM	3	3.9	14.88	9.9	2.25E-05	Lus10038046	AT3G15580	UBP
SDWt	Chr2:6056808	FASTmrMLM	3	10.4	0.17	5.94	4.80E-12	Lus10038025	AT3G62100	IAA30
NWW	Chr2:6662071	LFMM	3				2.92E-05	Lus10038699	AT3G54850	PUB14
NWW	Chr2:6662071	LFMM	7				1.88E-06	Lus10038696	AT1G01950	ARK2
NWDis	Chr2:7095057	mrMLM	3	3.4	10.54	10.1	7.81E-05	Lus10038670	AT3G10490	NAC052
SDWt	Chr2:16999042	FASTmrMLM	7	10.7	0.00	0.99	2.22E-12			
NWW	Chr2:18121207	pLARmEB	7	3.2	4.39	8.91	1.27E-04	Lus10025840	AT5G66870	ASL1
ARD	Chr2:18237395	pLARmEB	3	3.2	0.06	6.93	1.22E-04	Lus10033161	AT1G75590	SAUR
RDWt	Chr2:21629463	pLARmEB	7	13.2	0.59	13.86	7.00E-15			
SDWt	Chr2:22207057	FASTmrMLM	3	3.0	0.02	9.9	2.01E-04			
SDWt	Chr2:22207057	FASTmrMLM	7	7.7	0.00	9.9	2.37E-09			
RDWt	Chr2:22305709	pLARmEB	7	6.2	0.09	10.89	8.44E-08			
SDWt	Chr2:23342970	ISIS EM-BLASSO	3	3.7	0.02	6.93	3.96E-05	Lus10013276	AT1G69850	NRT1:2
SDWt	Chr2:23342970	FASTmrMLM	7	10.9	0.03	6.93	1.35E-12			
SDWt	Chr2:25271724	LFMM	7				8.61E-06			
SDWt	Chr2:25271724	pLARmEB	7	4.6	0.01	2.97	3.99E-06			
ARD	Chr3:1231621	pLARmEB	3	6.9	0.01	5.94	1.60E-08	Lus10037425	AT3G43790	ZIFL2
SDWt	Chr3:1363801	FASTmrMLM	7	37.7	0.00	10.89				
SDWt	Chr3:2215745	FASTmrMLM	7	13.7	0.00	5.94	1.78E-15			
SDWt	Chr3:3857902	FASTmrMLM	3	3.1	0.00	5.94	1.54E-04			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
ARD	Chr3:4092867	pLARmEB	3	5.5	0.45	17.82	4.52E-07			
SDWt	Chr3:4897266	FASTmrMLM	7	17.8	0.00	5.94				
RDWt	Chr3:5009686	pLARmEB	7	5.6	0.03	9.9	3.38E-07			
RDWt	Chr3:5432287	pLARmEB	7	7.8	0.08	12.87	2.06E-09			
ARD	Chr3:6835049	pLARmEB	3	18.1	0.04	7.92				
MaxR	Chr3:6925560	FASTmrMLM	7	3.8	13.34	3.96	3.04E-05	Lus10040829	AT2G28840	XBAT31
MedR	Chr3:6925560	ISIS EM-BLASSO	7	3.9	9.00	3.96	2.48E-05	Lus10040826	AT3G12110	ACT11
NWSA	Chr3:6925560	ISIS EM-BLASSO	7	4.0	10.04	3.96	1.67E-05	Lus10040817	AT5G07900	MTERF
NWL	Chr3:6925560	ISIS EM-BLASSO	7	4.4	11.55	3.96	6.40E-06	Lus10040818		
NWA	Chr3:6925560	pLARmEB	7	3.7	9.21	3.96	3.64E-05	Lus10040819		
NWL	Chr3:6925560	pLARmEB	7	4.4	11.41	3.96	6.40E-06	Lus10040820		
NWPer	Chr3:6925560	pLARmEB	7	4.1	10.31	3.96	1.33E-05			
SDWt	Chr3:7074112	FASTmrMLM	7	3.3	0.16	6.93	9.82E-05			
SDWt	Chr3:7668335	FASTmrMLM	7	5.8	0.09	4.95	2.22E-07			
ARD	Chr3:10151903	pLARmEB	3	4.5	0.14	5.94	5.57E-06			
NWW	Chr3:16939026	ISIS EM-BLASSO	7	4.6	8.89	3.96	4.21E-06	Lus10033528	AT4G39370	UBP27
NWW	Chr3:16939026	pLARmEB	7	4.2	8.89	3.96	1.16E-05	Lus10033524	AT2G40890	CYP98A3
SDWt	Chr3:17343476	mrMLM	3	3.7	7.48	8.08	3.68E-05	Lus10033446	AT4G08920	CRY1,
ARD	Chr3:17343476	pLARmEB	3	18.2	0.06	9.9	Inf	Lus10033447	AT1G31885	NIP3;1
RDWt	Chr3:18363136	pLARmEB	7	11.4	0.02	4.95	4.00E-13			
NWDis	Chr3:18772054	ISIS EM-BLASSO	7	4.2	9.68	4.95	1.08E-05			
NWW_Dep	Chr3:18933706	LFMM	3				2.31E-05			
RDWt	Chr3:22145720	pKWmEB	3	5.2	0.82	7.07	9.16E-07			
ARD	Chr3:25380098	FASTmrMLM	3	8.7	0.05	12.87	2.21E-10			
ARD	Chr3:25380098	FASTmrMLM	7	9.6	0.11	12.87	3.30E-11			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
NWDis	Chr3:25380098	mrMLM	3	3.7	12.72	13.13	3.90E-05	Lus10037572	AT4G32890	GATA9
ARD	Chr4:2910990	pLARmEB	3	26.9	0.18	10.89	Inf	Lus10029520	AT5G12870	MYB46
								Lus10029525	AT3G57670	NTT,WIP2
ARD	Chr4:4572761	pLARmEB	3	8.0	0.03	5.94	1.40E-09	Lus10002459	AT1G64960	ARM
								Lus10002458	AT4G37640	ACA2
ARD	Chr4:8061253	pLARmEB	3	11.1	0.03	19.8	8.53E-13	Lus10039524	AT3G57040	RR9,RR4
RDWt	Chr4:8745388	pLARmEB	7	5.5	0.12	6.93	5.29E-07			
NWW_Dep	Chr4:11667975	LFMM	3				8.24E-05	Lus10036736	AT1G26900	PPR
NWW	Chr4:11667975	LFMM	7				8.66E-06	Lus10036742	AT1G14300	ARM
NWW_Dep	Chr4:11667975	LFMM	7				4.62E-06			
RDWt	Chr4:12310229	pLARmEB	7	4.8	0.06	5.94	2.59E-06			
SDWt	Chr4:17127643	ISIS EM-BLASSO	7	4.7	0.00	1.98	3.39E-06			
SDWt	Chr4:17127643	pLARmEB	7	4.0	0.00	1.98	1.82E-05			
NWDep	Chr4:17242614	mrMLM	3	5.1	22.68	7.07	1.17E-06	Lus10020805	AT5G37020	ARF8
NWDep	Chr4:17242614	ISIS EM-BLASSO	3	3.7	16.61	6.93	3.61E-05	Lus10020804	AT5G37020	ARF8
NWDep	Chr4:17242614	FASTmrMLM	3	4.3	16.61	6.93	9.16E-06	Lus10020794	AT5G10360	RPS6B
NWDep	Chr4:17242614	pLARmEB	3	3.6	7.18	6.93	4.17E-05	Lus10020791	AT1G73370	SUS6
NWDep	Chr4:17242614	LFMM	3				2.69E-05			
NWDep	Chr4:17242614	ISIS EM-BLASSO	7	5.1	16.45	6.93	1.37E-06			
NWDis	Chr4:17408739	FASTmrEMMA	3	3.2	0.00	12.87	1.40E-04	Lus10020755	AT5G50890	ABH
NWDis	Chr4:17408739	FASTmrEMMA	7	3.1	0.00	12.87	1.59E-04			
RDWt	Chr4:17943797	pLARmEB	7	8.5	0.08	4.95	3.89E-10			
SRL	Chr4:18399285	pKWmEB	3	3.3	15.60	10.1	9.60E-05	Lus10029995	AT2G05810	ARM
SRL	Chr4:18399285	FASTmrMLM	3	4.8	11.05	9.9	2.91E-06	Lus10029996	AT2G05810	ARM
SRL	Chr4:18399285	pLARmEB	3	4.9	10.65	9.9	2.16E-06			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
SRL	Chr4:18399285	ISIS EM-BLASSO	3	4.9	10.65	9.9	2.16E-06			
ARD	Chr4:18399285	pLARmEB	3	19.4	2.27	9.9	Inf			
ARD	Chr4:18399285	ISIS EM-BLASSO	3	6.1	1.90	9.9	1.10E-07	Lus10029686	AT3G49600	UBP26
SDWt	Chr5:504766	FASTmrMLM	7	12.0	0.00	5.94	1.08E-13			
RDWt	Chr5:731588	pLARmEB	7	4.7	0.03	5.94	2.96E-06			
MedR	Chr5:1375386	mrMLM	3	4.0	12.94	9.09	1.76E-05	Lus10004757	AT5G47330	ABH
RDWt	Chr5:1375386	ISIS EM-BLASSO	3	3.0	0.76	8.91	1.99E-04			
RDWt	Chr5:1375386	FASTmrMLM	3	5.3	0.39	8.91	7.16E-07			
ARD	Chr5:1610010	pLARmEB	3	13.2	0.01	10.89	6.66E-15	Lus10008507	AT4G16520	G8F
NWSA	Chr5:2645287	LFMM	3				4.36E-05	Lus10032249	AT4G12440	APT4
RDWt	Chr5:2645287	FASTmrMLM	7	4.2	0.15	6.93	1.04E-05	Lus10032252	AT1G12110	NRT1
SDWt	Chr5:2645287	FASTmrMLM		11.1	0.01	6.93	9.69E-13			
RDWt	Chr5:2645287	ISIS EM-BLASSO		4.2	21.88	6.93	1.21E-05			
NWW	Chr5:2645287	LFMM					9.69E-06			
RDWt	Chr5:2645287	LFMM					4.83E-06			
RDWt	Chr5:9226031	pLARmEB	7	23.1	0.03	10.89	Inf			
SDWt	Chr5:9373438	FASTmrMLM	7	3.5	0.00	9.9	6.31E-05			
ARD	Chr5:10460806	ISIS EM-BLASSO	3	3.9	0.33	7.92	2.50E-05			
ARD	Chr5:10619678	pLARmEB	3	13.7	0.03	5.94	2.00E-15	Lus10021224	AT5G43630	TZP
MedR	Chr5:11019409	mrMLM	3	3.4	10.32	6.06	7.46E-05	Lus10008763	AT5G14580	PNP
								Lus10014903	AT4G26690	GPDL2
ARD	Chr5:12067166	pLARmEB	3	42.6	0.02	6.93	Inf	Lus10035312	AT2G13620	CHX15
NWDep	Chr5:15312783	mrMLM	3	5.6	19.26	15.15	4.03E-07	Lus10028301	AT3G50870	GATA18
NWDep	Chr5:15312783	FASTmrMLM	3	4.9	15.27	14.85	2.09E-06	Lus10024009	AT1G75250	RL6,RSM3
NWDep	Chr5:15312783	ISIS EM-BLASSO	3	4.2	15.27	14.85	1.20E-05	Lus10024014	AT5G66770	GRAS

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
NWDep	Chr5:15312783	FASTmrEMMA	3	3.9	13.35	14.85	2.53E-05	Lus10024022	AT4G36650	PBRP
NWSA	Chr5:15312783	mrMLM	3	3.1	11.48	15.15	1.67E-04	Lus10024023	AT5G25620	YUC6
NWDep	Chr5:15312783	pLARmEB	3	5.6	6.60	14.85	3.97E-07			
NWL	Chr5:15312783	pLARmEB	3	3.1	5.61	14.85	1.50E-04			
NWSA	Chr5:15312783	pLARmEB	3	3.1	4.97	14.85	1.67E-04			
NWA	Chr5:15312783	pLARmEB	3	3.1	4.64	14.85	1.70E-04			
NWDep	Chr5:15312783	LFMM	3				6.29E-06			
NWDep	Chr5:15312783	FASTmrEMMA	7	3.0	0.00	14.85	2.01E-04			
NWDep	Chr5:15312783	ISIS EM-BLASSO	7	5.7	10.41	14.85	2.82E-07			
NWDep	Chr5:15312783	mrMLM	7	3.0	15.48	15.15	2.01E-04			
RDWt	Chr5:15312783	pLARmEB	7	13.7	0.32	14.85	2.11E-15			
ARD	Chr5:15359894	pLARmEB	3	8.3	0.03	6.93	6.11E-10			
ARD	Chr5:15570504	pLARmEB	3	12.8	0.05	7.92	1.69E-14	Lus10024057	AT1G55250	HUB2
RDWt	Chr5:16359682	pLARmEB	7	3.9	0.56	12.87	2.34E-05			
SDWt	Chr6:3310382	pKWmEB	3	3.3	17.64	13.13	9.93E-05	Lus10019475/6	AT3G15354	SAP3
SDWt	Chr6:3310382	FASTmrMLM	7	6.6	0.35	12.87	3.44E-08	Lus10019474	AT3G48250	PPR
SDWt	Chr6:3310382	pLARmEB	7	3.3	11.45	12.87	9.42E-05	Lus10019471	AT4G24690	UBA/TS-N
								Lus10019470	AT5G52450	MATE efflux
NWW_Dep	Chr6:6417516	LFMM	3				1.28E-05	Lus10017778	AT5G08020	RPA70B
NWW_Dep	Chr6:6427626	LFMM	3				1.28E-05			
RDWt	Chr6:6698595	pLARmEB	7	36.4	0.03	3.96	Inf			
SDWt	Chr6:6704279	FASTmrMLM	7	8.4	0.00	1.98	5.06E-10			
RDWt	Chr6:6704279	pLARmEB	7	43.2	0.01	1.98	Inf			
SRL	Chr6:7732273	ISIS EM-BLASSO	7	3.2	5.47	9.9	1.37E-04			
NWW	Chr6:12030162	LFMM	3				5.78E-05			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
NWW	Chr6:12030162	LFMM	7				1.88E-06			
RDWt	Chr6:13079727	pLARmEB	7	11.6	0.01	1.98	2.79E-13			
RDWt	Chr6:13140060	pLARmEB	7	3.9	0.03	3.96	2.51E-05			
SDWt	Chr6:14867203	FASTmrMLM	7	8.0	0.01	9.9	1.16E-09			
RDWt	Chr6:15202179	pKWmEB	3	4.4	1.81	8.08	6.70E-06	Lus10014380/1	AT2G36400	GRF3
RDWt	Chr6:15418446	pLARmEB	7	12.9	0.00	6.93	1.19E-14			
NWW	Chr6:16928785	LFMM	3				2.92E-05	Lus10025270	AT3G50530	CRK
NWW_Dep	Chr6:16928785	LFMM	3				1.71E-04			
NWW	Chr6:16928785	LFMM	7				2.68E-06			
RDWt	Chr6:17011819	pLARmEB	7	6.9	0.01	10.89	1.80E-08			
RDWt	Chr7:1779542	pKWmEB	3	3.3	1.98	8.08	1.05E-04	Lus10028765	AT4G01370	MPK4
NWDis	Chr7:4694299	LFMM	7				8.56E-06			
NWDis	Chr7:4694715	LFMM	7				8.56E-06			
MedR	Chr7:4774423	mrMLM	3	3.0	7.72	6.06	1.90E-04	Lus10040127	AT1G73500	MKK9
								Lus10040125	AT3G02840	ARM
RDWt	Chr7:6087128	pLARmEB	7	8.1	0.03	3.96	1.00E-09			
NWW_Dep	Chr7:6346464	mrMLM	3	3.9	13.45	10.1	2.27E-05	Lus10008023	AT1G63460	GPX8
NWW_Dep	Chr7:6346464	ISIS EM-BLASSO	3	3.7	7.32	9.9	3.87E-05	Lus10008025	AT1G80600	WIN1
NWW_Dep	Chr7:6346464	FASTmrMLM	3	3.3	0.00	9.9	1.00E-04	Lus10008022	AT4G11600	GPX6
NWW_Dep	Chr7:6346464	FASTmrEMMA	7	3.7	0.00	9.9	3.72E-05			
ARD	Chr7:6346464	FASTmrMLM	7	5.0	0.21	9.9	1.57E-06			
NWDep	Chr7:6759944	ISIS EM-BLASSO	7	4.6	0.00	2.97	4.06E-06			
ARD	Chr7:15286061	pLARmEB	3	11.9	0.03	5.94	1.42E-13	Lus10038472	AT3G45640	MAPK3
								Lus10038460	AT5G12180	CPK17
ARD	Chr7:15553408	pLARmEB	3	50.0	0.06	6.93	Inf	Lus10038426	AT2G34830	WRKY35

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
NWW	Chr7:16072649	LFMM	7				8.20E-05			
ARD	Chr7:16356085	pLARmEB	3	4.2	0.05	5.94	1.07E-05	Lus10020049	AT3G21690	MATE efflux
ARD	Chr7:16356085	FASTmrMLM	7	4.5	0.04	5.94	4.82E-06			
RDWt	Chr7:17879110	pLARmEB	7	8.1	0.01	3.96	1.04E-09			
ARD	Chr8:3103449	pLARmEB	3	27.4	0.19	9.9	Inf			
RDWt	Chr8:4754353	FASTmrMLM	3	3.6	0.08	5.94	4.77E-05	Lus10023970	AT1G29450	SAUR
ARD	Chr8:4754354	pLARmEB		25.0	0.00	5.94	Inf	Lus10023966	AT2G17040	NAC036
RDWt	Chr8:10360902	pLARmEB	7	18.6	0.00	2.97	Inf			
SDWt	Chr8:14140850	FASTmrMLM	3	8.6	0.09	7.92	3.38E-10	Lus10016595	AT4G26640	WRKY20
								Lus10016596	AT5G48150	PAT1
SDWt	Chr8:15837449	FASTmrMLM	3	3.5	0.11	8.91	6.12E-05	Lus10022218	AT5G09690	MGT7
ARD	Chr8:17030355	pLARmEB	3	18.7	0.03	5.94	Inf	Lus10003016	AT1G13570	F_Box
RDWt	Chr8:17777940	pLARmEB	7	3.6	0.01	3.96	4.27E-05			
ARD	Chr8:20840576	pLARmEB	3	26.1	0.02	5.94	Inf	Lus10033932	AT3G20820	LRR
								Lus10033937	AT5G53300	UBC10
SRL	Chr8:21825897	pLARmEB	3	3.4	5.00	13.86	7.84E-05	Lus10010569	AT3G22560	AcylCoANAT
SRL	Chr8:21825897	ISIS EM-BLASSO	3	3.4	5.00	13.86	7.84E-05	Lus10010570	AT2G32030	AcylCoANAT
RDWt	Chr8:22357078	pLARmEB	7	7.0	0.09	15.84	1.31E-08			
SDWt	Chr8:22898342	FASTmrMLM	7	8.8	0.00	3.96	1.85E-10			
RDWt	Chr8:22919544	pKWmEB	3	6.1	0.00	6.06	1.05E-07	Lus10015412	AT5G26990	Drought
ARD	Chr9:630908	pLARmEB	3	26.4	0.23	9.9		Lus10010171	AT3G19970	ABH
RDWt	Chr9:1259084	pLARmEB	7	9.5	0.05	5.94	3.54E-11			
RDWt	Chr9:2989474	FASTmrMLM	3	3.1	0.01	6.93	1.62E-04	Lus10010403	AT1G32100	PRR1
NWW_Dep	Chr9:4628185	FASTmrMLM	7	3.7	0.00	15.84	3.98E-05			
RDWt	Chr9:5639608	pLARmEB	7	5.6	0.12	9.9	3.94E-07			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
RDWt	Chr9:15946848	pKWmEB	3	5.3	9.16	7.07	7.31E-07	Lus10021733	AT5G42870	PAH2
RDWt	Chr9:15946848	pLARmEB	3	4.0	0.99	6.93	1.66E-05			
SDWt	Chr9:15946848	FASTmrMLM	7	8.5	0.10	6.93	3.54E-10			
RDWt	Chr9:15946848	pLARmEB	7	66.2	0.87	6.93	Inf			
ARD	Chr9:16446802	pLARmEB	3	4.8	0.07	11.88	2.70E-06	Lus10042646	AT1G46264	SCZ
ARD	Chr9:16933932	pLARmEB	3	9.7	0.24	8.91	2.30E-11	Lus10042597	AT3G56800	CAM3
RDWt	Chr9:18388080	pLARmEB	7	14.1	0.09	5.94	7.77E-16			
SDWt	Chr9:19039213	FASTmrMLM	7	6.5	0.00	4.95	4.20E-08			
RDWt	Chr9:19040305	pLARmEB	7	6.8	0.01	2.97	1.95E-08			
ARD	Chr9:19061342	pLARmEB	3	9.8	0.13	11.88	1.87E-11	Lus10024830	AT4G14640	CAM8
NWL	Chr9:19061342	LFMM	3				4.27E-06	Lus10024833	AT2G38290	AMT2;1
NWPer	Chr9:19061342	LFMM	3				2.43E-06	Lus10024853	AT1G04240	IAA3
NWSA	Chr9:19061342	LFMM	3				1.91E-05			
NWW	Chr9:19061342	LFMM	3				1.28E-05			
NWW_Dep	Chr9:19061342	LFMM	3				3.67E-05			
NWPer	Chr9:19061342	FASTmrMLM	7	4.0	24.20	11.88	1.98E-05			
NWSA	Chr9:19061342	LFMM	7				2.12E-05			
NWA	Chr9:19061342	LFMM	7				1.07E-05			
NWL	Chr9:19061342	LFMM	7				7.59E-06			
NWPer	Chr9:19061342	LFMM	7				5.24E-06			
NWW	Chr9:19061342	LFMM	7				3.76E-07			
NWW_Dep	Chr9:19061342	LFMM	7				1.62E-06			
RDWt	Chr9:19061342	LFMM	7				9.70E-06			
RDWt	Chr9:19295240	FASTmrMLM	3	5.3	0.24	5.94	8.59E-07	Lus10024860	AT5G09690	MGT7
								Lus10024864	AT2G30590	WRKY21

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
								Lus10024877	AT2G42430	ASL18
ARD	Chr9:19859864	pLARmEB	3	11.3	0.06	8.91	5.24E-13	Lus10011913	AT2G20260	PSAE-2
ARD	Chr9:21921657	ISIS EM-BLASSO	3	3.3	0.05	5.94	1.03E-04	Lus10005686	AT1G08680.4	AGD14
ARD	Chr9:21921657	pLARmEB	3	88.8	0.03	5.94	Inf			
ARD	Chr9:21925137	pLARmEB	3	4.7	0.03	5.94	3.00E-06			
RDWt	Chr10:5257573	pKWmEB	3	3.4	2.33	8.08	7.69E-05	Lus10023805	AT5G45400	RPA70C
SDWt	Chr10:5257573	FASTmrMLM	7	6.3	0.02	7.92	6.90E-08			
RDWt	Chr10:5257573	pLARmEB	7	10.1	0.23	7.92	9.83E-12			
RDWt	Chr10:6106628	pKWmEB	3	4.0	2.92	7.07	1.82E-05	Lus10015037	AT3G09550	Ankyrin
RDWt	Chr10:6247028	pLARmEB	7	8.1	0.92	11.88	1.12E-09			
SDWt	Chr10:6457226	pKWmEB	3	3.0	2.05	6.06	1.88E-04	Lus10015076	AT1G61110	NAC025
RDWt	Chr10:7751366	pLARmEB	7	5.7	0.04	3.96	3.15E-07			
ARD	Chr10:9207013	pLARmEB	3	7.0	0.18	9.9	1.38E-08	Lus10039986	AT3G07490	AGD11
ARD	Chr10:11511478	pLARmEB	3	12.1	0.01	8.91	8.44E-14	Lus10032747	AT5G25900	CYP701A3
ARD	Chr10:13371316	pLARmEB	3	22.7	0.12	9.9	Inf			
ARD	Chr10:13391033	pLARmEB	3	10.8	0.07	9.9	1.91E-12			
RDWt	Chr10:14815706	pLARmEB	7	5.0	0.01	2.97	1.61E-06			
ARD	Chr10:16154433	ISIS EM-BLASSO	7	4.8	0.14	4.95	2.45E-06			
RDWt	Chr10:16677317		3	3.3	0.00	17.82	1.02E-04	Lus10022812	AT2G33150.1	PKT3
RDWt	Chr10:16677317		7	42.4	0.84	17.82	Inf			
ARD	Chr10:16680000	pLARmEB	3	28.5	0.05	19.8	Inf			
ARD	Chr10:16956264	pLARmEB	3	25.5	0.53	16.83	Inf			
ARD	Chr10:16957410	pLARmEB	3	5.5	0.88	17.82	4.88E-07			
SDWt	Chr10:17003973	FASTmrMLM	3	8.3	0.02	5.94	6.00E-10	Lus10020408	AT4G12970	EPFL9
SDWt	Chr10:17003973	FASTmrMLM	7	32.7	0.01	5.94	Inf			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
SDWt	Chr11:578136	FASTmrMLM	7	7.3	0.00	2.97	5.98E-09			
RDWt	Chr11:3194902	pLARmEB	7	14.8	0.02	4.95	1.11E-16			
SDWt	Chr11:3430032	FASTmrMLM	7	14.7	0.00	6.93	2.22E-16			
ARD	Chr11:3783939	pLARmEB	3	7.8	0.17	6.93	2.25E-09	Lus10042185	AT1G74740	CPK30
								Lus10042187	AT3G47870	LBD27,SCP
NWW_Dep	Chr11:5123720	LFMM	7				1.13E-07			
MedR	Chr11:5382153	FASTmrMLM	3	3.2	8.27	5.94	1.15E-04	Lus10038217	AT3G07330	CSLC6
								Lus10038218	AT3G07360	PUB9
								Lus10038224	AT3G07390	ARF-AIR12
								Lus10038225	AT3G25290	ARF
SDWt	Chr11:6948325	FASTmrMLM	7	5.5	0.00	7.92	4.93E-07			
SDWt	Chr11:8154007	ISIS EM-BLASSO	3	4.8	16.53	7.92	2.68E-06	Lus10036375	AT2G46050	PPR
SDWt	Chr11:8154007	LFMM	3				7.74E-06	Lus10036374	AT4G00350	MATE efflux
SDWt	Chr11:8154007	FASTmrMLM	7	3.5	0.86	7.92	5.82E-05			
SDWt	Chr11:8154007	LFMM	7				7.21E-06			
SDWt	Chr11:8176149	LFMM	3				2.24E-06	Lus10036372	AT1G63800	UBC5
SDWt	Chr11:8176149	FASTmrMLM	7	4.3	2.78	6.93	9.41E-06			
SDWt	Chr11:8176149	LFMM	7				1.50E-06			
RDWt	Chr11:14272770	pLARmEB	7	19.9	0.01	9.9	Inf			
ARD	Chr11:15115639	pLARmEB	3	34.0	0.02	7.92	Inf			
SDWt	Chr11:15133773	FASTmrMLM	3	9.2	0.00	6.93	7.15E-11	Lus10012927	AT4G29230	NAC075
NWV	Chr12:255713	ISIS EM-BLASSO	3	3.2	11.40	5.94	1.13E-04	Lus10019965	AT2G23810	TET8
NWV	Chr12:255713	pLARmEB	3	3.1	11.34	5.94	1.49E-04	Lus10019982	AT4G17770	TPS5
NWV	Chr12:255749	LFMM	3				4.70E-05			
SDWt	Chr12:348747	FASTmrMLM	7	3.6	0.00	2.97	4.21E-05			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
RDWt	Chr12:348747	pLARmEB	7	31.2	0.06	2.97	Inf			
SDWt	Chr12:799432	FASTmrMLM	3	4.9	0.00	5.94	2.10E-06	Lus10006759	AT4G21200	GA2OX8
SDWt	Chr12:895847	FASTmrMLM	7	8.1	0.00	3.96	1.05E-09			
SDWt	Chr12:2003623	FASTmrMLM	3	4.4	0.10	11.88	6.44E-06	Lus10023297	AT4G18390	TCP2
SDWt	Chr12:2003623	FASTmrMLM	7	16.3	0.12	11.88	Inf	Lus10023283	AT2G34640	PTAC12
SDWt	Chr12:3091575	FASTmrMLM	7	12.2	0.03	7.92	7.33E-14			
SL	Chr12:3690290	ISIS EM-BLASSO	3	3.3	12.58	5.94	Inf	Lus10001638	AT4G29900	ACA10
								Lus10001637	AT5G57090	PIN2
ARD	Chr12:5504203	pLARmEB	3	21.5	0.13	6.93	Inf	Lus10037827	AT1G04850	UBA
								Lus10037820	AT3G17205	UPL6
ARD	Chr12:5870441	pLARmEB	3	31.3	0.09	12.87	Inf	Lus10037785	AT5G49520	WRKY48
RDWt	Chr12:6352775	pLARmEB	7	59.5	0.13	7.92	Inf			
ARD	Chr12:6618716	pLARmEB	3	10.1	0.02	5.94	9.04E-12	Lus10006523	AT3G16510	CaLB
ARD	Chr12:7572669	pLARmEB	3	18.7	0.04	9.9	Inf			
SDWt	Chr12:12200657	mrMLM	3	3.5	10.05	18.18	5.80E-05	Lus10033041	AT2G31880	SOBIR1
								Lus10033043	AT3G01650	RGLG1
								Lus10033044	AT2G26710	CYP72B1
SDWt	Chr12:17419506	FASTmrMLM	7	4.6	0.00	2.97	4.52E-06			
RDWt	Chr12:20228564	pKWmEB	3	3.3	1.71	6.06	9.16E-05	Lus10031478	AT4G12560	CPR30
ARD	Chr12:20244330	ISIS EM-BLASSO	3	3.0	0.40	7.92	1.92E-04	Lus10031480	AT4G12560	CPR30
RDWt	Chr13:1344162	pLARmEB	7	13.4	0.00	6.93	3.66E-15			
RDWt	Chr13:2333295	pLARmEB	7	13.7	0.00	4.95	1.78E-15			
ARD	Chr13:3031249	pLARmEB	3	16.0	0.10	7.92	Inf	Lus10025986	AT2G18170	MPK7
RDWt	Chr13:3157081	pKWmEB	3	4.4	0.00	12.12	6.90E-06			
ARD	Chr13:6621308	pLARmEB	3	10.5	0.17	6.93	3.21E-12	Lus10034544	AT1G20160	ATSBT5.2

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
SDWt	Chr13:13450469	FASTmrMLM	7	9.4	0.01	1.98	4.37E-11			
SDWt	Chr13:15851979	LFMM	7				8.06E-06			
RDWt	Chr13:15851979	pLARmEB	7	4.1	0.01	1.98	1.25E-05			
SDWt	Chr13:17568553	FASTmrMLM	7	5.5	0.00	3.96	4.99E-07			
ARD	Chr13:17670145	pLARmEB	3	8.3	0.01	5.94	5.84E-10	Lus10030723	AT1G76420	NAC368
ARD	Chr13:17935031	FASTmrMLM	7	5.5	0.06	5.94	4.98E-07			
SDWt	Chr13:18631902	FASTmrMLM	3	3.6	0.01	5.94	4.65E-05	Lus10030921	AT1G60860	AGD2
								Lus10030567	AT4G21200	GA2OX8
RDWt	Chr14:995061	pLARmEB	7	24.6	0.14	8.91	Inf			
ARD	Chr14:1596540	pLARmEB	3	10.0	0.10	6.93	1.08E-11	Lus10028713	AT1G34190	NAC017
ARD	Chr14:2355271	pLARmEB	3	5.9	0.01	5.94	2.09E-07	Lus10013393	AT3G43600	AO3
								Lus10013387	AT1G21240	WAK3
								Lus10013383	AT1G21270	WAK2
ARD	Chr14:3694337	pLARmEB	3	13.4	0.01	6.93	4.11E-15	Lus10021401	AT1G08190	VAM2,ZIP2
SDWt	Chr14:3695509	FASTmrMLM	3	4.5	0.02	9.9	5.81E-06	Lus10021410	AT2G28190	CSD2
SDWt	Chr14:3842467	FASTmrMLM	7	9.7	0.00	1.98	2.11E-11			
RDWt	Chr14:4123574	pKWmEB	3	4.7	2.71	8.08	3.21E-06	Lus10021466	AT1G08010	GATA11
RDWt	Chr14:4123574	pLARmEB	7	36.3	0.36	7.92	Inf	Lus10021467	AT4G30080	ARF16
RDWt	Chr14:4767918	pLARmEB	7	9.9	0.15	4.95	1.52E-11			
SDWt	Chr14:8815881	FASTmrMLM	7	4.2	0.00	4.95	1.19E-05			
NWL	Chr14:13363192	mrMLM	3	4.6	9.77	9.09	4.55E-06	Lus10032911	AT3G28860	MDR1,
NWPer	Chr14:13363192	mrMLM	3	3.7	9.05	9.09	3.99E-05	Lus10032912	AT5G24660	LSU2
NWL	Chr14:13363192	pLARmEB	3	3.2	5.33	8.91	1.39E-04			
NWL	Chr14:13363192	ISIS EM-BLASSO	3	3.1	4.75	8.91	1.66E-04			
NWSA	Chr14:13363192	mrMLM	7	3.2	8.87	9.09	1.11E-04			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
NWA	Chr14:13363192	mrMLM	7	3.9	8.44	9.09	2.52E-05			
NWL	Chr14:13363192	mrMLM	7	3.8	10.69	9.09	2.83E-05			
NWPer	Chr14:13363192	mrMLM	7	3.5	9.86	9.09	5.94E-05			
RDWt	Chr14:15018214	pLARmEB	7	3.4	0.01	4.95	8.25E-05			
NWW_Dep	Chr14:15462441	ISIS EM-BLASSO	7	5.0	14.77	2.97	1.80E-06			
RDWt	Chr14:15462441	pLARmEB	7	6.5	0.05	2.97	4.59E-08			
SDWt	Chr14:15952502	FASTmrMLM	7	4.2	0.00	4.95	1.12E-05			
RDWt	Chr14:15956346	pLARmEB	7	15.6	0.06	4.95	Inf			
SDWt	Chr14:16353008	FASTmrMLM	7	18.6	0.00	5.94	Inf			
RDWt	Chr14:16353008	pLARmEB	7	8.1	0.12	5.94	1.02E-09			
ARD	Chr14:16745088	pLARmEB	3	14.4	0.02	9.9	3.33E-16	Lus10017988	AT1G48380	HYP7
								Lus10017984	AT1G73370	SUS6
ARD	Chr14:16927209	pLARmEB	3	8.7	0.06	5.94	2.66E-10	Lus10005537	AT5G53950	CUC2
SDWt	Chr14:16927209	FASTmrMLM	7	9.6	0.00	5.94	3.17E-11			
NWV	Chr14:18171993	LFMM	3				5.03E-05	Lus10038981	AT2G27030	CAM5
SDWt	Chr14:18171993	FASTmrMLM	7	3.4	0.25	10.89	7.63E-05	Lus10038977	AT1G77760	NR1
NWW	Chr14:18171993	LFMM	7				4.00E-05			
RDWt	Chr15:41099	pLARmEB	7	5.6	0.03	14.85	4.17E-07			
ARD	Chr15:2735924	pLARmEB	3	25.5	0.20	7.92	Inf	Lus10029410	AT5G22380	NAC090
ARD	Chr15:2735924	FASTmrMLM	3	9.1	0.15	7.92	9.13E-11			
RDWt	Chr15:2885976	pLARmEB	7	73.6	0.03	5.94	Inf			
ARD	Chr15:3479585	FASTmrMLM	7	7.6	0.03	3.96	3.43E-09			
RDWt	Chr15:3849840	pLARmEB	7	17.4	0.01	4.95	Inf			
SDWt	Chr15:9172034	FASTmrMLM	7	43.6	0.07	4.95	Inf			
ARD	Chr15:9340878	pLARmEB	3	6.0	0.01	5.94	1.31E-07			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
RDWt	Chr15:9525891	pLARmEB	7	7.5	0.07	12.87	4.66E-09			
ARD	Chr15:9795079	pLARmEB	3	34.5	0.14	14.85	Inf	Lus10041216	AT5G19280	RAG1
ARD	Chr15:10515587	pLARmEB	3	22.1	0.23	9.9	Inf	Lus10041396	AT3G61050	NTMC2T4
MedR	Chr15:10531332	mrMLM	3	3.8	8.52	6.06	2.65E-05			
NWSA	Chr15:11371216	mrMLM	3	3.6	13.23	6.06	5.03E-05	Lus10001139	AT5G63090	LOB
NWA	Chr15:11371216	mrMLM	3	4.1	13.05	6.06	1.48E-05	Lus10001137	AT5G40440	MKK3
NWW	Chr15:11371216	mrMLM	3	4.6	11.76	6.06	4.68E-06			
NWPer	Chr15:11371216	mrMLM	3	3.6	10.37	6.06	5.16E-05			
NWL	Chr15:11371216	mrMLM	3	3.0	10.21	6.06	1.79E-04			
NWA	Chr15:11371216	pLARmEB	3	3.8	7.15	5.94	2.58E-05			
NWV	Chr15:11371216	FASTmrMLM	3	3.0	6.64	5.94	1.85E-04			
NWV	Chr15:11371216	ISIS EM-BLASSO	3	3.7	6.62	5.94	3.54E-05			
NWV	Chr15:11371216	pLARmEB	3	3.8	6.61	5.94	2.56E-05			
NWA	Chr15:11371216	ISIS EM-BLASSO	3	3.1	6.50	5.94	1.77E-04			
NWSA	Chr15:11371216	pLARmEB	3	3.6	6.50	5.94	5.03E-05			
NWW	Chr15:11371216	FASTmrMLM	3	3.0	5.94	5.94	1.99E-04			
NWL	Chr15:11371216	pLARmEB	3	3.2	5.46	5.94	1.21E-04			
NWL	Chr15:11371216	ISIS EM-BLASSO	3	3.3	5.36	5.94	1.04E-04			
NWA	Chr15:11371216	FASTmrMLM	3	3.1	0.00	5.94	1.46E-04			
NWV	Chr15:11371216	LFMM	3				6.97E-05			
NWSA	Chr15:11371216	FASTmrMLM	7	3.1	7.28	5.94	1.46E-04			
NWW	Chr15:11371216	FASTmrMLM	7	4.2	9.52	5.94	1.23E-05			
NWSA	Chr15:11371216	ISIS EM-BLASSO	7	3.5	7.92	5.94	5.39E-05			
NWA	Chr15:11371216	ISIS EM-BLASSO	7	3.4	8.51	5.94	7.70E-05			
NWL	Chr15:11371216	ISIS EM-BLASSO	7	3.3	6.61	5.94	1.06E-04			

Trait	QTN[1]	Model	Dataset[2]	LOD[3]	R^2 (%)[4]	MAF[5]	*P* value	Predicted gene	At orthologue[6]	Gene product
NWPer	Chr15:11371216	ISIS EM-BLASSO	7	3.6	7.47	5.94	5.23E-05			
NWW	Chr15:11371216	LFMM	7				3.02E-05			
NWSA	Chr15:11371216	mrMLM	7	4.1	12.67	6.06	1.27E-05			
NWA	Chr15:11371216	mrMLM	7	4.4	13.67	6.06	7.20E-06			
NWW	Chr15:11371216	mrMLM	7	4.4	14.17	6.06	7.60E-06			
NWA	Chr15:11371216	pLARmEB	7	4.1	9.02	5.94	1.24E-05			
NWL	Chr15:11371216	pLARmEB	7	3.3	6.53	5.94	1.06E-04			
NWPer	Chr15:11371216	pLARmEB	7	3.3	6.87	5.94	9.27E-05			
RDWt	Chr15:13192686	pLARmEB	7	9.5	0.17	5.94	4.09E-11			
ARD	Chr15:14639205	pLARmEB	3	28.9	0.11	5.94	Inf	Lus10014822	AT3G54220	SCR,SGR1
								Lus10014821	AT1G71890	SUC5

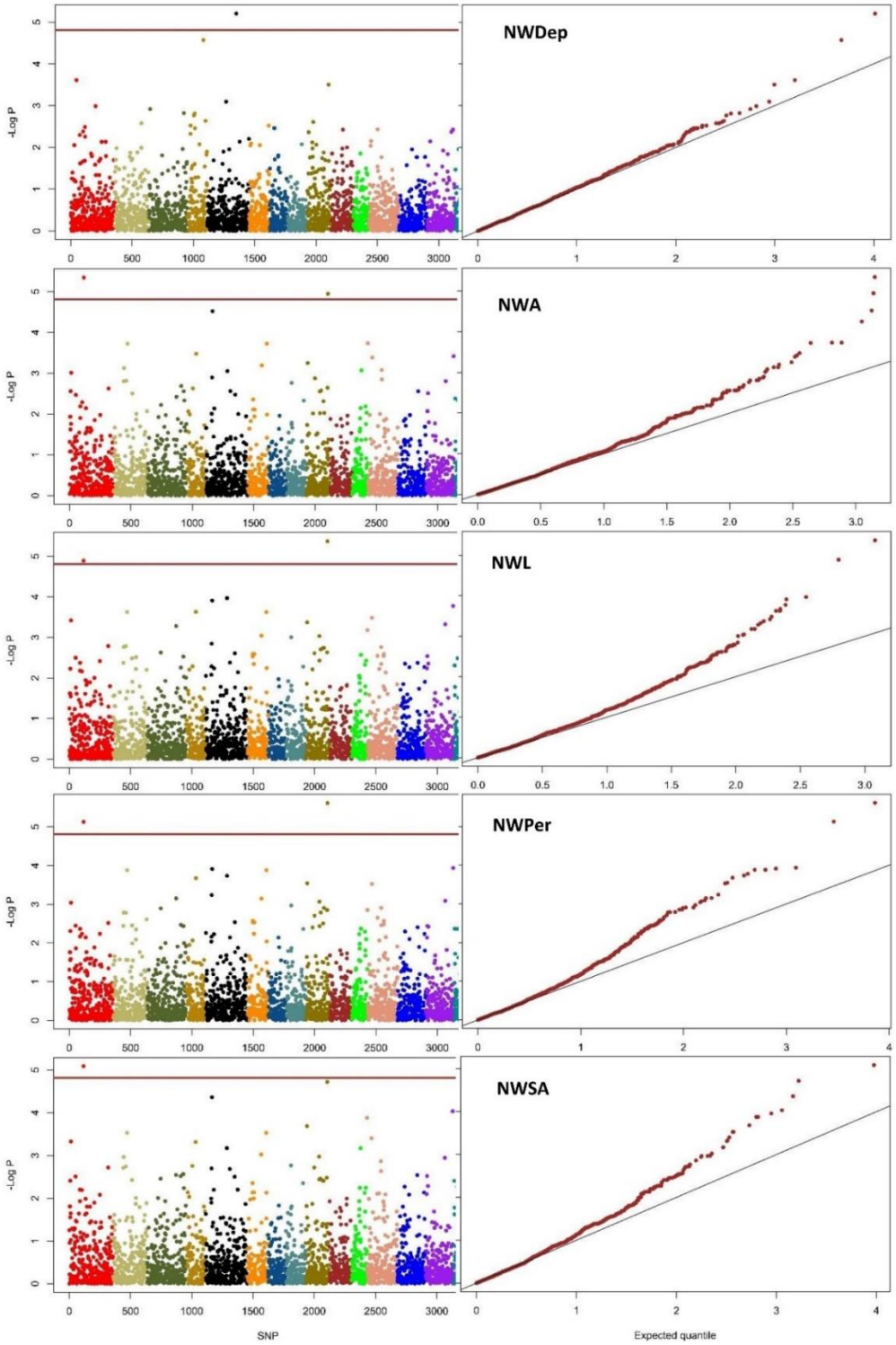
[1]Quantitative trait nucleotide chromosome number and position are indicated
[2] Dataset 3=3K and 7=7K
[3]LOD=logarithm of odds
[4]R^2=coefficient of determination indicating phenotypic variance explained due to allelic effect
[5]MAF=minor allele frequency
[6]*Arabidobsis thaliana* orthologue for the predicted gene in flax linked to the trait indicated within 100kb up and downstream of the detected QTN.

II-4 Table II-4 Associated traits and QTNs detected based on Bonferroni criterion from the two datasets

Trait	QTN[1]	Dataset[2]	*P* value	Predicted flax gene	Arabidopsis orthologue	Gene product
NWW_Dep	Chr1:7339983	7	5.39E-07	Lus10029174	AT3G53710	AGD6(ARF_GAP domain6)
NWL	Chr1:9614041	3	1.29E-05	Lus10004206	AT3G44350	NAC061
NWPer	Chr1:9614041	3	7.43E-06	Lus10004214	AT3G11620	ABH
NWSA	Chr1:9614041	3	8.16E-06	Lus10004225	AT3G11570	TBL8
NWV	Chr1:9614041	3	8.57E-06			
NWW	Chr1:9614041	3	1.52E-05			
NWW	Chr1:9614041	7	2.29E-07			
RDWt	Chr1:9614041	7	8.91E-06			
NWW	Chr2:6662368	7	1.88E-06			
NWW_Dep	Chr4:11669191	7	4.62E-06			
RDWt	Chr5:2645658	7	4.83E-06			
NWDep	Chr5:15312783	3	6.29E-06			
NWW_Dep	Chr6:6417737	3	1.28E-05	Lus10017778	AT5G08020	RPA70B
NWW_Dep	Chr6:6427847	3	1.28E-05			
NWW	Chr6:12030481	7	1.88E-06			
NWW	Chr6:16929211	7	2.68E-06			
NWL	Chr9:19061342	3	4.27E-06	Lus10024833	AT2G38290	AMT2,AMT2;1
NWPer	Chr9:19061342	3	2.43E-06	Lus10024853	AT1G04240	IAA3,SHY2
NWW	Chr9:19061342	3	1.28E-05			
NWPer	Chr9:19061342	7	5.24E-06			
NWW	Chr9:19061342	7	3.76E-07			
NWW_Dep	Chr9:19061342	7	1.62E-06			
NWW_Dep	Chr11:5124099	7	1.13E-07			
SDWt	Chr11:8154007	3	7.74E-06	Lus10036374	AT4G00350	MATE efflux
SDWt	Chr11:8176642	3	2.24E-06	Lus10036372	AT1G63800	UBC5
SDWt	Chr11:8176642	7	1.5E-06			

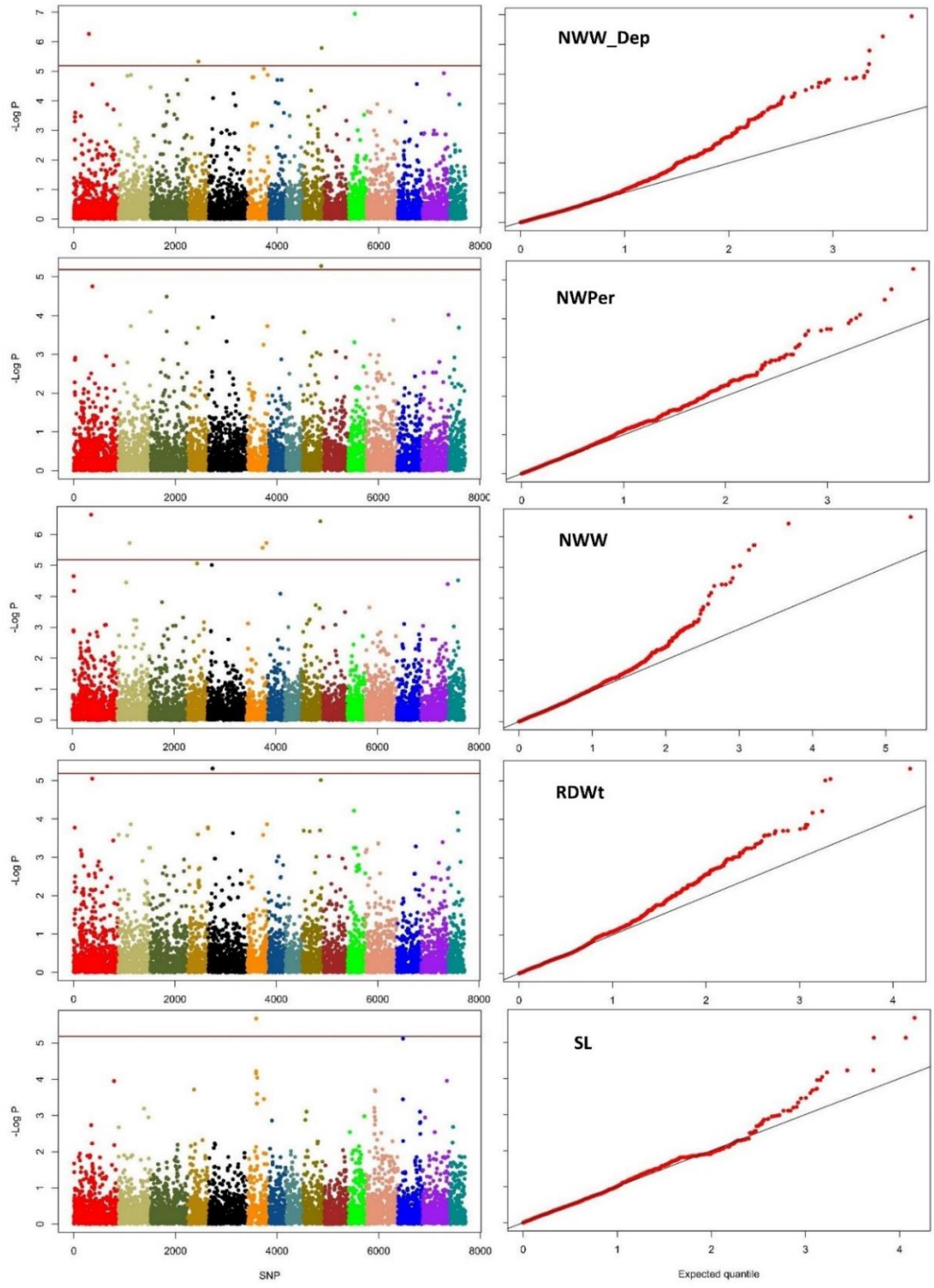
[1]Quantitative trait nucleotide; chromosome number and position are indicated

[2] Dataset 3=3K and 7=7K



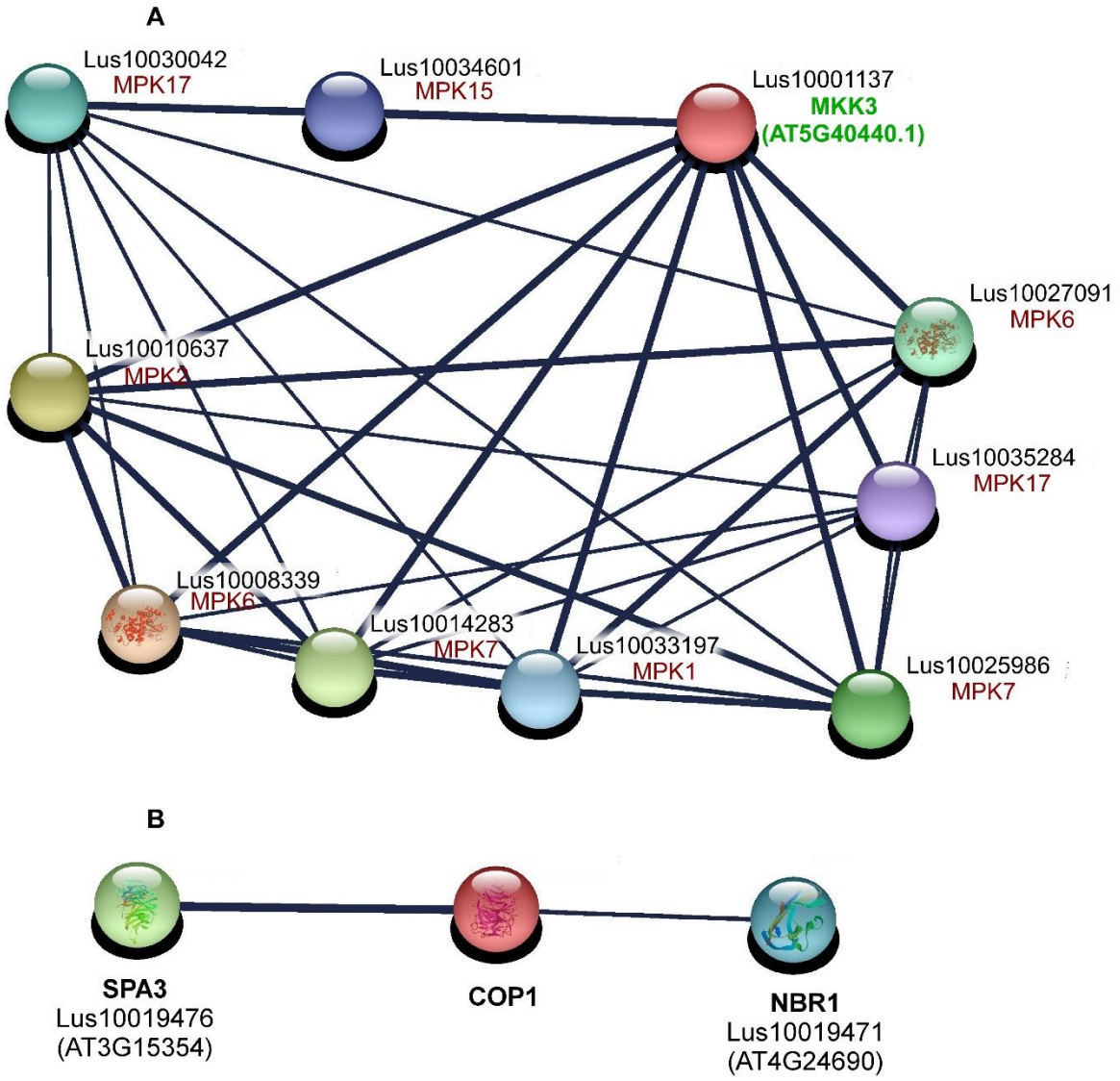
II-5 Fig. II-1 Manhattan and QQ plots showing QTN with significantly high peak based on Bonferroni criterion in 3K dataset.

The horizontal brown line indicate the threshold (P=0.05/n=0.05/3243= 1.54178E-05). Each color in Manhattan indicate a chromosome where chromosomes order from 1 to 15 from left to right.



II-6 Fig. II-2 Manhattan and QQ plots showing QTN with significantly high peak based on Bonferroni criterion in 7K dataset.

The horizontal brown line indicate the threshold (P=0.05/n=0.05/7707= 6.48761E-06). Each color in Manhattan indicate a chromosome where chromosomes order from 1 to 15 from left to right.



II-7 Fig. II-3 Protein interaction network of candidate genes at root and shoot QTN loci.

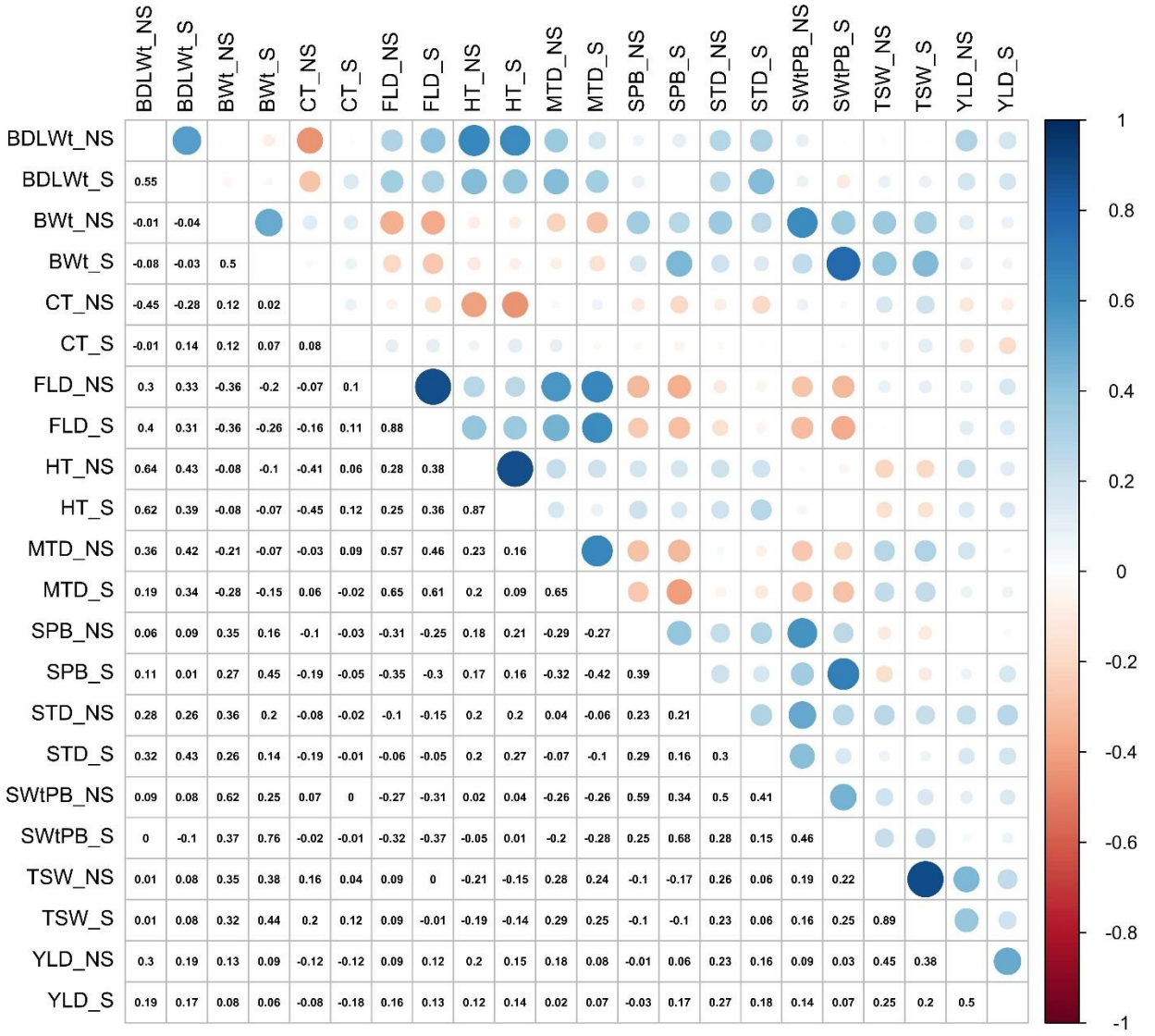
(**A**) Interaction of MAPK gene at QTN locus Chr15:11371866 with other MAPK genes of the flax genome. MKK3 is the MAPK linked to the target locus. The names in brown font are MAPK genes located elsewhere in the flax genome. (**B**) SPA3 and UBA (NBR1) at QTN locus Chr6:3310382 are predicted to interact via COP1. The names with Lus prefix are gene codes in flax genome (You et al 2018), names in parenthesis are Arabidopsis orthologues. The thickness of the lines indicate the strength of the predicted interactions.

III Supplementary materials for chapter 4

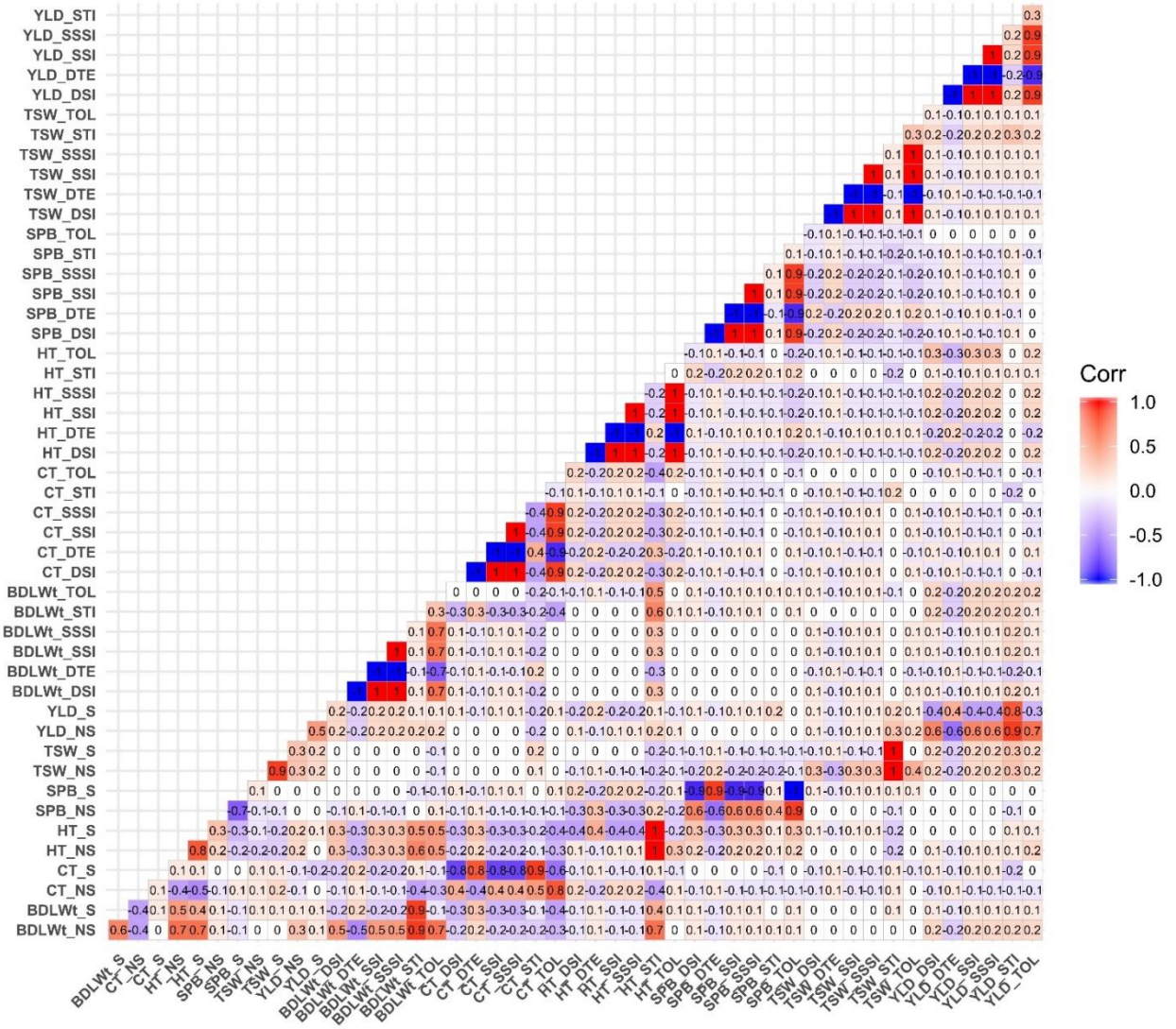
III-1 Table III-1 Cropping months weather data during experimental years 2016-2018

	Morden			Ottawa		
Year Month	PRECTOT[1]	MRH[2]	MT[3]	PRECTOT[1]	MRH[2]	MT[3]
2016	**559.52**	**65.7464**	**13.81748**	**462.24**	**69.57285**	**14.90234**
April	56.44	64.144	3.149667	47.54	69.06267	2.408667
May	91.49	56.75548	14.52323	35.33	72.20161	12.87742
June	86.96	63.31	18.389	64.14	68.85267	18.04533
July	118.41	68.0729	20.41645	60.02	66.85548	21.70484
August	70.01	62.97226	19.61065	90.84	67.86581	22.46032
September	88.18	68.68233	14.15567	54.03	66.11167	17.446
October	48.03	76.25226	6.291935	110.34	75.90871	9.154194
2017	**311.12**	**59.52262**	**14.05916**	**868.61**	**79.71776**	**14.74435**
April	15.63	59.33833	5.680333	120.97	78.731	6.970333
May	39.95	58.93871	12.37903	162.16	80.39516	11.70774
June	72.56	60.00933	18.06633	135.86	80.59	17.45167
July	46.09	57.10065	21.81258	170.81	80.46065	19.71258
August	39.34	54.74903	19.4171	78.74	78.57419	18.58387
September	85.88	64.681	14.23767	56.43	78.269	16.976
October	11.67	62.01742	6.685806	143.64	80.95387	11.71677
2018	**335.19**	**58.24626**	**13.67435**	**546.18**	**76.84682**	**14.41953**
April	6.26	51.85533	2.116333	104.34	80.18333	1.812667
May	50.08	52.77839	16.23968	55.87	75.09516	14.23548
June	95.12	59.14933	20.34867	83.92	76.652	17.29333
July	54.81	54.19935	21.78419	112.98	75.85226	21.94226
August	32.78	50.81871	20.83258	72.06	78.59548	21.40774
September	49.1	63.603	11.62733	55.73	74.733	17.17967
October	47.04	75.31548	2.548065	61.28	76.84968	6.840645

[1]Total precipitation; [2]Mean relative humidity; [3]Mean temperature



III-2 Fig. III-1 Spearman correlation of all traits under the irrigated (NS) and non-irrigated (S) conditions.



III-3 Fig. III-2 Pearson correlation of the six respondent traits under the two watering condition and their corresponding stress indices.

III-4 Table III-2 Traits and their associated QTN from all methods.

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
CT_TOL	FASTmrMLM	Chr1:22417	4.56	4.50E-06	5.1	11.3			
BDLWt_DSI	MLM	Chr1:1129073		1.80E-05	17.1	9.4	Lus10035975	AT2G17270	PHT3;3
BDLWt_DTE	MLM	Chr1:1129073		1.80E-05	17.1	9.4	Lus10035971	AT1G68150	WRKY9
BDLWt_SSI	MLM	Chr1:1129073		1.80E-05	17.1	9.4	Lus10035970	AT2G38290	AMT2
BDLWt_SSSI	MLM	Chr1:1129073		1.80E-05	17.1	9.4	Lus10025692	AT2G38290	AMT2
BDLWt_STI	pLARmEB	Chr1:2649874	4.15	1.20E-05	2.2	6.6	Lus10042505	AT1G53440	LRR
BDLWt_S	FASTmrMLM	Chr1:2649874	3.46	6.60E-05	5.5	6.6	Lus10025941	AT3G24715	HCR1
BDLWt_S	ISIS EM-BLASSO	Chr1:2649874	3.55	5.20E-05	6.2	6.6			
HT_NS	FASTmrEMMA	Chr1:3082807	3.28	1.00E-04	7.0	21.7	Lus10042401	AT2G02480	STI
CT_STI	FASTmrMLM	Chr1:3237594	3.24	1.10E-04	6.6	5.7			
CT_STI	ISIS EM-BLASSO	Chr1:3237594	3.05	1.80E-04	4.3	5.7			
TSW_S	FASTmrEMMA	Chr1:3887702	3.78	3.00E-05	7.2	23.6	Lus10008595	AT3G49400	WD40R
TSW_STI	FASTmrEMMA	Chr1:3887702	3.22	1.20E-04	6.3	23.6	Lus10008599	AT5G60870	RCC1
						0.0	Lus10008607	AT1G53600	TPR
TSW_DSI	mrMLM	Chr1:4520461	4.34	7.90E-06	16.6	15.4	Lus10027782	AT2G38080	LMCO4
TSW_SSI	mrMLM	Chr1:4520461	4.34	7.90E-06	16.6	15.4	Lus10027786	AT3G08940	LHCB4
TSW_SSSI	mrMLM	Chr1:4520461	4.34	7.90E-06	16.6	15.4	Lus10027787	AT3G51880	HMGB1
						0.0	Lus10027788	AT5G01510	RUS5
CT_TOL	ISIS EM-BLASSO	Chr1:5713605	3.10	1.60E-04	5.6	10.4	Lus10012574	AT5G14040	PHT3;1
						0.0	Lus10012579	AT2G38130	MAK3
						0.0	Lus10012582	AT4G36220	FAH1
TSW_NS	mrMLM	Chr1:7029139	3.56	5.20E-05	18.2	15.4	Lus10029120	AT1G05190	emb2394
TSW_S	ISIS EM-BLASSO	Chr1:7029139	4.04	1.60E-05	12.7	15.1	Lus10029115	AT1G60770	TPR
TSW_STI	mrMLM	Chr1:7029139	4.69	3.40E-06	20.2	15.4			
TSW_STI	FASTmrEMMA	Chr1:7029139	4.40	6.70E-06	13.9	15.1			
TSW_STI	ISIS EM-BLASSO	Chr1:7029139	6.51	4.40E-08	19.0	15.1			
TSW_DSI	mrMLM	Chr1:9230907	4.44	6.10E-06	6.6	13.5	Lus10016200	AT2G38910	CPK20

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
TSW_SSI	mrMLM	Chr1:9230907	4.44	6.10E-06	6.6	13.5	Lus10016191	AT4G35310	CPK5
TSW_SSSI	mrMLM	Chr1:9230907	4.44	6.10E-06	6.6	13.5	Lus10016190	AT3G10660	CPK2
						0.0	Lus10016196	AT5G20720	CPN10/20/21
NHT_NS	ISIS EM-BLASSO	Chr1:9371618	4.14	1.30E-05	9.1	49.1	Lus10016166	AT3G52450	PUB22
						0.0	Lus10004191	AT3G11840	PUB24
BDLWt_DSI	MLM	Chr1:12482542		6.10E-05	14.8	9.4	Lus10020589	AT2G27600	SKD1
BDLWt_DTE	MLM	Chr1:12482542		6.10E-05	14.8	9.4	Lus10020590	AT4G14770	TCX2
BDLWt_SSI	MLM	Chr1:12482542		6.10E-05	14.8	9.4			
BDLWt_SSSI	MLM	Chr1:12482542		6.10E-05	14.8	9.4			
BDLWt_DSI	MLM	Chr1:12657010		6.10E-05	14.8	9.4	Lus10020610	AT1G08320	TGA9
BDLWt_DTE	MLM	Chr1:12657010		6.10E-05	14.8	9.4	Lus10020612	AT3G25520	RPL5A
BDLWt_SSI	MLM	Chr1:12657010		6.10E-05	14.8	9.4			
BDLWt_SSSI	MLM	Chr1:12657010		6.10E-05	14.8	9.4			
HT_DSI	FASTmrMLM	Chr1:17731344	3.02	1.90E-04	9.9	5.7	Lus10026819	AT1G46264	HSFB4
HT_DSI	FASTmrEMMA	Chr1:17731344	3.08	1.70E-04	0.0	5.7	Lus10000889	AT5G12180	CPK17
HT_DSI	pLARmEB	Chr1:17731344	3.24	1.10E-04	10.1	5.7			
HT_DSI	ISIS EM-BLASSO	Chr1:17731344	3.75	3.20E-05	10.7	5.7			
HT_DTE	FASTmrMLM	Chr1:17731344	3.87	2.40E-05	10.6	5.7			
HT_DTE	pLARmEB	Chr1:17731344	3.95	2.00E-05	10.0	5.7			
HT_DTE	ISIS EM-BLASSO	Chr1:17731344	4.27	9.20E-06	10.8	5.7			
HT_SSI	FASTmrMLM	Chr1:17731344	3.24	1.10E-04	10.6	5.7			
HT_SSI	FASTmrEMMA	Chr1:17731344	3.08	1.70E-04	0.0	5.7			
HT_SSI	ISIS EM-BLASSO	Chr1:17731344	3.75	3.20E-05	10.8	5.7			
HT_SSSI	FASTmrMLM	Chr1:17731344	3.02	1.90E-04	9.9	5.7			
HT_SSSI	FASTmrEMMA	Chr1:17731344	3.08	1.70E-04	0.0	5.7			
HT_SSSI	ISIS EM-BLASSO	Chr1:17731344	3.75	3.20E-05	10.7	5.7			
CT_NS	ISIS EM-BLASSO	Chr1:18066829	3.51	5.90E-05	6.1	8.5	Lus10009701	AT5G19420	RCC1
SPB_S	FASTmrMLM	Chr1:20699101	3.17	1.30E-04	7.0	13.2	Lus10015848	AT3G16857	RR1

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
SPB_S	ISIS EM-BLASSO	Chr1:20699101	3.17	1.30E-04	7.0	13.2	Lus10015853	AT2G45670	PEAT2
						0.0	Lus10015862	AT3G60870	AHL18
BDLWt_STI	FASTmrMLM	Chr1:20932841	3.77	3.10E-05	5.1	5.7	Lus10015861	AT2G45440	DHDPS2
BDLWt_STI	pLARmEB	Chr1:20932841	3.37	8.20E-05	1.8	5.7			
BDLWt_STI	ISIS EM-BLASSO	Chr1:20932841	3.21	1.20E-04	4.1	5.7			
BDLWt_DSI	MLM	Chr1:21002257		6.60E-05	14.6	8.5	Lus10020434	AT3G18270	CYP77A5P
BDLWt_DTE	MLM	Chr1:21002257		6.60E-05	14.6	8.5	Lus10020435	AT4G33630	EX1
BDLWt_SSI	MLM	Chr1:21002257		6.60E-05	14.6	8.5	Lus10020436	AT5G64370	PYD3
BDLWt_SSSI	MLM	Chr1:21002257		6.60E-05	14.6	8.5	Lus10020437	AT4G37370	CYP81D8
						0.0	Lus10020439	AT1G29510	SAUR68
CT_NS	ISIS EM-BLASSO	Chr1:25733811	3.75	3.20E-05	6.2	12.3	Lus10003114	AT5G39150	LRR-RLK
						0.0	Lus10003116	AT5G39150	LRR-RLK
YLD_DSI	mrMLM	Chr1:28692299	3.05	1.80E-04	15.3	6.7	Lus10006062	AT1G14290	SBH2
YLD_SSI	mrMLM	Chr1:28692299	3.05	1.80E-04	15.3	6.7	Lus10006065	AT1G35670	CDPK2
YLD_DTE	mrMLM	Chr1:28692299	3.05	1.80E-04	15.3	6.7	Lus10006067	AT1G34420	LRRK
YLD_SSSI	mrMLM	Chr1:28692299	3.05	1.80E-04	15.3	6.7	Lus10006058	AT1G14290	SBH2
SPB_S	mrMLM	Chr2:3211248	3.98	1.90E-05	9.1	8.7			
NHT_NS	FASTmrMLM	Chr2:3816576	3.90	2.20E-05	6.7	8.5	Lus10020105	AT5G24930	COL4
NHT_NS	pLARmEB	Chr2:3816576	3.90	2.20E-05	6.7	8.5	Lus10020107	AT4G31730	GDU1
HT_STI	FASTmrMLM	Chr2:3816576	3.17	1.30E-04	3.1	8.5	Lus10020120	AT4G31590	CSLC5
HT_STI	pLARmEB	Chr2:3816576	4.43	6.20E-06	1.9	8.5	Lus10020121	AT4G07960	CSLC12
HT_STI	ISIS EM-BLASSO	Chr2:3816576	4.87	2.20E-06	4.8	8.5	Lus10020116	AT2G24640	UBP19
SPB_DTE	FASTmrMLM	Chr2:4509381	3.63	4.40E-05	8.9	7.6	Lus10013611	AT3G59480	FRK7
SPB_DSI	FASTmrMLM	Chr2:4509381	3.63	4.40E-05	8.9	7.6	Lus10020233	AT5G16270	SCC1/REC8
SPB_SSI	FASTmrMLM	Chr2:4509381	3.63	4.40E-05	8.9	7.6	Lus10020116	AT2G24640	UBP19
SPB_SSI	pLARmEB	Chr2:4509381	4.11	1.30E-05	0.7	7.6	Lus10013615	AT5G18100	CSD3
SPB_SSSI	FASTmrMLM	Chr2:4509381	3.63	4.40E-05	8.9	7.6	Lus10020230	AT5G16260	ELF9
SPB_DTE	mrMLM	Chr2:5592445	4.38	7.00E-06	15.8	10.6	Lus10038088	AT3G07490	AGD11

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
SPB_DSI	mrMLM	Chr2:5592445	4.38	7.00E-06	15.8	10.6	Lus10038086	AT4G33150	LKR/SDH
SPB_SSI	mrMLM	Chr2:5592445	4.38	7.00E-06	15.8	10.6			
SPB_SSSI	mrMLM	Chr2:5592445	4.38	7.00E-06	15.8	10.6			
TSW_DTE	mrMLM	Chr2:16676745	3.16	1.30E-04	15.8	11.5	Lus10016289	AT2G26490	JGB
						0.0	Lus10016298	AT1G19270	DA1
NYLD_NS	ISIS EM-BLASSO	Chr2:21621421	5.33	7.20E-07	15.8	7.6	Lus10035138	AT1G68400	LRR
						0.0	Lus10035141	AT5G15680	ARMR
						0.0	Lus10035142	AT5G15680	ARMR
						0.0	Lus10035143	AT1G68260	ALT3
						0.0	Lus10035152	AT1G68260	ALT4
CT_DSI	FASTmrMLM	Chr2:23123754	5.17	1.10E-06	16.7	9.4	Lus10013252	AT1G34300	LPK
CT_DSI	pLARmEB	Chr2:23123754	5.42	5.80E-07	6.0	9.4	Lus10013244	AT1G11340	S-LLPK
CT_DSI	ISIS EM-BLASSO	Chr2:23123754	5.42	5.80E-07	16.7	9.4	Lus10013250	AT2G01830	WOL1
CT_SSI	FASTmrMLM	Chr2:23123754	5.17	1.10E-06	16.7	9.4	Lus10013240	AT2G01850	XTH27
CT_SSI	pLARmEB	Chr2:23123754	3.33	8.90E-05	7.5	9.4			
CT_SSI	ISIS EM-BLASSO	Chr2:23123754	5.42	5.80E-07	16.7	9.4			
CTR	FASTmrMLM	Chr2:23123754	5.17	1.10E-06	16.7	9.4			
CTR	ISIS EM-BLASSO	Chr2:23123754	5.17	1.10E-06	16.7	9.4			
CT_TOL	pLARmEB	Chr2:23123754	3.35	8.50E-05	6.0	9.4			
CT_TOL	ISIS EM-BLASSO	Chr2:23123754	3.68	3.80E-05	8.4	9.4			
CT_SSSI	FASTmrMLM	Chr2:23123754	5.17	1.10E-06	16.7	9.4			
CT_SSSI	pLARmEB	Chr2:23123754	5.28	8.10E-07	7.5	9.4			
CT_SSSI	ISIS EM-BLASSO	Chr2:23123754	5.42	5.80E-07	16.7	9.4			
SPB_SSI	pLARmEB	Chr2:24765844	3.10	1.60E-04	0.2	11.3			
YLD_DSI	mrMLM	Chr2:24999711	4.72	3.10E-06	10.0	8.7	Lus10004620	AT3G26600	ARO4
YLD_SSI	mrMLM	Chr2:24999711	4.72	3.10E-06	10.0	8.7	Lus10004630	AT1G69010	BIM2
YLD_DTE	mrMLM	Chr2:24999711	4.72	3.10E-06	10.0	8.7			
YLD_SSSI	mrMLM	Chr2:24999711	4.72	3.10E-06	10.0	8.7			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
BDLWt_DSI	MLM	Chr2:3045789		7.40E-05	14.4	8.5			
BDLWt_DTE	MLM	Chr2:3045789		7.40E-05	14.4	8.5	Lus10014342	AT3G10490	NAC052
BDLWt_SSI	MLM	Chr2:3045789		7.40E-05	14.4	8.5			
BDLWt_SSSI	MLM	Chr2:3045789		7.40E-05	14.4	8.5			
BDLWt_DSI	MLM	Chr2:3045853		7.40E-05	14.4	8.5			
BDLWt_DTE	MLM	Chr2:3045853		7.40E-05	14.4	8.5			
BDLWt_SSI	MLM	Chr2:3045853		7.40E-05	14.4	8.5			
BDLWt_SSSI	MLM	Chr2:3045853		7.40E-05	14.4	8.5			
BDLWt_DSI	MLM	Chr2:7199506		2.30E-05	16.6	10.4	Lus10038651	AT3G52190	PHF1
BDLWt_DTE	MLM	Chr2:7199506		2.30E-05	16.6	10.4			
BDLWt_SSI	MLM	Chr2:7199506		2.30E-05	16.6	10.4			
BDLWt_SSSI	MLM	Chr2:7199506		2.30E-05	16.6	10.4			
YLD_NS	mrMLM	Chr3:1000904	4.06	1.50E-05	10.5	8.7	Lus10021946	AT3G06190	BPM2
SPB_DSI	pLARmEB	Chr3:3391512	3.08	1.70E-04	0.6	5.7	Lus10019206	AT2G04030	Hsp88/HSP90
SPB_SSSI	pLARmEB	Chr3:3391512	3.90	2.30E-05	0.8	5.7			
SPB_DTE	mrMLM	Chr3:3839306	3.62	4.40E-05	10.8	7.7	Lus10037038	AT5G65060	AGL70, FCL3, MAF3
SPB_DSI	mrMLM	Chr3:3839306	3.62	4.40E-05	10.8	7.7	Lus10037039	AT5G10140	AGL25,FLC,FLF
SPB_SSI	mrMLM	Chr3:3839306	3.62	4.40E-05	10.8	7.7	Lus10037040	AT1G24260	AGL9,SEP3)
SPB_SSSI	mrMLM	Chr3:3839306	3.62	4.40E-05	10.8	7.7			
HT_NS	mrMLM	Chr3:3890807	4.41	6.50E-06	9.4	20.2	Lus10037027	AT1G27950	LTPG1
HT_NS	FASTmrMLM	Chr3:3890807	4.22	1.00E-05	5.8	19.8	Lus10037032	AT1G27320	HK3
HT_NS	ISIS EM-BLASSO	Chr3:3890807	3.65	4.10E-05	5.4	19.8			
HT_STI	FASTmrMLM	Chr3:3890807	3.62	4.40E-05	3.4	19.8			
HT_STI	pLARmEB	Chr3:3890807	3.71	3.60E-05	1.9	19.8			
HT_STI	ISIS EM-BLASSO	Chr3:3890807	5.11	1.20E-06	5.3	19.8			
YLD_NS	mrMLM	Chr3:7068603	3.75	3.30E-05	10.0	11.5	Lus10040840	AT5G59340	WOX2
BDLWt_DSI	mrMLM	Chr3:7401595	4.21	1.10E-05	14.5	5.8	Lus10004612	AT1G29860	WRKY71

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
BDLWt_SSI	mrMLM	Chr3:7401595	4.21	1.10E-05	14.5	5.8	Lus10004611	AT4G18160	KCO6,TPK3
BDLWt_DTE	mrMLM	Chr3:7401595	4.21	1.10E-05	14.5	5.8			
BDLWt_SSSI	mrMLM	Chr3:7401595	4.21	1.10E-05	14.5	5.8			
HT_NS	FASTmrMLM	Chr3:8455048	4.52	5.10E-06	6.2	19.8	Lus10005270	AT1G09200	HISTONE 3.1
						0.0	Lus10005271	AT1G09200	HISTONE 3.1
CT_STI	ISIS EM-BLASSO	Chr3:9279281	4.20	1.10E-05	22.9	9.4	Lus10019365	AT2G25110	SDL,SDF2
CT_TOL	pLARmEB	Chr3:16885895	4.68	3.50E-06	7.3	6.6	Lus10020500	AT1G09970	RLK7
BDLWt_DSI	FASTmrMLM	Chr3:23320193	4.33	8.00E-06	8.1	8.5	Lus10033386	AT3G21180	ACA9
BDLWt_DSI	pLARmEB	Chr3:23320193	4.02	1.70E-05	1.8	8.5	Lus10033391/2	AT5G13750	ZIFL1
BDLWt_DSI	ISIS EM-BLASSO	Chr3:23320193	3.69	3.80E-05	7.0	8.5			
BDLWt_SSI	FASTmrMLM	Chr3:23320193	4.33	8.00E-06	8.1	8.5			
BDLWt_SSI	ISIS EM-BLASSO	Chr3:23320193	3.69	3.80E-05	7.0	8.5			
BDLWt_DTE	FASTmrMLM	Chr3:23320193	4.33	8.00E-06	8.1	8.5			
BDLWt_DTE	ISIS EM-BLASSO	Chr3:23320193	3.61	4.60E-05	7.0	8.5			
BDLWt_SSSI	FASTmrMLM	Chr3:23320193	4.33	8.00E-06	8.1	8.5			
BDLWt_SSSI	pLARmEB	Chr3:23320193	4.54	4.80E-06	1.6	8.5			
BDLWt_SSSI	ISIS EM-BLASSO	Chr3:23320193	3.69	3.80E-05	7.0	8.5			
BDLWt_DSI	MLM	Chr3:9262936		8.10E-05	14.2	12.3	Lus10019361	AT4G37650	SGR7,SHR
BDLWt_DTE	MLM	Chr3:9262936		8.10E-05	14.2	12.3	Lus10019365	AT2G25110	SDF2
BDLWt_SSI	MLM	Chr3:9262936		8.10E-05	14.2	12.3			
BDLWt_SSSI	MLM	Chr3:9262936		8.10E-05	14.2	12.3			
CT_STI	MLM	Chr3:9279281		7.20E-06	22.6	9.4			
BDLWt_DSI	MLM	Chr3:19216959		1.10E-05	18.1	10.4	Lus10033886	AT3G54700	PHT1;7
BDLWt_DTE	MLM	Chr3:19216959		1.10E-05	18.1	10.4			
BDLWt_SSI	MLM	Chr3:19216959		1.10E-05	18.1	10.4			
BDLWt_SSSI	MLM	Chr3:19216959		1.10E-05	18.1	10.4			
BDLWt_DSI	MLM	Chr3:24874158		1.10E-04	13.6	8.5	Lus10037694	AT3G02850	SKOR
BDLWt_DTE	MLM	Chr3:24874158		1.10E-04	13.6	8.5			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
BDLWt_SSI	MLM	Chr3:24874158		1.10E-04	13.6	8.5			
BDLWt_SSSI	MLM	Chr3:24874158		1.10E-04	13.6	8.5			
NCT_NS	FASTmrMLM	Chr4:9280669	3.73	3.40E-05	10.1	12.3	Lus10013044	AT1G10760	SOP1
NCT_NS	ISIS EM-BLASSO	Chr4:9280669	3.73	3.40E-05	10.1	12.3	Lus10013040	AT1G70530	CRK3
						0.0	Lus10013039	AT1G70520	CRK2
						0.0	Lus10013056	AT5G28540	Hsp 70 (BIP3)
						0.0	Lus10013055	AT1G09080	Hsp 70 (BIP3)
SPB_DTE	mrMLM	Chr4:13309980	4.83	2.40E-06	14.6	8.7	Lus10015717	AT5G61400	PPR
SPB_DTE	FASTmrMLM	Chr4:13309980	3.39	7.80E-05	6.6	8.5	Lus10015716	AT5G61400	PPR
SPB_DTE	ISIS EM-BLASSO	Chr4:13309980	5.53	4.50E-07	11.5	8.5	Lus10015715	AT5G61410	RPE (emb2728)
SPB_DSI	mrMLM	Chr4:13309980	4.83	2.40E-06	14.6	8.7			
SPB_DSI	FASTmrMLM	Chr4:13309980	3.39	7.80E-05	6.6	8.5			
SPB_DSI	FASTmrEMMA	Chr4:13309980	4.96	1.80E-06	0.0	8.5			
SPB_DSI	pLARmEB	Chr4:13309980	3.50	5.90E-05	0.5	8.5			
SPB_DSI	ISIS EM-BLASSO	Chr4:13309980	5.81	2.30E-07	11.5	8.5			
SPB_SSI	mrMLM	Chr4:13309980	4.83	2.40E-06	14.6	8.7			
SPB_SSI	FASTmrMLM	Chr4:13309980	3.39	7.80E-05	6.6	8.5			
SPB_SSI	FASTmrEMMA	Chr4:13309980	4.96	1.80E-06	0.0	8.5			
SPB_SSI	pLARmEB	Chr4:13309980	3.39	7.80E-05	0.4	8.5			
SPB_SSI	ISIS EM-BLASSO	Chr4:13309980	5.87	2.00E-07	11.5	8.5			
SPB_SSSI	mrMLM	Chr4:13309980	4.83	2.40E-06	14.6	8.7			
SPB_SSSI	FASTmrMLM	Chr4:13309980	3.39	7.80E-05	6.6	8.5			
SPB_SSSI	FASTmrEMMA	Chr4:13309980	4.96	1.80E-06	0.0	8.5			
SPB_SSSI	ISIS EM-BLASSO	Chr4:13309980	5.87	2.00E-07	11.5	8.5			
TSW_DSI	mrMLM	Chr4:14422439	3.84	2.60E-05	9.5	9.6	Lus10041483	AT1G25540	PFT1
TSW_SSI	mrMLM	Chr4:14422439	3.84	2.60E-05	9.5	9.6	Lus10041487	AT1G13260	EDF4,RAV1
TSW_SSSI	mrMLM	Chr4:14422439	3.84	2.60E-05	9.5	9.6	Lus10041492	AT1G25580	ANAC008,SOG1
SPB_TOL	pLARmEB	Chr4:15035662	3.23	1.10E-04	5.3	6.6	Lus10041613	AT3G12750	ZIP1

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
						0.0	Lus10041615	AT1G16030	Hsp70b
						0.0			
SPB_DSI	pLARmEB	Chr4:15585809	4.44	6.10E-06	0.3	6.6	Lus10041717	AT4G14910	HISN5B
SPB_SSI	pLARmEB	Chr4:15585809	5.18	1.00E-06	0.4	6.6			
BDLWt_DSI	MLM	Chr4:10207561		7.40E-06	18.9	11.3	Lus10007267	AT3G01280	VDAC1
BDLWt_DTE	MLM	Chr4:10207561		7.40E-06	18.9	11.3	Lus10007263	AT5G39820	NAC094
BDLWt_SSI	MLM	Chr4:10207561		7.40E-06	18.9	11.3			
BDLWt_SSSI	MLM	Chr4:10207561		7.40E-06	18.8	11.3			
BDLWt_DSI	MLM	Chr4:19400123		3.50E-05	15.8	6.6	Lus10039806	AT3G15730	PLDALPHA1
BDLWt_DTE	MLM	Chr4:19400123		3.50E-05	15.8	6.6			
BDLWt_SSI	MLM	Chr4:19400123		3.50E-05	15.8	6.6			
BDLWt_SSSI	MLM	Chr4:19400123		3.50E-05	15.8	6.6			
HT_STI	FASTmrMLM	Chr5:10294	4.66	3.60E-06	4.6	9.4			
HT_NS	mrMLM	Chr5:1375386	5.99	1.50E-07	15.9	8.7	Lus10029692	AT1G01720	NAC002,AF1
HT_NS	pLARmEB	Chr5:1375386	3.46	6.50E-05	5.2	8.5	Lus10029691	AT1G77460	CSI3
HT_NS	ISIS EM-BLASSO	Chr5:1375386	6.00	1.50E-07	13.3	8.5	Lus10029690	AT1G77460	CSI3
HT_STI	mrMLM	Chr5:1375386	5.87	2.00E-07	15.7	8.7			
HT_STI	pLARmEB	Chr5:1375386	3.79	2.90E-05	4.6	8.5			
HT_STI	ISIS EM-BLASSO	Chr5:1375386	6.33	6.70E-08	9.6	8.5			
HT_STI	ISIS EM-BLASSO	Chr5:1620515	6.90	1.70E-08	7.9	8.5	Lus10008499	AT4G11600	GPX6
HT_STI	FASTmrMLM	Chr5:1650789	3.68	3.80E-05	5.0	10.4			
HT_STI	pLARmEB	Chr5:1650789	5.05	1.40E-06	2.6	10.4			
SPB_S	mrMLM	Chr5:3935326	4.23	1.00E-05	9.7	27.9	Lus10008282	AT1G32930	GALT31A
						0.0	Lus10008285	AT4G35580	NTL9
NYLD_NS	FASTmrMLM	Chr5:4624604	3.37	8.20E-05	10.1	10.4	Lus10034793	AT5G59550	RING-DUF1117
						0.0	Lus10034794	AT5G13750	ZIFL1
						0.0	Lus10034797	AT3G20210	DELTAVPE
YLD_STI	ISIS EM-BLASSO	Chr5:4935595	3.71	3.50E-05	5.0	5.7	Lus10034840	AT3G21180	ACA9

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
BDLWt_STI	FASTmrMLM	Chr5:5628560	3.01	2.00E-04	5.4	10.4	Lus10034892	AT3G26060	ATPRX Q
TSW_TOL	mrMLM	Chr5:9161068	3.82	2.70E-05	7.5	7.7	Lus10007126	AT3G08690	ATUBC11,UBC11
						0.0	Lus10007121	AT1G04860	ATUBP2,UBP2
TSW_DSI	FASTmrMLM	Chr5:11604499	3.25	1.10E-04	5.9	18.9	Lus10014853	AT5G67360	ARA12
TSW_SSSI	FASTmrMLM	Chr5:11604499	3.25	1.10E-04	5.9	18.9			
HT_DSI	mrMLM	Chr5:13281348	4.04	1.60E-05	10.7	5.8	Lus10029853	AT1G09190	TPR
HT_DTE	mrMLM	Chr5:13281348	4.04	1.60E-05	10.7	5.8	Lus10029850	AT3G62550	
HT_SSI	mrMLM	Chr5:13281348	4.04	1.60E-05	10.7	5.8			
HT_SSSI	mrMLM	Chr5:13281348	4.04	1.60E-05	10.7	5.8			
SPB_DSI	FASTmrEMMA	Chr5:16385870	5.75	2.60E-07	0.0	9.4	Lus10039598	AT5G13690	CYL1,NAGLU
SPB_SSI	FASTmrEMMA	Chr5:16385870	5.75	2.60E-07	0.0	9.4	Lus10039597	AT5G52820	NLE(WD40)
SPB_SSSI	FASTmrEMMA	Chr5:16385870	5.75	2.60E-07	0.0	9.4			
TSW_DTE	mrMLM	Chr5:16422430	3.20	1.20E-04	8.8	14.4			
TSW_DSI	mrMLM	Chr6:3094197	3.56	5.20E-05	9.1	7.7			
TSW_SSI	mrMLM	Chr6:3094197	3.56	5.20E-05	9.1	7.7			
TSW_SSSI	mrMLM	Chr6:3094197	3.56	5.20E-05	9.1	7.7			
TSW_NS	FASTmrMLM	Chr6:3310382	5.02	1.50E-06	16.6	12.3	Lus10019476	AT3G15354	SPA3
TSW_NS	pLARmEB	Chr6:3310382	5.02	1.50E-06	15.3	12.3	Lus10019475	AT3G15354	SPA3
TSW_S	FASTmrMLM	Chr6:3310382	3.47	6.40E-05	9.8	12.3	Lus10019474	AT3G48250	BIR6
TSW_S	pLARmEB	Chr6:3310382	3.47	6.40E-05	9.8	12.3	Lus10019482	AT3G48430	REF6
TSW_STI	pLARmEB	Chr6:3310382	7.54	3.80E-09	8.8	12.3			
BDLWt_TOL	mrMLM	Chr6:4915106	3.26	1.10E-04	10.4	8.7			
BDLWt_STI	pLARmEB	Chr6:5531448	3.40	7.60E-05	0.0	12.3	Lus10036644	AT3G48090	EDS1
HT_STI	mrMLM	Chr6:13339308	3.61	4.60E-05	9.3	11.5	Lus10013823	AT2G36870	XTH32
HT_STI	FASTmrMLM	Chr6:13339308	5.73	2.80E-07	7.7	11.3	Lus10013822	AT2G36870	XTH32
HT_STI	pLARmEB	Chr6:13339308	5.83	2.20E-07	2.3	11.3	Lus10013836	AT5G44370	PHT4;6
HT_NS	mrMLM	Chr6:13365708	3.62	4.40E-05	6.0	7.7			
HT_NS	FASTmrMLM	Chr6:13365708	4.99	1.60E-06	6.0	7.6			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
HT_STI	ISIS EM-BLASSO	Chr6:13365708	4.38	7.10E-06	3.4	7.6			
BDLWt_DSI	MLM	Chr6:2475973		4.10E-06	20.0	10.4	Lus10043402	AT5G22290	NAC089
BDLWt_DTE	MLM	Chr6:2475973		4.10E-06	20.0	10.4			
BDLWt_SSI	MLM	Chr6:2475973		4.10E-06	20.0	10.4			
BDLWt_SSSI	MLM	Chr6:2475973		4.10E-06	20.0	10.4			
BDLWt_DSI	MLM	Chr6:14958574		4.50E-06	19.8	11.3	Lus10014415	AT3G53200	MYB27
BDLWt_DTE	MLM	Chr6:14958574		4.50E-06	19.8	11.3			
BDLWt_SSI	MLM	Chr6:14958574		4.60E-06	19.8	11.3			
BDLWt_SSSI	MLM	Chr6:14958574		4.50E-06	19.8	11.3			
BDLWt_DSI	MLM	Chr7:2903232		8.80E-05	14.1	14.2	Lus10011287/8	AT3G18830	PMT5
BDLWt_DTE	MLM	Chr7:2903232		8.80E-05	14.1	14.2			
BDLWt_SSI	MLM	Chr7:2903232		8.80E-05	14.1	14.2			
BDLWt_SSSI	MLM	Chr7:2903232		8.80E-05	14.1	14.2			
BDLWt_DSI	MLM	Chr7:15761400		9.20E-05	14.0	5.7	Lus10038381/2	AT4G20270	BAM3
BDLWt_DTE	MLM	Chr7:15761400		9.20E-05	14.0	5.7			
BDLWt_SSI	MLM	Chr7:15761400		9.30E-05	14.0	5.7			
BDLWt_SSSI	MLM	Chr7:15761400		9.10E-05	14.0	5.7			
BDLWt_DSI	MLM	Chr7:15761419		9.20E-05	14.0	5.7			
BDLWt_DTE	MLM	Chr7:15761419		9.20E-05	14.0	5.7			
BDLWt_SSI	MLM	Chr7:15761419		9.30E-05	14.0	5.7			
BDLWt_SSSI	MLM	Chr7:15761419		9.10E-05	14.0	5.7			
HT_NS	mrMLM	Chr8:2514743	7.60	3.30E-09	16.4	7.7	Lus10025174	AT2G42500	PP2A-3
HT_NS	FASTmrMLM	Chr8:2514743	7.74	2.40E-09	13.3	7.6	Lus10025166	AT1G76520	PILS3
HT_NS	pLARmEB	Chr8:2514743	6.96	1.50E-08	5.1	7.6			
HT_NS	ISIS EM-BLASSO	Chr8:2514743	8.08	1.10E-09	13.3	7.6			
NHT_NS	mrMLM	Chr8:2514743	3.17	1.30E-04	9.5	7.7	Lus10025163	AT1G76430	PHT1,9
HT_STI	mrMLM	Chr8:2514743	6.24	8.40E-08	15.6	7.7			
HT_STI	FASTmrMLM	Chr8:2514743	7.59	3.40E-09	12.8	7.6			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
HT_STI	pLARmEB	Chr8:2514743	9.54	3.40E-11	5.1	7.6			
HT_STI	ISIS EM-BLASSO	Chr8:2514743	10.64	2.60E-12	13.6	7.6			
SPB_NS	ISIS EM-BLASSO	Chr8:3308923	3.56	5.20E-05	10.0	6.6			
SPB_SSSI	pLARmEB	Chr8:3308923	3.83	2.70E-05	0.5	6.6			
YLD_DSI	mrMLM	Chr8:14140850	3.51	5.90E-05	6.0	7.7	Lus10016595	AT4G26640	WRKY20
YLD_SSI	mrMLM	Chr8:14140850	3.51	5.90E-05	6.0	7.7	Lus10016594	AT5G55560	WNK-kinase
YLD_DTE	mrMLM	Chr8:14140850	3.51	5.90E-05	6.0	7.7	Lus10016590	AT2G36570	LRR
YLD_SSSI	mrMLM	Chr8:14140850	3.51	5.90E-05	6.0	7.7			
SPB_S	mrMLM	Chr8:16095917	3.91	2.20E-05	13.7	8.7	Lus10022278	AT2G38470	WRKY33
BDLWt_STI	mrMLM	Chr8:16534117	4.83	2.40E-06	17.0	7.7	Lus10004554	AT2G35350	PLL1
BDLWt_STI	pLARmEB	Chr8:16534117	3.88	2.40E-05	2.9	7.6			
BDLWt_STI	ISIS EM-BLASSO	Chr8:16534117	4.70	3.30E-06	8.3	7.6			
TSW_S	FASTmrMLM	Chr8:21219971	4.12	1.30E-05	6.6	8.5	Lus10034005	AT4G15530.5	PPDK
TSW_S	FASTmrEMMA	Chr8:21219971	5.56	4.20E-07	10.5	8.5	Lus10033981	AT3G06350	EMB3004
TSW_S	pLARmEB	Chr8:21219971	4.12	1.30E-05	6.6	8.5	Lus10033999	AT3G21510	AHP1
TSW_S	ISIS EM-BLASSO	Chr8:21219971	5.32	7.40E-07	10.2	8.5	Lus10033996	AT4G03020	WD-40R
CT_STI	MLM	Chr8:9117708		1.20E-05	21.4	7.5	Lus10001657	AT1G19440	KCS4
CT_STI	MLM	Chr8:16606284		1.40E-05	21.0	7.5	Lus10004574	AT1G32640	MYC2,ZBF1
CT_STI	MLM	Chr8:16698991		2.00E-06	25.7	9.4	Lus10000312	AT1G32330	HSFA1D
CT_STI	MLM	Chr8:16722804		2.30E-06	25.3	10.4	Lus10000894	AT1G31970	STRS1
htTol	ISIS EM-BLASSO	Chr9:901876	3.81	2.80E-05	11.0	7.6	Lus10010215	AT3G19820	DIM1
						0.0	Lus10010220	AT5G67240	SDN3
BDLWt_DSI	mrMLM	Chr9:4203006	7.73	2.50E-09	29.1	5.8	Lus10040353	AT2G45570	CYP76C2
BDLWt_DSI	FASTmrMLM	Chr9:4203006	4.78	2.70E-06	12.9	5.7	Lus10040335	AT3G12360	ITN1
BDLWt_SSI	mrMLM	Chr9:4203006	7.73	2.50E-09	29.1	5.8	Lus10040357	AT4G10090	ELP6
BDLWt_SSI	FASTmrMLM	Chr9:4203006	4.78	2.70E-06	12.9	5.7	Lus10040333	AT5G04530	KCS19
BDLWt_DTE	mrMLM	Chr9:4203006	7.73	2.50E-09	29.1	5.8	Lus10040348	AT2G39290	PGS1
BDLWt_DTE	FASTmrMLM	Chr9:4203006	4.78	2.70E-06	12.9	5.7	Lus10040347	AT3G55040	GSTL2

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
BDLWt_DTE	ISIS EM-BLASSO	Chr9:4203006	4.52	5.00E-06	12.9	5.7			
BDLWt_SSSI	mrMLM	Chr9:4203006	7.73	2.50E-09	29.1	5.8			
BDLWt_SSSI	FASTmrMLM	Chr9:4203006	4.78	2.70E-06	12.9	5.7			
BDLWt_SSSI	pLARmEB	Chr9:4203006	5.64	3.50E-07	2.9	5.7			
NYLD_NS	ISIS EM-BLASSO	Chr9:7297801	3.89	2.30E-05	8.2	9.4	Lus10024511	AT4G11650	OSM34
SPB_DTE	FASTmrMLM	Chr9:15446958	6.48	4.70E-08	18.8	5.7	Lus10021766	AT5G66850	MAPKKK5
SPB_DTE	ISIS EM-BLASSO	Chr9:15446958	6.25	8.10E-08	21.7	5.7	Lus10002787	AT2G26510	PDE135
SPB_DSI	FASTmrMLM	Chr9:15446958	6.48	4.70E-08	18.8	5.7	Lus10021762	AT2G23290	MYB70
SPB_DSI	pLARmEB	Chr9:15446958	5.66	3.30E-07	1.3	5.7			
SPB_SSI	FASTmrMLM	Chr9:15446958	6.48	4.70E-08	18.8	5.7			
SPB_SSI	pLARmEB	Chr9:15446958	8.33	5.90E-10	1.7	5.7			
SPB_SSSI	FASTmrMLM	Chr9:15446958	6.48	4.70E-08	18.8	5.7			
SPB_SSSI	pLARmEB	Chr9:15446958	4.92	1.90E-06	1.2	5.7			
HT_SSSI	pLARmEB	Chr9:16446802	3.67	3.90E-05	8.0	12.3	Lus10042646	AT1G46264	HSFB4,SCZ
CT_STI	mrMLM	Chr9:18937872	6.47	4.80E-08	34.7	9.6	Lus10024816	AT4G37370	CYP81D8
						0.0	Lus10024817	AT4G37330	CYP81D4
						0.0	Lus10024818	AT4G37320	CYP81D5
						0.0	Lus10024811	AT4G36930	SPT
BDLWt_DSI	MLM	Chr9:3599111		1.60E-05	17.3	7.5	Lus10040475	AT1G49320	USPL1
BDLWt_DTE	MLM	Chr9:3599111		1.60E-05	17.3	7.5			
BDLWt_SSI	MLM	Chr9:3599111		1.60E-05	17.3	7.5			
BDLWt_SSSI	MLM	Chr9:3599111		1.70E-05	17.3	7.5			
BDLWt_DSI	MLM	Chr9:3979356		9.10E-05	14.0	6.6	Lus10040391	AT3G54890	LHCA1
BDLWt_DTE	MLM	Chr9:3979356		9.10E-05	14.0	6.6			
BDLWt_SSI	MLM	Chr9:3979356		9.10E-05	14.0	6.6			
BDLWt_SSSI	MLM	Chr9:3979356		9.10E-05	14.0	6.6			
BDLWt_DSI	MLM	Chr9:4203006		1.40E-05	17.6	5.7			
BDLWt_DTE	MLM	Chr9:4203006		1.40E-05	17.6	5.7			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
BDLWt_SSI	MLM	Chr9:4203006		1.40E-05	17.6	5.7			
BDLWt_SSSI	MLM	Chr9:4203006		1.40E-05	17.6	5.7			
BDLWt_DSI	MLM	Chr9:4398116		1.00E-06	22.9	5.7	Lus10040304	AT5G02490	Hsp70-2
BDLWt_DTE	MLM	Chr9:4398116		1.00E-06	22.9	5.7	Lus10040300	AT5G02500	HSP70-1
BDLWt_SSI	MLM	Chr9:4398116		1.00E-06	22.9	5.7			
BDLWt_SSSI	MLM	Chr9:4398116		1.00E-06	22.9	5.7			
BDLWt_DSI	MLM	Chr9:17615028		1.20E-05	17.8	12.3	Lus10042518	AT5G22290	NAC089
BDLWt_DTE	MLM	Chr9:17615028		1.20E-05	17.8	12.3	Lus10042516	AT1G11260	STP1
BDLWt_SSI	MLM	Chr9:17615028		1.20E-05	17.9	12.3			
BDLWt_SSSI	MLM	Chr9:17615028		1.20E-05	17.8	12.3			
CT_STI	MLM	Chr9:18937269		4.80E-06	23.6	9.4			
CT_STI	MLM	Chr9:18937872		1.00E-05	21.7	9.4			
CT_STI	MLM	Chr9:18938885		1.00E-05	21.7	9.4			
CT_STI	MLM	Chr9:18948993		4.80E-06	23.6	9.4			
SPB_TOL	pLARmEB	Chr10:11201761	3.49	6.10E-05	6.4	13.2	Lus10032728	AT1G19300	GLZ1,PARVUS
SPB_DTE	mrMLM	Chr10:11227648	3.81	2.80E-05	8.2	13.5	Lus10032734	AT3G50770	CML41
SPB_DSI	mrMLM	Chr10:11227648	3.81	2.80E-05	8.2	13.5	Lus10032736	AT1G19270	DA1
SPB_SSI	mrMLM	Chr10:11227648	3.81	2.80E-05	8.2	13.5			
SPB_SSSI	mrMLM	Chr10:11227648	3.81	2.80E-05	8.2	13.5			
BDLWt_STI	FASTmrMLM	Chr10:11652051	3.37	8.20E-05	6.7	7.6			
SPB_DTE	mrMLM	Chr10:13671937	4.12	1.30E-05	9.2	12.5	Lus10042707	AT1G77590	Lac9
SPB_DTE	FASTmrMLM	Chr10:13671937	4.23	1.00E-05	5.4	12.3	Lus10042698	AT4G08500	MAPKKK8
SPB_DTE	ISIS EM-BLASSO	Chr10:13671937	3.94	2.00E-05	5.8	12.3	Lus10042700	AT1G77610	UDP-GALT1
SPB_DSI	mrMLM	Chr10:13671937	4.12	1.30E-05	9.2	12.5			
SPB_DSI	FASTmrMLM	Chr10:13671937	4.23	1.00E-05	5.4	12.3			
SPB_DSI	pLARmEB	Chr10:13671937	4.62	4.00E-06	0.4	12.3			
SPB_DSI	ISIS EM-BLASSO	Chr10:13671937	3.77	3.10E-05	5.8	12.3			
SPB_SSI	mrMLM	Chr10:13671937	4.12	1.30E-05	9.2	12.5			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
SPB_SSI	FASTmrMLM	Chr10:13671937	4.23	1.00E-05	5.4	12.3			
SPB_SSI	pLARmEB	Chr10:13671937	5.16	1.10E-06	0.5	12.3			
SPB_SSI	ISIS EM-BLASSO	Chr10:13671937	3.66	4.10E-05	5.8	12.3			
SPB_SSSI	mrMLM	Chr10:13671937	4.12	1.30E-05	9.2	12.5			
SPB_SSSI	FASTmrMLM	Chr10:13671937	4.23	1.00E-05	5.4	12.3			
SPB_SSSI	pLARmEB	Chr10:13671937	4.89	2.10E-06	0.4	12.3			
SPB_SSSI	ISIS EM-BLASSO	Chr10:13671937	3.66	4.10E-05	5.8	12.3			
SPB_NS	ISIS EM-BLASSO	Chr10:14905874	3.20	1.20E-04	7.9	8.5	Lus10042933	AT4G37870	PCK1
SPB_DTE	FASTmrMLM	Chr10:14905874	3.06	1.70E-04	5.5	8.5	Lus10042929	AT5G65670	IAA9
SPB_DTE	ISIS EM-BLASSO	Chr10:14905874	4.52	5.00E-06	10.9	8.5			
SPB_DSI	FASTmrMLM	Chr10:14905874	3.06	1.70E-04	5.5	8.5			
SPB_DSI	FASTmrEMMA	Chr10:14905874	4.74	2.90E-06	0.0	8.5			
SPB_DSI	pLARmEB	Chr10:14905874	6.08	1.20E-07	0.9	8.5			
SPB_DSI	ISIS EM-BLASSO	Chr10:14905874	3.95	2.00E-05	10.9	8.5			
SPB_SSI	FASTmrMLM	Chr10:14905874	3.06	1.70E-04	5.5	8.5			
SPB_SSI	FASTmrEMMA	Chr10:14905874	4.74	2.90E-06	0.0	8.5			
SPB_SSI	pLARmEB	Chr10:14905874	4.48	5.60E-06	0.6	8.5			
SPB_SSI	ISIS EM-BLASSO	Chr10:14905874	4.59	4.30E-06	10.9	8.5			
SPB_SSSI	FASTmrMLM	Chr10:14905874	3.06	1.70E-04	5.5	8.5			
SPB_SSSI	FASTmrEMMA	Chr10:14905874	4.74	2.90E-06	0.0	8.5			
SPB_SSSI	pLARmEB	Chr10:14905874	3.87	2.40E-05	0.9	8.5			
SPB_SSSI	ISIS EM-BLASSO	Chr10:14905874	4.59	4.30E-06	10.9	8.5			
BDLWt_DSI	MLM	Chr10:12528281		1.50E-06	22.0	11.3	Lus10010578/9	AT3G13225	FNBP4
BDLWt_DTE	MLM	Chr10:12528281		1.50E-06	22.0	11.3			
BDLWt_SSI	MLM	Chr10:12528281		1.60E-06	22.0	11.3			
BDLWt_SSSI	MLM	Chr10:12528281		1.50E-06	22.0	11.3			
BDLWt_DSI	MLM	Chr10:15197308		9.90E-05	13.9	9.4	Lus10018651	AT4G07960	CSLC12
BDLWt_DTE	MLM	Chr10:15197308		9.90E-05	13.9	9.4			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
BDLWt_SSI	MLM	Chr10:15197308		9.90E-05	13.9	9.4			
BDLWt_SSSI	MLM	Chr10:15197308		9.90E-05	13.9	9.4			
YLD_TOL	mrMLM	Chr11:3124555	4.06	1.50E-05	12.2	10.6	Lus10042035	AT2G41510	CKX1
YLD_TOL	FASTmrMLM	Chr11:3124555	3.38	8.10E-05	6.3	10.4	Lus10042040	AT3G57330	ACA11
YLD_TOL	FASTmrEMMA	Chr11:3124555	3.23	1.20E-04	0.0	10.4			
YLD_TOL	pLARmEB	Chr11:3124555	3.38	8.10E-05	5.8	10.4			
HT_SSSI	pLARmEB	Chr11:3226540	3.68	3.80E-05	8.1	9.4	Lus10042079	AT1G10710	PHS1
						0.0	Lus10042076	AT5G50230	ATG16
YLD_DSI	FASTmrMLM	Chr11:3972867	5.56	4.20E-07	15.5	7.6	Lus10042219	AT1G29930	CAB1
YLD_DSI	FASTmrEMMA	Chr11:3972867	5.03	1.50E-06	19.2	7.6	Lus10042229	AT5G58380	SIP1
YLD_DSI	ISIS EM-BLASSO	Chr11:3972867	5.56	4.20E-07	15.5	7.6	Lus10042218	AT3G49400	WD40R
YLD_SSI	FASTmrMLM	Chr11:3972867	5.56	4.20E-07	15.5	7.6	Lus10042215	AT5G60870	RCC1(RUG3)
YLD_SSI	FASTmrEMMA	Chr11:3972867	5.03	1.50E-06	19.2	7.6	Lus10042216	AT1G04260	MPIP7
YLD_SSI	ISIS EM-BLASSO	Chr11:3972867	5.03	1.50E-06	15.5	7.6	Lus10042231/2	AT1G06950	TIC110
YLD_DTE	FASTmrMLM	Chr11:3972867	5.03	1.50E-06	15.5	7.6			
YLD_DTE	FASTmrEMMA	Chr11:3972867	6.08	1.20E-07	19.2	7.6			
YLD_DTE	pLARmEB	Chr11:3972867	4.82	2.50E-06	7.6	7.6			
YLD_DTE	ISIS EM-BLASSO	Chr11:3972867	5.03	1.50E-06	15.5	7.6			
YLD_TOL	FASTmrMLM	Chr11:3972867	3.69	3.70E-05	7.6	7.6			
YLD_TOL	pLARmEB	Chr11:3972867	3.69	3.70E-05	6.9	7.6			
YLD_TOL	ISIS EM-BLASSO	Chr11:3972867	3.28	1.00E-04	6.1	7.6			
YLD_SSSI	FASTmrMLM	Chr11:3972867	5.56	4.20E-07	15.5	7.6			
YLD_SSSI	FASTmrEMMA	Chr11:3972867	5.03	1.50E-06	19.2	7.6			
YLD_SSSI	ISIS EM-BLASSO	Chr11:3972867	5.56	4.20E-07	15.5	7.6			
YLD_NS	FASTmrMLM	Chr11:4223430	4.50	5.30E-06	10.4	6.6	Lus10026362	AT5G45890	SAG12
YLD_NS	ISIS EM-BLASSO	Chr11:4223430	3.56	5.20E-05	6.8	6.6	Lus10026377/8	AT2G42010	PLDBETA1
TSW_STI	pLARmEB	Chr11:7466316	4.43	6.20E-06	3.1	5.7	Lus10036525	AT2G28880	emb1997
TSW_STI	ISIS EM-BLASSO	Chr11:7466316	3.23	1.20E-04	8.2	5.7	Lus10036520	AT5G53000	TAP46

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
TSW_TOL	mrMLM	Chr11:18971874	3.89	2.30E-05	4.9	6.7	Lus10031275	AT3G63530	BB2
						0.0	Lus10031280	AT4G36250	ALDH3F1
BDLWt_DSI	MLM	Chr11:5277154		6.90E-06	19.0	9.4	Lus10038178	AT4G31940	CYP82C4
BDLWt_DTE	MLM	Chr11:5277154		6.90E-06	19.0	9.4			
BDLWt_SSI	MLM	Chr11:5277154		6.90E-06	19.0	9.4			
BDLWt_SSSI	MLM	Chr11:5277154		6.90E-06	19.0	9.4			
HT_DSI	ISIS EM-BLASSO	Chr12:3878445	3.62	4.50E-05	4.6	7.6	Lus10005864	AT5G58850	MYB119
HT_SSI	ISIS EM-BLASSO	Chr12:3878445	3.62	4.50E-05	4.5	7.6			
HT_SSSI	ISIS EM-BLASSO	Chr12:3878445	3.62	4.50E-05	4.6	7.6			
BDLWt_DSI	FASTmrMLM	Chr12:6352775	8.84	1.80E-10	25.9	10.4	Lus10016846	AT3G16857	RR1
BDLWt_DSI	FASTmrEMMA	Chr12:6352775	5.79	2.40E-07	0.0	10.4	Lus10016831	AT3G05640	EGR1
BDLWt_DSI	pLARmEB	Chr12:6352775	9.19	7.80E-11	6.7	10.4	Lus10016833	AT5G10530	LECRK-IX
BDLWt_DSI	ISIS EM-BLASSO	Chr12:6352775	8.86	1.70E-10	28.0	10.4			
BDLWt_SSI	FASTmrMLM	Chr12:6352775	8.84	1.80E-10	25.9	10.4			
BDLWt_SSI	FASTmrEMMA	Chr12:6352775	5.79	2.40E-07	0.0	10.4			
BDLWt_SSI	pLARmEB	Chr12:6352775	7.31	6.50E-09	7.9	10.4			
BDLWt_SSI	ISIS EM-BLASSO	Chr12:6352775	8.86	1.70E-10	28.0	10.4			
BDLWt_DTE	FASTmrMLM	Chr12:6352775	8.84	1.80E-10	25.9	10.4			
BDLWt_DTE	pLARmEB	Chr12:6352775	9.48	3.90E-11	6.4	10.4			
BDLWt_DTE	ISIS EM-BLASSO	Chr12:6352775	9.48	3.90E-11	28.0	10.4			
BDLWt_SSSI	FASTmrMLM	Chr12:6352775	8.84	1.80E-10	25.9	10.4			
BDLWt_SSSI	FASTmrEMMA	Chr12:6352775	5.79	2.40E-07	0.0	10.4			
BDLWt_SSSI	pLARmEB	Chr12:6352775	8.05	1.20E-09	6.4	10.4			
BDLWt_SSSI	ISIS EM-BLASSO	Chr12:6352775	8.86	1.70E-10	28.0	10.4			
CT_NS	FASTmrMLM	Chr12:7437241	3.74	3.40E-05	10.2	15.1	Lus10001113	AT3G14090	EXO70D3
						0.0	Lus10001112	AT1G72470	EXO70D1
TSW_DSI	FASTmrMLM	Chr12:10910146	4.34	7.80E-06	11.7	17.9	Lus10030142	AT1G79280	NUA
TSW_DSI	pLARmEB	Chr12:10910146	4.02	1.70E-05	9.2	17.9	Lus10030134	AT5G19450	CDPK19

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
TSW_DSI	ISIS EM-BLASSO	Chr12:10910146	3.23	1.10E-04	12.4	17.9			
TSW_SSI	FASTmrMLM	Chr12:10910146	3.95	2.00E-05	11.7	17.9			
TSW_SSI	pLARmEB	Chr12:10910146	3.84	2.60E-05	7.7	17.9			
TSW_SSI	ISIS EM-BLASSO	Chr12:10910146	3.24	1.10E-04	12.4	17.9			
TSW_DTE	FASTmrMLM	Chr12:10910146	3.95	2.00E-05	11.7	17.9			
TSW_DTE	ISIS EM-BLASSO	Chr12:10910146	3.23	1.10E-04	12.4	17.9			
TSW_TOL	mrMLM	Chr12:10910146	4.05	1.60E-05	16.6	18.3			
TSW_TOL	FASTmrMLM	Chr12:10910146	3.90	2.20E-05	12.6	17.9			
TSW_TOL	pLARmEB	Chr12:10910146	4.15	1.20E-05	8.6	17.9			
TSW_TOL	ISIS EM-BLASSO	Chr12:10910146	4.23	1.00E-05	12.6	17.9			
TSW_SSSI	FASTmrMLM	Chr12:10910146	4.34	7.80E-06	11.7	17.9			
TSW_SSSI	pLARmEB	Chr12:10910146	4.43	6.30E-06	7.6	17.9			
TSW_SSSI	ISIS EM-BLASSO	Chr12:10910146	3.23	1.10E-04	12.4	17.9			
BDLWt_SSI	pLARmEB	Chr12:17400297	3.07	1.70E-04	0.0	5.7	Lus10043237	AT2G20020	CAF1
CT_TOL	mrMLM	Chr12:19686955	3.52	5.70E-05	12.5	5.8	Lus10031607	AT4G28110	MYB41
YLD_TOL	mrMLM	Chr12:20557728	4.65	3.70E-06	16.7	7.7	Lus10031398	AT3G06720	IMPA1
YLD_TOL	FASTmrMLM	Chr12:20557728	4.93	1.90E-06	11.2	7.6			
YLD_TOL	pLARmEB	Chr12:20557728	4.93	1.90E-06	10.2	7.6			
YLD_TOL	ISIS EM-BLASSO	Chr12:20557728	6.34	6.50E-08	12.8	7.6			
CT_TOL	MLM	Chr12:1853487		1.10E-06	26.7	7.5	Lus10023339	AT3G45640	MPK3
CT_TOL	MLM	Chr12:1853750		3.40E-06	23.9	8.5	Lus10023332	AT2G28610	PRS1,WOX3
CT_NS	FASTmrMLM	Chr13:3465862	4.37	7.30E-06	9.2	7.6	Lus10026051	AT1G69400	WD40R
CT_NS	ISIS EM-BLASSO	Chr13:3465862	3.56	5.10E-05	6.5	7.6	Lus10026039	AT5G66680	DGL1
TSW_S	ISIS EM-BLASSO	Chr13:3484923	3.68	3.90E-05	7.0	17.0	Lus10026057	AT2G44450	BGLU15
TSW_TOL	mrMLM	Chr13:3886742	5.89	1.90E-07	21.0	6.7	Lus10026154	AT3G12280	RBR1
YLD_STI	FASTmrMLM	Chr13:4208860	3.80	2.90E-05	8.4	7.6	Lus10019778	AT2G38770	EMB2765
YLD_STI	pLARmEB	Chr13:4208860	3.07	1.70E-04	3.3	7.6	Lus10019780	AT5G12380	ANNAT8
YLD_STI	ISIS EM-BLASSO	Chr13:4208860	4.32	8.10E-06	11.6	7.6			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
YLD_NS	mrMLM	Chr13:4581161	4.22	1.00E-05	20.0	6.7	Lus10019701/2	AT5G13490	AAC2
YLD_NS	FASTmrMLM	Chr13:4581161	3.04	1.80E-04	5.6	6.6	Lus10019706	AT5G13450	ATP5
YLD_STI	FASTmrMLM	Chr13:4581161	3.09	1.60E-04	5.0	6.6	Lus10019703	AT3G63530	EOD1
TSW_NS	mrMLM	Chr13:11108918	3.74	3.30E-05	13.5	7.7	Lus10032862	AT1G74970	RPS9,TWN3
TSW_NS	FASTmrMLM	Chr13:11108918	4.40	6.80E-06	8.0	8.5			
TSW_NS	pLARmEB	Chr13:11108918	4.40	6.80E-06	7.4	8.5			
TSW_NS	ISIS EM-BLASSO	Chr13:11108918	4.14	1.30E-05	7.2	8.5	Lus10032868	AT2G26490	WD40R
TSW_S	FASTmrMLM	Chr13:11108918	3.66	4.00E-05	5.6	8.5	Lus10032861	AT5G67310	CYP81G1
TSW_S	pLARmEB	Chr13:11108918	3.66	4.00E-05	5.7	8.5	Lus10032869	AT4G31940	CYP82C4
TSW_S	ISIS EM-BLASSO	Chr13:11108918	4.09	1.40E-05	7.1	8.5			
TSW_STI	mrMLM	Chr13:11108918	3.76	3.20E-05	13.4	7.7			
TSW_STI	FASTmrMLM	Chr13:11108918	4.47	5.80E-06	7.6	8.5			
TSW_STI	pLARmEB	Chr13:11108918	4.60	4.10E-06	2.6	8.5			
TSW_STI	ISIS EM-BLASSO	Chr13:11108918	5.81	2.30E-07	9.9	8.5			
SPB_TOL	pLARmEB	Chr13:11699277	3.99	1.80E-05	8.7	9.4	Lus10009756	AT1G45230	TGS1
BDLWt_DSI	MLM	Chr13:4363772		1.30E-05	17.8	11.3			
BDLWt_DTE	MLM	Chr13:4363772		1.30E-05	17.8	11.3			
BDLWt_SSI	MLM	Chr13:4363772		1.30E-05	17.8	11.3			
BDLWt_SSSI	MLM	Chr13:4363772		1.30E-05	17.8	11.3			
BDLWt_DSI	MLM	Chr13:19063115		9.90E-05	13.9	9.4	Lus10030494	AT2G02820	MYB88
BDLWt_DTE	MLM	Chr13:19063115		9.90E-05	13.9	9.4	Lus10030492	AT2G02860	SUC3,SUT2
BDLWt_SSI	MLM	Chr13:19063115		9.90E-05	13.9	9.4			
BDLWt_SSSI	MLM	Chr13:19063115		9.90E-05	13.9	9.4			
NHT_NS	mrMLM	Chr14:205508	5.64	3.50E-07	19.8	7.7	Lus10009476	AT1G72180	CEPR2
NHT_NS	FASTmrMLM	Chr14:205508	4.18	1.10E-05	9.6	7.6	Lus10009481	AT4G09960	AGL11,STK
NHT_NS	pLARmEB	Chr14:205508	4.18	1.10E-05	9.6	7.6			
NHT_NS	ISIS EM-BLASSO	Chr14:205508	3.23	1.10E-04	8.9	7.6			
HT_DSI	FASTmrMLM	Chr14:205508	3.06	1.70E-04	10.9	7.6			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
HT_DSI	pLARmEB	Chr14:205508	4.08	1.50E-05	9.9	7.6			
HT_DTE	FASTmrMLM	Chr14:205508	3.84	2.60E-05	10.8	7.6	Lus10009472	AT1G71692	AGL12,XAL1
HT_DTE	pLARmEB	Chr14:205508	3.84	2.60E-05	9.9	7.6	Lus10009480	AT1G15360	SHN1,WIN1
HT_DTE	ISIS EM-BLASSO	Chr14:205508	3.95	2.00E-05	10.4	7.6			
HT_SSI	FASTmrMLM	Chr14:205508	4.08	1.50E-05	10.8	7.6			
BDLWt_NS	FASTmrMLM	Chr14:205508	5.11	1.20E-06	15.5	7.6			
BDLWt_NS	ISIS EM-BLASSO	Chr14:205508	5.11	1.20E-06	15.5	7.6			
BDLWt_STI	FASTmrMLM	Chr14:205508	3.69	3.70E-05	6.3	7.6			
BDLWt_STI	pLARmEB	Chr14:205508	4.11	1.40E-05	2.6	7.6			
BDLWt_STI	ISIS EM-BLASSO	Chr14:205508	3.70	3.70E-05	6.5	7.6			
HT_SSSI	FASTmrMLM	Chr14:205508	3.06	1.70E-04	10.9	7.6			
CT_TOL	FASTmrMLM	Chr14:1479557	4.05	1.60E-05	6.6	9.4	Lus10028694	AT1G71696	SOL1
CT_TOL	ISIS EM-BLASSO	Chr14:1479557	3.84	2.60E-05	5.6	9.4			
CT_DSI	FASTmrMLM	Chr14:3671638	4.21	1.10E-05	8.9	6.6	Lus10021393/4	AT2G19070	SHT
CT_DSI	pLARmEB	Chr14:3671638	3.73	3.40E-05	3.2	6.6	Lus10021395	AT5G48930	HCT
CT_DSI	ISIS EM-BLASSO	Chr14:3671638	3.73	3.40E-05	8.9	6.6			
CT_SSI	FASTmrMLM	Chr14:3671638	4.21	1.10E-05	8.9	6.6			
CT_SSI	pLARmEB	Chr14:3671638	3.97	1.90E-05	3.6	6.6			
CT_SSI	ISIS EM-BLASSO	Chr14:3671638	3.73	3.40E-05	8.9	6.6			
CTR	FASTmrMLM	Chr14:3671638	4.21	1.10E-05	8.9	6.6			
CTR	ISIS EM-BLASSO	Chr14:3671638	4.21	1.10E-05	8.9	6.6			
CT_TOL	ISIS EM-BLASSO	Chr14:3671638	3.14	1.40E-04	3.5	6.6			
CT_SSSI	FASTmrMLM	Chr14:3671638	4.21	1.10E-05	8.9	6.6			
CT_SSSI	pLARmEB	Chr14:3671638	3.46	6.60E-05	3.6	6.6			
CT_SSSI	ISIS EM-BLASSO	Chr14:3671638	3.73	3.40E-05	8.9	6.6			
TSW_DSI	mrMLM	Chr14:5390042	3.04	1.80E-04	5.3	8.7	Lus10014156/7	AT3G08850	RAPTOR1B
TSW_SSI	mrMLM	Chr14:5390042	3.04	1.80E-04	5.3	8.7			
TSW_DTE	mrMLM	Chr14:5390042	3.66	4.00E-05	7.4	8.7			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
TSW_SSSI	mrMLM	Chr14:5390042	3.04	1.80E-04	5.3	8.7			
BDLWt_TOL	FASTmrEMMA	Chr14:6181390	3.31	9.40E-05	8.0	19.8	Lus10015684	AT5G51600	MAP65-3,PLE
BDLWt_STI	pLARmEB	Chr14:13250645	3.54	5.40E-05	2.7	6.6	Lus10032887	AT5G56270	WRKY2
BDLWt_S	FASTmrMLM	Chr14:13428471	5.10	1.30E-06	10.4	7.6	Lus10032919	AT5G24590	TIP
BDLWt_S	pLARmEB	Chr14:13428471	5.54	4.40E-07	9.2	7.6	Lus10032926	AT1G05260	RCI3
BDLWt_S	ISIS EM-BLASSO	Chr14:13428471	4.77	2.80E-06	10.6	7.6			
CT_STI	mrMLM	Chr14:16353008	3.53	5.50E-05	5.1	5.8	Lus10018062	AT4G34740	CIA1
YLD_NS	FASTmrMLM	Chr14:16861788	3.21	1.20E-04	5.0	7.6	Lus10017963	AT2G31880	EVR,SOBIR1
BDLWt_DSI	MLM	Chr14:7308449		2.90E-05	16.2	14.2	Lus10004079	AT3G63300	FKD1
BDLWt_DTE	MLM	Chr14:7308449		2.90E-05	16.2	14.2	Lus10004076	AT1G19670	CLH1,CORI1
BDLWt_SSI	MLM	Chr14:7308449		2.90E-05	16.2	14.2			
BDLWt_SSSI	MLM	Chr14:7308449		2.90E-05	16.2	14.2			
YLD_DSI	mrMLM	Chr15:7035946	6.04	1.30E-07	15.4	8.7	Lus10002673	AT2G46820	PSAP,PSI-P
YLD_SSI	mrMLM	Chr15:7035946	6.04	1.30E-07	15.4	8.7	Lus10002670	AT1G13090	CYP71B28
YLD_DTE	mrMLM	Chr15:7035946	6.04	1.30E-07	15.4	8.7	Lus10002671	AT3G26330	CYP71B37
YLD_SSSI	mrMLM	Chr15:7035946	6.04	1.30E-07	15.4	8.7			
YLD_STI	FASTmrMLM	Chr15:12802900	3.50	5.90E-05	6.6	12.3	Lus10010356	AT3G04340	emb2458
YLD_STI	FASTmrEMMA	Chr15:12802900	3.92	2.10E-05	9.8	12.3	Lus10010353	AT1G78240	TSD2
YLD_STI	pLARmEB	Chr15:12802900	4.34	7.80E-06	3.9	12.3			
TSW_DSI	mrMLM	Chr15:14329302	4.50	5.30E-06	6.7	7.7	Lus10037864	AT2G45830	AGL15-2 (DTA2)
TSW_DSI	FASTmrMLM	Chr15:14329302	3.53	5.60E-05	9.2	7.6	Lus10037856	AT4G00710	BSK3
TSW_DSI	pLARmEB	Chr15:14329302	4.65	3.70E-06	7.2	7.6			
TSW_DSI	ISIS EM-BLASSO	Chr15:14329302	4.86	2.20E-06	9.5	7.6			
TSW_SSI	mrMLM	Chr15:14329302	4.50	5.30E-06	6.7	7.7			
TSW_SSI	FASTmrMLM	Chr15:14329302	4.43	6.30E-06	9.1	7.6			
TSW_SSI	pLARmEB	Chr15:14329302	3.09	1.60E-04	5.9	7.6			
TSW_SSI	ISIS EM-BLASSO	Chr15:14329302	4.50	5.30E-06	9.5	7.6			
TSW_DTE	FASTmrMLM	Chr15:14329302	4.43	6.30E-06	9.1	7.6			

Trait	Method[1]	QTN[2]	LOD[3]	*P* value	r2 (%)[4]	MAF (%)[5]	Candidate gene	At Orth[6]	Pred_Protien
TSW_DTE	ISIS EM-BLASSO	Chr15:14329302	4.86	2.20E-06	9.5	7.6			
TSW_TOL	mrMLM	Chr15:14329302	4.21	1.10E-05	9.7	7.7			
TSW_TOL	FASTmrMLM	Chr15:14329302	4.52	5.00E-06	6.0	7.6			
TSW_TOL	pLARmEB	Chr15:14329302	3.97	1.90E-05	4.1	7.6			
TSW_TOL	ISIS EM-BLASSO	Chr15:14329302	3.50	5.90E-05	6.0	7.6			
TSW_SSSI	mrMLM	Chr15:14329302	4.50	5.30E-06	6.7	7.7			
TSW_SSSI	FASTmrMLM	Chr15:14329302	3.53	5.60E-05	9.2	7.6			
TSW_SSSI	pLARmEB	Chr15:14329302	3.39	7.70E-05	5.8	7.6			
TSW_SSSI	ISIS EM-BLASSO	Chr15:14329302	4.86	2.20E-06	9.5	7.6			

[1]MLM indicates single locus GWAS using mixed linear model, all the remaining models are multilocus based GWAS.
[2]QTN=quantitative trait nucleotide; number before and after the underscore indicate the chromosome and position of the QTN on the chromosome respectively.
[3]LOD=logarithm of odds.
[4]R^2=Coefficient of determination indicating PVE.
[5]MAF=Minor allele frequency.
[6]At Orth= *Arabidopsis thaliana* orthologue.

III-5 Table III-3 QTN detected by more than two methods

QTN[1]	Trait variable	Number of methods	LOD[2]	-Log*P*[3]	R^2(%)[4]	MAF(%)[5]
Chr1:17732021	HT_DSI	4	3-3.8	3.72-4.49	0-10.7	5.7
	HT_DTE	3	3.9-4.3	4.62-5.04	10-10.8	5.7
	HT__SSI	3	3.1-3.8	3.78-4.49	0-10.8	5.7
	HT_SSSI	3	3-3.8	3.72-4.49	0-10.7	5.7
Chr1:20699803	BL25_SDCt_S	2	3.2-3.2	3.88-3.88	7-	13.2
Chr1:20933543	BDLWt_STI	3	3.2-3.8	3.92-4.51	1.8-5.2	5.7
Chr1:2649874	BDLWt_S	2	3.5-3.6	4.18-4.28	5.5-6.2	6.6
Chr1:3237615	CT_STI	2	3.1-3.2	3.75-3.95	4.3-6.6	5.7
Chr1:7029199	TSW_STI	3	4.4-6.5	5.17-7.36	13.9-20.2	15.1
Chr10:13672926	BL25_SDCt_DSI	4	3.8-4.6	4.51-5.4	0.4-9.2	12.3
	BL25_SDCt_DTE	3	3.9-4.2	4.69-5	5.4-9.2	12.3
	BL25_SDCt_SSI	4	3.7-5.2	4.39-5.97	0.5-9.2	12.3
	BL25_SDCt_SSSI	4	3.7-4.9	4.39-5.68	0.4-9.2	12.3
Chr10:14906881	BL25_SDCt_DSI	4	3.1-6.1	3.76-6.92	0-10.9	8.5
	BL25_SDCt_DTE	2	3.1-4.5	3.76-5.3	5.5-10.9	8.5
	BL25_SDCt_SSI	4	3.1-4.7	3.76-5.53	0-10.9	8.5
	BL25_SDCt_SSSI	4	3.1-4.7	3.76-5.53	0-10.9	8.5
Chr11:3124895	YLD_TOL	4	3.2-4.1	3.94-4.81	0-12.2	10.4
Chr11:3973246	YLD_DSI	3	5-5.6	5.82-6.37	15.5-19.2	7.6
	YLD_DTE	4	4.8-6.1	5.61-6.91	7.6-19.2	7.6
	YLD_SSI	3	5-5.6	5.82-6.37	15.5-19.2	7.6
	YLD_SSSI	3	5-5.6	5.82-6.37	15.5-19.2	7.6
	YLD_TOL	3	3.3-3.7	4-4.43	6.1-7.6	7.6
Chr11:4223809	YLD_NS	2	3.6-4.5	4.29-5.28	6.8-10.4	6.6
Chr11:7466809	TSW_STI	2	3.2-4.4	3.94-5.2	3.1-8.2	5.7
Chr12:10911022	TSW_DSI	3	3.2-4.3	3.94-5.11	9.2-12.4	17.9
	TSW_DTE	2	3.2-4	3.94-4.7	11.7-12.5	17.9
	TSW_SSI	3	3.2-4	3.95-4.7	7.7-12.5	17.9
	TSW_SSSI	3	3.2-4.4	3.94-5.2	7.6-12.4	17.9
	TSW_TOL	4	3.9-4.2	4.65-5	8.6-16.6	17.9
Chr12:20558721	YLD_TOL	4	4.7-6.3	5.43-7.19	10.2-16.7	7.6
Chr12:6353095	BDLWt_DSI	4	5.8-9.2	6.62-10.11	0-28	10.4
	BDLWt_DTE	3	8.8-9.5	9.75-10.41	6.4-28	10.4
	BDLWt_SSI	4	5.8-8.9	6.62-9.78	0-28	10.4
	BDLWt_SSSI	4	5.8-8.9	6.62-9.78	0-28	10.4
Chr13:11109568	TSW_NS	4	3.7-4.4	4.48-5.17	7.2-13.6	7.7
	TSW_S	3	3.7-4.1	4.4-4.85	5.7-7.1	8.5

QTN[1]	Trait variable	Number of methods	LOD[2]	-Log*P*[3]	R^2(%)[4]	MAF(%)[5]
	TSW_STI	4	3.8-5.8	4.5-6.64	2.6-13.4	7.7
Chr13:3466056	CT_NS	2	3.6-4.4	4.29-5.14	6.5-9.2	7.6
Chr13:4209079	YLD_STI	3	3.1-4.3	3.77-5.09	3.3-11.6	7.6
Chr13:4581423	YLD_NS	2	3-4.2	3.74-4.98	5.6-20	6.6
Chr14:13429427	BDLWt_NS	3	4.8-5.5	5.56-6.35	9.2-10.6	7.6
Chr14:1479760	CT_TOL	2	3.8-4	4.58-4.8	5.6-6.6	9.4
Chr14:205605	BDLWt_STI	3	3.7-4.1	4.43-4.86	2.6-6.5	7.6
	HT_DSI	2	3.1-4.1	3.76-4.84	9.9-10.9	7.6
	HT_DTE	3	3.8-4	4.59-4.7	9.9-10.8	7.6
	BDLWt_NS	2	5.1	5.91	15.5	7.6
	HT_S	4	3.2-5.6	3.94-6.46	8.9-19.8	7.6
Chr14:3672177	CT_DSI	3	3.7-4.2	4.47-4.97	3.2-8.9	6.6
	CT_DTE	2	4.2	4.97	8.9	6.6
	CT_SSI	3	3.7-4.2	4.47-4.97	3.6-8.9	6.6
	CT_SSSI	3	3.5-4.2	4.18-4.97	3.6-8.9	6.6
Chr15:12803550	YLD_STI	3	3.5-4.3	4.23-5.11	3.9-9.8	12.3
Chr15:14329974	TSW_DSI	4	3.5-4.9	4.25-5.65	6.7-9.5	7.6
	TSW_DTE	2	4.4-4.9	5.2-5.65	9.1-9.5	7.6
	TSW_SSI	4	3.1-4.5	3.79-5.28	5.9-9.5	7.6
	TSW_S_SSI	4	3.4-4.9	4.11-5.65	5.8-9.5	7.6
	TSW_TOL	4	3.5-4.5	4.23-5.3	3.5-4.5	7.6
Chr2:23124799	CT_DSI	3	5.2-5.4	5.98-6.23	6-16.7	9.4
	CT_DTE	2	5.2	5.98	16.7	9.4
	CT_SSI	3	3.3-5.4	4.05-6.23	7.5-16.7	9.4
	CT_SSSI	3	5.2-5.4	5.98-6.23	7.5-16.7	9.4
	CT_TOL	2	3.4-3.7	4.07-4.42	6-8.4	9.4
Chr2:3816726	HT_STI	3	3.2-4.9	3.87-5.66	1.9-4.9	8.5
	HT_S	2	3.9	4.65	6.7	8.5
Chr2:4509531	BL25_SDCt_SSI	2	3.6-4.1	4.36-4.87	0.7-8.9	7.6
Chr3:23321080	BDLWt_DSI	3	3.7-4.3	4.42-5.1	1.8-8.1	8.5
	BDLWt_DTE	2	3.6-4.3	4.34-5.1	7-8.1	8.5
	BDLWt_SSI	2	3.7-4.3	4.42-5.1	7-8.1	8.5
	BDLWt_SSSI	3	3.7-4.5	4.42-5.32	1.6-8.1	8.5
Chr3:3891064	HT_STI	3	3.6-5.1	4.35-5.91	1.9-5.3	19.8
	HT_NS	3	3.7-4.4	4.39-5.19	5.4-9.4	19.8
Chr4:13311248	BL25_SDCt_DSI	5	3.4-5.8	4.11-6.63	0-14.6	8.5
	BL25_SDCtR	3	3.4-5.5	4.11-6.34	6.6-14.6	8.5
	BL25_SDCt_SSI	5	3.4-5.9	4.11-6.7	0-14.6	8.5
	BL25_SDCt_SSSI	4	3.4-5.9	4.11-6.7	0-14.6	8.5

Appendices

QTN[1]	Trait variable	Number of methods	LOD[2]	-Log*P*[3]	R^2(%)[4]	MAF(%)[5]
Chr4:9281754	CT_S	2	3.7	4.47	10.1	12.3
Chr5:1375696	HT_STI	3	3.8-6.3	4.53-7.18	4.6-15.8	8.5
	HT_NS	3	3.5-6	4.19-6.83	5.2-15.9	8.5
Chr5:1651099	HT_STI	2	3.7-5.1	4.42-5.85	2.6-5	10.4
Chr6:13339627	HT_STI	3	3.6-5.8	4.34-6.66	2.3-9.3	11.3
Chr6:13366027	HT_NS	2	3.6-5	4.35-5.79	6	7.6
Chr6:3310405	TSW_NS	2	5	5.82	15.3-16.6	12.3
	TSW_S	2	3.5-3.5	4.2-4.2	9.8-9.8	12.3
Chr8:16534685	BDLWt_STI	3	3.9-4.8	4.63-5.62	2.9-17	7.6
Chr8:21220902	TSW_S	4	4.1-5.6	4.87-6.38	6.6-10.5	8.5
Chr8:2514780	HT_STI	4	6.2-10.6	7.08-11.59	5.1-15.6	7.6
	HT_NS	4	7-8.1	7.82-8.97	5.1-16.4	7.6
Chr9:15447798	BL25_SDCt_DSI	2	5.7-6.5	6.49-7.33	1.3-18.8	5.7
	BL25_SDCt_DTE	2	6.3-6.5	7.09-7.33	18.8-21.7	5.7
	BL25_SDCt_SSI	2	6.5-8.3	7.33-9.23	1.7-18.8	5.7
	BL25_SDCt_SSSI	2	4.9-6.5	5.71-7.33	1.2-18.8	5.7
Chr9:4203205	BDLWt_DSI	2	4.8-7.7	5.56-8.61	12.9-29.1	5.7
	BDLWt_DTE	3	4.5-7.7	5.3-8.61	12.9-29.1	5.7
	BDLWt_SSI	2	4.8-7.7	5.56-8.61	12.9-29.1	5.7
	BDLWt_SSSI	3	4.8-7.7	5.56-8.61	2.9-29.1	5.7

[1]QTN=Quantitative trait nucleotides
[2]LOD=Logarithm of odds
[3]-Log*P*= -log of *P* value
[4]R2=Coefficient of determination indicating percent phenotypic variance explained
[5]MAF=Minor allele frequency