Dissertations, Master's Theses and Master's Reports

2018

Quantitative Genetics of Populus

Roba Bdeir

Michigan Technological University, rbdeir@mtu.edu

Recommended Citation

Bdeir, Roba, "Quantitative Genetics of Populus", Open Access Dissertation, Michigan Technological University, 2018.
https://digitalcommons.mtu.edu/etdr/709

QUANTITATIVE GENETICS OF *POPULUS*

By

Roba Bdeir

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Biochemistry and Molecular Biology

MICHIGAN TECHNOLOGICAL UNIVERSITY

2018

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Biochemistry and Molecular Biology.

School of Forest Resources and Environmental Science

Dissertation Co-Advisor: *Oliver Gailing*

Dissertation Co-Advisor: *Victor Busov*

Committee Member: *Gerald Tuskan*

Committee Member: *Thomas Werner*

School Dean: *Andrew J. Storer*

Dedication

Dedicated to my beloved parents

Mohammad Aiman Bdeir

&

Randa Alsahli (1974-2009)

Table of Contents

Preface

The second chapter is a manuscript published in BMC Plant Biology (Bdeir et al, 2017, DOI 10.1186/s12870-017-1166-4), on which I am first author. My co-advisors, Dr. Oliver Gailing and Dr. Victor Busov, and co-authors Dr. Yordan Yordano, Dr. Wellington Muchero and Dr. Gerald Tuskan helped conceive and design the study. All provided helpful input on the presented paper and helped in writing and revising the manuscript.

The third chapter comprises a manuscript submitted to BMC Plant Biology on which I am first author. My co-advisors, Dr. Oliver Gailing and Dr. Victor Busov, and co-authors Dr. Yordan Yordanov and Dr. Gerald Tuskan helped conceive and design the study. Dr. Wellington Muchero helped by running the association mapping tests. All provided helpful input on the presented paper and helped in writing and revising the manuscript.

The fourth chapter comprises a manuscript draft to be submitted to Tree Genetics and Genomics on which I am first author. My co-advisors, Dr. Oliver Gailing and Dr. Victor Busov, and co-authors Dr. Matias Kirst, Dr. Wellington Muchero and Dr. Gerald Tuskan helped conceive and design the study. All provided helpful input on the presented paper and helped in writing and revising the manuscript.

Acknowledgments

I would like to extend thanks to the many people, in many countries, who so generously contributed to the work presented in this thesis.

Special mention goes to my enthusiastic supervisors, Oliver Gailing and Victor Busov. My PhD has been an amazing experience and I thank you both wholeheartedly, not only for their tremendous academic support, but also for giving me so many wonderful opportunities. I am also hugely appreciative for sharing their expertise so willingly.

Similar, profound gratitude goes to Gerald Tuskna and Wellington Muchero, who has been a truly dedicated mentor. I am particularly indebted to Wellington for his constant faith in my work, and for his support when so generously hosting me in Oregon. I have very fond memories of my time there.

Special mention goes to Thomas Werner, Matias Krist, Yordan Yordanov and Soha Albukhari, for helping me along my journey.

Finally, but by no means least, thanks go to mom, dad, Moe, my husband Ahmad and my bundle of joy Randa for their unbelievable support. They are the most important people in my world and I dedicate this thesis to them.

Abstract

Evidence for bark, stem and stomatal density adaptation to different climates in the model species *Populus* is seen in both the natural population as well as in the greenhouse, but the genetic basis of these adaptation remains poorly understood. The present thesis investigates bark texture, bark thickness, diameter and stomatal density variations among *Populus* population using two quantitative genetics methods to attempt understand the genetic system controlling inheritance of these traits and to associate them with respective genes. The first approach aimed at detecting quantitative trait loci (QTL) associated with all phenotypic traits in an interspecific hybrid pedigree (*Populus trichocarpa* x *P. deltoides* and *P. deltoides*) collected across several years and sites. Identifying QTL for these important traits will pinpoint to polymorphisms linked to functional genes and provide a list of candidate genes and a baseline for future work. QTL specific to bark texture were highly reproducible in shared intervals across sites, years and replicates. Furthermore, significant positive correlations and co-localization between traits QTL suggest pleiotropic regulators or closely linked genes. Since bark texture showed promising results using the low-resolution QTL method, it was then analyzed further using high-resolution population genomic approach, the genome-wide association study (GWAS). Association mapping defined the genomic regions associated with natural variation in bark texture in a clonally replicated provenance trials of *P. trichocarpa* across three sites, multiple years. The association mapping used with a high-density SNP array allowed us to detect narrow genomic intervals (1-20 kb) with high reproducibility and shared candidate genes for bark texture variation. Several candidate genes were identified related putative function and their location close to QTL maxima were highlighted and are worth further investigation using functional genomics or forward genetics approaches. The results should be exploited for the future conservation and breeding of *Populus* species.

Chapter 1: General Introduction

The genus *Populus*, belonging to the family of Salicaceae, contains 30 to 35 species recognized by taxonomists [1] are forest trees (poplars, cottonwoods, aspens) with tremendous economic and ecological value, as well as valuable characteristics of basic scientific interest. Due to the inherent difficulties of studying such very large long-lived organisms, a model system for forest trees was developed. Poplars have several advantages as a model tree, including fast growth, ease of propagation and of cloning, propensity to hybridize, a small genome and many more [2]. Furthermore, their multiple beneficial uses to society are countless by providing wood, pulp and fiber, contributing to a sustainable development, as well as restoration of forest land and mitigation of climate change.

The spread of fast-growing poplar hybrid plantations across the world was the results of the increasing worldwide demand for wood products as well as the rising world population growth. Breeding programs were established in several countries including Belgium, France, Netherlands, Germany, UK, Canada and the USA [3, 4] to improve poplars for growth, wood products, bioenergy, biomass and environmental enhancement. In China alone, 20 million ha of poplar and willows tree plantings were established by 1978 known as the "Great Green Wall" in efforts to compensate for the many cut forests, resort vegetation and to protect the cities from dust and sand storms [5]. While in the western USA, since the early 1980s private industry planted large blocks of hybrid poplars totaling 20,000 ha [6].

The common cottonwoods of North America (*Populus deltoides* Bartr. ex Marsh *and P. trichocarpa* Torr. & Gray) and the black poplar of Europe *(P. nigra* L.) are the most important species for poplar breeding programs worldwide [7]. *P. deltoides* naturally thrives in the eastern, southern and mid-western USA and southern Canada and has been regarded as the fastest-growing native tree in North America [8], while *P. trichocarpa* occurs in Pacific coastal areas from California to Alaska in the northwestern North America [8]. The hybrid *P. trichocarpa* × *P. deltoides* shows hybrid vigour (heterosis) with many improved qualities including rapid growth rate, excellent form, site adaptability, ease of propagation and genetic pliability [7, 8]. The knowledge from the pioneering poplar genetic and silvicultural work on *P. trichocarpa* × *P. deltoides* hybrids

along with the recent release of genomic resources and the *P. trichocarpa* genome [9] will further support the identification of the underlying genetics involved in complex traits important for plants breeding.

In forest genetics, poplars have been adopted by tree physiologists as a model system due to this rapid growth, ease for clonal propagation and strong heterosis in interspecific hybridization [2, 10, 11]. Furthermore, it offers abundant genetic variation in tree morphology, anatomy and physiology where poplars are known to their plasticity and adaptations to varied environments and specific conditions [12]. For these treasons and many more, *Populus* is the center of forest genetics studies from molecular to ecological studies to investigation of large-scale systems to improve our understanding of tree growth and development to use in protection of forest ecosystems and practical tree breeding.

The traditional tree improvement for the selection of trees with imperative traits of economic importance an impractically slow and expensive process due to the long generation interval and since selection is based on physical traits [2, 11, 12] and since most poplars will not flower earlier than 4 years old. Additionally, classical genetic tools such as inbred lines cannot be produced rapidly enough due to the inability of self-pollination. Consequently, we need a sustainable program for this long-lived tree with basic knowledge of its genetic infrastructure including breeding and maintenance of living clone banks. Interspecific hybridization to produce large progeny arrays and multi-generation pedigrees are the basis for genetic mapping experiments designed to identify important genomic regions and the underlying genetics affecting important tree phenotypes [2, 10, 11]. Having the whole-genome sequence of *P. trichocarpa* readily available and the use of genetic markers in the naturally occurring molecular diversity provided an excellent resource for population and quantitative genetics and genomics of the genus.

Molecular markers are specific DNA sequences especially useful in designing breeding strategies to characterize and evaluate genetic diversity in a population, specifically aiding in maximizing genetic gain and monitoring the efficiency of tree improvement traits [13-15]. They are of significant value in multiple trait breeding and are used more efficiently in applied plant breeding, such as breeding for two traits found

in two separate parents are chosen to produce offspring with both desired traits [14]. The discovery of gene linkage, genes on the same chromosome, led to the establishment of linkage maps further aided by the development of DNA molecular marker technology [16, 17]. DNA markers are known to be valuable tools for crop improvements in rice, wheat, maize, barely and many more plant species [18-21]. As for the genus *Populus*, several types of genetic markers and genetic maps have been developed [7, 22, 23].

Genetic mapping, either QTL (Quantitative Trait Loci) or GWAS (Genome Wide Association Study) is based on the fact that during meiosis genes and markers segregate via chromosome recombination; thus, genes/markers that are close together (linked) will be transmitted together from parent to progeny more frequently. This leads to the determination of the recombinant frequency and converting them into centi Morgans (cM) as a genetic map unit [24-26]. Once the population is selected, either F_2 recombinant inbred lines or a naturally occurring population, the identification of DNA markers, polymorphisms, between individual trees is critical for constructing a linkage map. The final step involves marker scoring and coding on individual tree within a population and conducting linkage analysis using computer programs and creating a linkage map [24-26]. It is used to analyze and detect an association between the trait of interest and the genotype of markers pointing to the underlying genes responsible for the desirable trait.

The conventional QTL mapping is based on the association between genotype and phenotype in a F_2 population, derived by backcrossing the F_1 hybrid to one of the parents [25]. QTL mapping has high statistical power for detecting a QTL, however, it provides low resolution. Furthermore, only allelic diversity that segregates between the parents of the particular F_2 cross can be assayed; thus, we are limited to the genetic diversity present in the parents of our segregating population [25]. The second type GWAS or association mapping (also named linkage disequilibrium mapping) is based on identifying genes from unorganized natural populations. The low levels of linkage disequilibrium (LD) in natural populations of *Populus* makes candidate gene association studies very accurate since polymorphisms are expected to be within <1 kb of the SNP (single nucleotide polymorphisms) used to detect it [12]. Association mapping offers a very fine resolution, to the base pair variation, thus overcoming the limitations found in QTL mapping

technique. Nonetheless, it has its own drawbacks, such as losing power for detection for any rare alleles since the power for detecting a loci using GWAS will be determined by the frequency of alleles [27, 28]. Consequently, GWAS are often complementary to QTL mapping when conducted together they mitigate each other's limitations.

Goals and Objectives

The primary goal of this study was to identify allelic effects of genes controlling economically important traits, including bark and stomatal density, using two population genomics methods, the conventional QTL mapping and association mapping (GWAS). The main objectives of the present study were:

1. Study the genetic variation among different clones and populations of *Populus* across replicates and years.

2. Identify positional candidate genes that underlie QTL for bark texture, bark thickness, diameter growth and stomatal density.

3. Test reproducibility and consistency of QTL across years and environments.

4. Comparison between GWAS mapping with QTL mapping results and quantification of the number of candidate genes within the QTL intervals.

5. Identification of the differences in the genetic control of bark texture in *P. trichocarpa* and *P. deltoides*.

References

1.   Eckenwalder JE: **Systematics and evolution of *Populus*.** *Biology of Populus and its Implications for Management and Conservation* 1996, **7**:30.
2.   Bradshaw H, Ceulemans R, Davis J, Stettler R: **Emerging model systems in plant biology: poplar (*Populus*) as a model forest tree.** *Journal of Plant Growth Regulation* 2000, **19**(3):306-313.
3.   Riemenschneider DE, Stanton BJ, Vallée G, Périnet P: **Poplar breeding strategies.** *Poplar Culture in North America* 2001(Part A):43-76.
4.   Stettler R, Bradshaw T, Heilman P, Hinckley T: **Biology of *Populus* and its implications for management and conservation.**: NRC Research Press; 1996.
5.   Veste M, Gao J, Sun B, Breckle S-W: **The Green Great Wall–combating desertification in China.** *Geographische Rundundschan* 2006, **2**:14-20.
6.   Dickmann DI, Isebrands JG, Eckenwalder JE, Richardson J: **Poplar culture in north America.**: NRC Research Press; 2002.
7.   Cervera M-T, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, Van Slycken J, Van Montagu M, Boerjan W: **Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers.** *Genetics* 2001, **158**(2):787-809.
8.   Isebrands JG, Richardson J: **Poplars and willows: trees for society and the environment.**: CABI; 2014.
9.   Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A: **The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray).** *Science* 2006, **313**(5793):1596-1604.
10. Wullschleger SD, Jansson S, Taylor G: **Genomics and forest biology: *Populus* emerges as the perennial favorite.** In.: Am Soc Plant Biol; 2002.
11. Taylor G: ***Populus*: *Arabidopsis* for forestry. Do we need a model tree?** *Annals of Botany* 2002, **90**(6):681-689.
12. DiFazio SP, Slavov GT, Joshi CP: ***Populus*: a premier pioneer system for plant genomics.** In: *Joshi C, DiFazio SP, Kole C, editors Genetics, genomics and breeding of poplar Enfield, NH: Science Publishers.* 2011: 1-28.
13. White TL, Adams WT, Neale DB: **Forest genetics.**: Cabi; 2007.
14. Liu BH: **Statistical genomics: linkage, mapping, and QTL analysis.** Boca Raton, FA: CRC press; 2017.
15. Bradshaw H, Stettler RF: **Molecular genetics of growth and development in *Populus*. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree.** *Genetics* 1995, **139**(2):963-973.
16. Sturtevant AH: **History of genetics.**: Harper and Row; New York And London; 1965.
17. Botstein D, White RL, Skolnick M, Davis RW: **Construction of a genetic linkage map in man using restriction fragment length polymorphisms.** *American Journal of Human Genetics* 1980, **32**(3):314.
18. Kumar LS: **DNA markers in plant improvement: an overview.** *Biotechnology Advances* 1999, **17**(2-3):143-182.

19. Blair M, Panaud O, McCouch S: **Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.).** *Theoretical and Applied Genetics* 1999, **98**(5):780-792.
20. McLauchlan A, Ogbonnaya FC, Hollingsworth B, Carter M, Gale K, Henry RJ, Holton TA, Morell MK, Rampling L, Sharp PJ: **Development of robust PCR-based DNA markers for each homoeo-allele of granule-bound starch synthase and their application in wheat breeding programs.** *Australian Journal of Agricultural Research* 2001, **52**(12):1409-1416.
21. Andersen JR, Lübberstedt T: **Functional markers in plants.** *Trends in Plant Science* 2003, **8**(11):554-560.
22. Bradshaw H, Villar M, Watson B, Otto K, Stewart S, Stettler R: **Molecular genetics of growth and development in Populus. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers**. *Theoretical and Applied Genetics* 1994, **89**(2-3):167-178.
23. Rahman MH, Dayanandan S, Rajora OP: **Microsatellite DNA markers in Populus tremuloides**. *Genome* 2000, **43**(2):293-297.
24. Jones N, Ougham H, Thomas H: **Markers and mapping: we are all geneticists now.** *The New Phytologist* 1997, **137**(1):165-177.
25. Collard B, Jahufer M, Brouwer J, Pang E: **An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts.** *Euphytica* 2005, **142**(1-2):169-196.
26. Haley CS, Andersson L: **Linkage mapping of quantitative trait loci in plants and animals.** *Genome mapping: a practical approach* 1997.
27. Korte A, Farlow A: **The advantages and limitations of trait analysis with GWAS: a review.** *Plant Methods* 2013, **9**(1):29.
28. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN: **Genome-wide association studies for complex traits: consensus, uncertainty and challenges.** *Nature Reviews Genetics* 2008, **9**(5):356.

Chapter 2[1]: Quantitative Trait Locus mapping of *Populus* bark features and stem diameter

Abstract

Bark plays important roles in photosynthate transport and storage, along with physical and chemical protection. Bark texture varies extensively among species, from smooth to fissured to deeply furrowed, but its genetic control is unknown. This study sought to determine the main genomic regions associated with natural variation in bark features and stem diameter. Quantitative trait loci (QTL) were mapped using an interspecific pseudo-backcross pedigree (*Populus trichocarpa* x *P. deltoides* and *P. deltoides*) for bark texture, bark thickness and diameter collected across three years, two sites and three biological replicates per site. QTL specific to bark texture were highly reproducible in shared intervals across sites, years and replicates. Significant positive correlations and co-localization between traits QTL suggest pleiotropic regulators or closely linked genes. A list of candidate genes with related putative function, location close to QTL maxima and with the highest expression level in the phloem, xylem and cambium was identified. Candidate genes for bark texture included an ortholog of *Arabidopsis* ANAC104 (PopNAC128), which plays a role in lignified fiber cell and ray development, as well as Pinin and Fasciclin (PopFLA) genes with a role in cell adhesion, cell shape and migration. The results presented in this study provide a basis for future genomic characterization of genes found within the QTL for bark texture, bark thickness and diameter in order to better understand stem and bark development in *Populus* and other woody perennial plants. The QTL mapping approach identified a list of prime candidate genes for further validation using functional genomics or forward genetics approaches.

Introduction

Bark, the outermost surface of stems and branches in woody plants, encompasses all tissues outside the vascular cambium and includes the secondary phloem, secondary cortex and the periderm [1, 2]. Bark's outer layer, or phellem, is composed of mostly dead tissues that form a protective barrier between plant and the abiotic and biotic environment, while the inner layer, or phloem, serves as a conduit for transport and storage of photosynthate [1, 2]. Despite its important roles including photosynthate transport [3], photosynthesis [4, 5], storage [6], mechanical support [7] and protection [8-11], the molecular basis of bark formation remains poorly understood (for reviews see: [12, 13]).

Bark texture varies among species, and even among genotypes within species, and has notable phenotypic diversity ranging from smooth, peeling, fractured, fissured to plated [14]. Within genera, bark texture differs between related species, e.g., in mature *Populus trichocarpa* (Torr. & Gray), bark is smooth or lightly flaky, while in *P. deltoides* (Bartr. ex Marsh), bark is rough and highly furrowed [15]. Bark's high morphological diversity suggests that variation in texture may be an important component of variation in plant ecological strategies. It has been reported that in ash and beech smooth bark genotypes are less susceptible to insect and fungal diseases [16, 17]. In addition, bark thickness and moisture content are correlated with enhanced fire resistance [9, 18, 19] and in cork oak the phellem is also the basis of the cork manufacturing industry [20]. Despite the biological, ecological and industrial value of bark, the genetic basis of bark's features remains undefined.

To further understand the variation in bark texture, we need a better understanding of outer bark development. The outer bark includes all tissues formed by the phellogen, consisting of dead hollow cork cells [1, 2], and originating from the outermost layer of the secondary phloem [2]. Romero [21] has proposed that discontinuous periderms may be the result of variation in radial meristematic activity in the phellogen in apparent response to the mechanical stresses imposed by radial growth, whereas, smooth textured barks may be derived from the formation of a single periderm and continuous shedding of phellem. However, since most plant species develop several periderms over the course of time, smooth bark scales can develop from preceding

17

periderms from beneath the stem surface while uneven thick and thin layers result in the bark splitting, and in a peeling bark appearance [21]. Finally, Romero [21] also suggests that scaly and fissured bark develops when bark growth is discontinued and overlapping layers of periderms are formed. While these descriptions provide indications on how bark texture can vary, there remains a lack of understanding of phellem development at the molecular level. Additionally, variation in radial meristematic activity in the cambium may affect both diameter growth and bark texture as result of mechanical stresses. Especially Lateral Organ Boundaries (LBD) genes were found to be important regulators of woody perennial growth in poplar [22]. Specifically, the two LBD genes, PtaLBD1 and PtaLBD4 are expressed at the cambium/phloem boundary and are involved in ray cell and secondary phloem development. Two other LBD genes, PtaLBD15 and PtaLBD18, are expressed at the cambium/xylem boundary and are involved in secondary xylem development.

Quantitative Trait Locus (QTL) mapping in segregating populations is a powerful tool to 1) uncover genes underlying naturally occurring phenotypic variation and 2) dissect the genetic basis of phenotypic traits [23]. QTL-based approaches have often been implemented to study the complex genetic architecture underlying wood formation, including lignin, diameter, height, biomass and various wood chemistry traits [24-31]. However, only a few studies have explored bark features and are limited to bark thickness in *Eucalyptus globulus* [32], *Pinus* hybrids [33] and *Boehmeria nivea* [34]. In *Populus*, QTL mapping and gene expression analyses were used to link sequence polymorphisms and variation in transcript levels [35-37].

Thus, in this study, we investigate bark texture, bark thickness and diameter variation in the pseudo-backcross Family 52-124 derived from a cross between a *P. trichocarpa* x *P. deltoides* hybrid and *P. deltoides* [25, 38]. Novaes et al. [25] performed a QTL mapping study in the same mapping pedigree for 20 biomass and wood chemistry traits (including stem diameter) under different nitrogen treatments and identified a total of 63 QTL distributed across 14 chromosomes. In the present study, QTL mapping was done using a genetic map with high marker density anchored to the *P. trichocarpa* whole-genome assembly [39], which allowed us to determine the map position of QTL and identify underlying candidate genes. Specifically, we performed QTL analyses for bark

texture, bark thickness and stem diameter and report results across three years and two geographic locations. These three traits have been chosen since we suspect an interconnection between their development. Specific objectives were to: 1) identify positional candidate genes that underlie QTL for bark texture, bark thickness and diameter growth and 2) test reproducibility and consistency of QTL across years and environments.

Materials and methods

Mapping population

An interspecific hybrid poplar pseudo-backcross pedigree (Family 52-124) composed of 396 genotypes was created by crossing the hybrid female clone 52-225 (TD), an F_1 hybrid derived from *P. trichocarpa* (TT, clone 93-968) × *P. deltoides* (DD, clone ILL-101), with *P. deltoides* (DD, clone D124) (Fig. S1a) [25]. The F_1 hybrid, clone 52-225, had smooth bark and was crossed with *P. deltoides* clone D124 with rough bark. The mapping population was planted at Boardman, OR (45°50'8"N, 119°33'48"W) in 2010 with two replicates of 396 genotypes in a three-block replication for a total of six ramets per cloned genotype (Fig. S1b). The same pedigree was planted in Morgantown, WV (39°39′32″N 79°54′19″W), in 2006 with four replicates of the 396 genotypes used in this study.

Plant material and construction of genetic linkage map

We used the genetic map of the mapping family 52-124 comprised of 3,568 SNP markers with known genomic positions for QTL identification. SNP genotyping, marker curation and genetic map construction were previously described by Muchero *et al.* [31].

Phenotypic measurements

Phenotypic data for bark texture (BT), bark thickness (BTh) and diameter at breast height (D) for all 396 full-sibs were analyzed in this study. Specifically, for the Oregon site (OR), bark texture data were collected in year 3, 4 and 5 by visual inspection, whereas diameter and bark thickness data were collected only in year 3 by using diameter tape and a bark thickness gauge on two opposites sides of the stem. For the West Virginia site (WV), bark texture data were collected in year 4 and 6 by visual inspection; diameter

and bark thickness traits were collected only for year 4 using a caliper and ruler. Bark texture was assigned a qualitative score based on a scale from 1 (smooth) to 4 (furrowed with deep grooves) (Fig. 1). Especially, the replicated multi-year measurements allowed for a reliable identification of QTL in Boardman, Oregon. Some QTL in Morgantown, Virginia, could have remained undetected as result of the lower number of replicates and years.

Pair-wise phenotypic correlations were calculated as Pearson correlation coefficients for all three traits across the two different geographic regions and the three years using WinSTAT software [40] to assess covariance within blocks, and within and between years (Table S1). To assess variation within blocks, the two replicates for each of three blocks, named as 1a, 1b, 2a, 2b, 3a & 3b, respectively, were analyzed at the OR site in year 3. In year 4 and 5, only block 1 and 2 were measured.

QTL analysis

The data were evaluated for the presence of outliers and recording errors were corrected or deleted. WinSTAT [40] was used to check for normal distribution of residuals. Transformations were deemed unnecessary (Fig. S2). The traits were analyzed with the interval mapping method implemented in MapQTL6 [41] to map putative and suggestive QTL intervals on the genetic linkage map and to test for reproducibility across years and environments. Composite interval mapping with the Multiple-QTL Model (MQM) method was then used to further refine the QTL regions. Markers closely linked to a putative QTL were selected as cofactors and the selected markers were used as genetic background controls in subsequent MQM mapping. We selected additional cofactors until no additional QTL was detected. Mean phenotypic values across the two replicates per site were analyzed separately for each of the three blocks across three years for the OR site and across two years for the WV site. The putative QTL were subjected to 1,000 genome-wide (GW) and chromosome-wide (CW) permutation tests [42] to determine LOD significance thresholds at the 0.05 significance level (Table S2). A putative QTL was declared when it was detected in at least two replicates or in one replicate in different years or sites, with at least one of those instances exceeding the chromosome-wide LOD threshold. To account for minor deviations from normality in

20

some cases, we also performed the non-parametric Kruskal-Wallis test which is the equivalent to the one-way analysis of variance [43].

Candidate genes

Genes underlying genome-anchored QTL intervals (Table S3, S4) were identified from the *Populus* genome assembly V3.0 [44] in the Phytozome database via BioMart tool (https://phytozome.jgi.doe.gov). A complete gene list with InterPro descriptions was collected including both putative and unknown functions. Expression profiles of the gene models from various tissues: bud, leaf, various parts of root and stem (expression FPKM), were downloaded from the publicly available PhytoMine database (https://phytozome.jgi.doe.gov/phytomine/begin.do). The Affymetrix microarray expression raw date profiles for the traits bark and mature phloem, developing phloem, cambium, developing cambium and mature cambium were obtained from the NCBI, GEO database (https://www.ncbi.nlm.nih.gov/geo/, GEO accession number GSE30507) [45]. The raw data were normalized using the RMA algorithm [46] and further analyzed statistically using TM4:MeV software [47, 48], utilizing Affymetrix probe annotation [49].

The genes' expression in developing and mature phloem/xylem and cambium was then assessed for each QTL interval for all traits based on publicly available data. For each QTL cluster, genes in the map interval with the highest LOD score and high expression in phloem, cambium and xylem tissues (above the 90th percentile) were compiled in a list (Table S5).

Position of LBD genes

The position of the Lateral Organ Boundaries Domain (LBD) genes with putative role in bark development and diameter growth were identified by using the BLAST tool in the *Populus* genome assembly V3.0 in the Phytozome (https://phytozome.jgi.doe.gov/) database against well-established *Arabidopsis thaliana* LBD genes.

Results

Analysis of phenotypic correlations among traits and trait frequency distributions

Bark phenotypes ranged from smooth (1) to deeply furrowed bark (4) (Fig. 1). Shallowly fissured bark typical for *P. trichocarpa* was not found in this backcross pedigree (*P. deltoides* x *P. trichocarpa* hybrid backcrossed with *P. deltoides*). The interspecific crossing parent, *P. trichocarpa* (clone 93-968) x *P. deltoides* (clone ILL-101), had relatively smooth bark (mean value: 1.33, SD: 0.47). The other crossing parent, *P. deltoides* clone D124, had a rough bark texture (mean: 2.66, SD: 0.47). The grandparent *P. deltoides* (DD, clone ILL-101) had a slightly furrowed bark texture while grandparent *P. trichocarpa* (TT, clone 93-968) had smooth and slightly fissured bark (field observations on adult trees, no genotypes and measurements available in the field trials). Bark texture showed the highest correlation within blocks (r=0.91 to 0.93, p≤0.0001). The phenotypic correlations at the OR site within the same year for bark texture ranged from r=0.58 to 0.76; for diameter, r=0.38 to 0.45 and for bark thickness, r=0.40 to 0.56, all at p≤0.0001 (Table S1). Comparing mean values of traits among years, bark texture values were significantly correlated among years at the OR site (r=0.51 to 0.77, p≤0.0001). The correlations were weaker for the WV site, but still highly significant (r=0.39, p≤0.0001). Finally, at year 3, bark thickness showed a strong positive correlation with both bark texture (r=0.32 to 0.69, mean value r=0.49, p≤0.0001) and diameter (r=0.17 to 0.75, mean value r=0.43, p≤0.0001) within the OR site (Table S1, blue and green sections), however, bark texture and diameter showed inconsistency in correlation values and significance ranging from r=0.15 (p≤0.05) to r=0.47 (p≤0.0001) (Table S1, red section). Overall, traits showed high correlations among replicates and years, and traits were correlated with each other. Across site correlations were only significant for bark texture ranging from r=0.25 to 0.40 (p≤0.01-0.0001).

QTL analysis and detection across contrasting environments

Seven major QTL clusters were detected for bark texture on seven individual chromosomes I, II, VI, VIII, XIII and XVIII (Table S2), with all clusters containing at least three individual QTL above the GW threshold. For diameter, three QTL clusters with significance above the GW threshold were detected on chromosome I, VI and

XVIII; in addition, two suggestive QTL above the CW threshold were detected on chromosome VIII and XII (Table S2). Bark thickness showed three QTL clusters above the GW threshold on chromosome I, VI and XVIII, and four QTL above the CW threshold on chromosome II, VIII and XII (Table S2). Chromosome VIII likely contains two separate QTL since they map to distinct chromosomal positions. All 94 individual QTL detected for the three traits across various chromosomes were successfully anchored to the *Populus* genome assembly (Fig. 2; Table S2). For the seven bark texture QTL clusters, the percentage of phenotypic variance explained (PVE) ranged from 3.6 to 12.8% for QTL above the GW threshold, while for diameter and bark thickness, it ranged from 5.4 to 8.4% and 4.5 to 9.6%, respectively (Table S2). For QTL on chromosome II, VI, VIII and XII the *deltoides* genotype DD was associated with the lower value for bark texture (Table S2), while for QTL on chromosome I, XIII and XVIII the DT genotype was associated with a lower value for bark texture.

Figure 3 shows a graphical outline of LOD score profiles for bark texture QTL versus map location across all seven chromosomes before (left) and after cofactor selection (right). The QTL for bark thickness and diameter, described above, overlap with six out of the seven bark texture QTL clusters using the interval mapping approach (Fig. 2). Specifically, QTL for bark texture overlap with diameter and bark thickness QTL on chromosome I, VI and XII, and solely with bark thickness QTL on chromosome II, VIII and XVII (Fig. 2). Overall, reproducibility and co-location within the same map interval are observed across experimental replicates within sites and years and in some cases across sites.

Description of QTL clusters

Based on significance, consistency and reproducibility of the QTL across sites, years and replicates, specifically for bark texture, we classified the QTL clusters according to four criteria: 1) significance (LOD scores), 2) reproducibility across biological replicates, 3) reproducibility over time (years) and 4) reproducibility across environments (sites). All seven QTL clusters were significant with at least three individual QTL having LOD scores above the GW threshold and were reproducible across blocks (biological replicates) within the same year and across two years for the OR

site. Four QTL clusters on chromosome I, II, VIII and XIII, were reproducible between sites across very different environments. In each QTL cluster most of the QTL were associated with bark texture. Below we provide a detailed description of all seven bark texture QTL clusters ranked according to the four criteria and their association with bark texture, bark thickness and diameter.

QTL cluster on chromosome I was associated with all three traits (Fig. 2). For diameter four out the five QTL were significant at the GW threshold and for bark thickness three out of the four QTL were above the GW threshold (Table S2). For bark texture, twelve QTL were detected across all replicates and years in both OR and WV and ranged from 101 to 192 cM (before cofactor selection) with highly reproducible LOD maxima after cofactor selection consistently around 165 cM (Table S2; Fig. 3a). For the OR site, seven out of the ten QTL in this cluster were above the GW threshold all being detected in years 3 and 5; whereas at the WV site, the two QTL for both years were above the CW threshold. All twelve individual QTL mapped reproducibly to the same map interval and LOD maxima positions at the GW threshold were typically associated with no more than three markers in close proximity, around 3 cM (Fig. 3a). Notably, the results showed consistency and high reproducibility, first between blocks within individual years, second within the same site, and third across sites. Overlapping QTL intervals and maxima with high significance across all years are presented in Table S2 and in Figures 2, 3.

The second bark texture QTL cluster was detected across all replicates and years in OR, and in year 4 in WV, and was mapped on chromosome VIII within the chromosomal region 48 to 104 cM before cofactor selection with varied LOD maxima positions after cofactor selection (~50 cM, ~65 cM or 74 cM) (Table S2; Fig. 3a). When comparing this cluster with other traits, one overlapping QTL for bark thickness (significant at the CW level), peaking at 67 cM was detected. For the OR site, six of the ten QTL for bark texture in this cluster were significant above the GW threshold; whereas at the WV site, only one QTL was found above the CW threshold. All eleven individual QTL within this cluster mapped reproducibly within the same map interval, however positions of QTL maxima varied for the OR site (Fig. 3a). The overall reproducibility

within the OR site, across both sites and across all years, was high and QTL were found within the same chromosomal region covering 25 cM.

The third bark texture QTL cluster mapped on chromosome XIII and was detected across all years and replicates in OR, and in year 4 in WV, within the chromosomal region 83 to 118 cM before cofactor selection and reproducible LOD maxima after cofactor selection was found around 108 cM (Table S2; Fig. 3a), which interestingly co-located with the Lateral Organ Boundaries Domain gene LBD15c (Potri.013G156200), a candidate gene for xylem development [22]. No QTL for bark thickness or diameter were found in this QTL interval. For the OR site, seven out of the nine QTL in this cluster were significant above the GW threshold across all years; whereas at the WV site, only one QTL was detected and found above the GW threshold. All ten individual QTL mapped reproducibly within the same map interval and positions of QTL maxima were typically associated with no more than four markers in close proximity covering a map interval around 5 cM (Fig. 3a). The results showed overall consistency and reproducibility between block replicates within individual years, within the OR site and across the two sites and across most years (Table S2, Fig. 3a).

The fourth bark texture QTL cluster on chromosome II was detected in both OR and WV, across all years and replicates for both sites, within a chromosomal interval from 0 to 70 cM (before cofactor selection) but displayed inconsistent and variable LOD maxima after cofactor selection (Table S2; Fig. 3b). When comparing this cluster with bark thickness and diameter, one overlapping QTL for bark thickness (significant at the CW threshold), peaking at 0 cM, was found. Eight out of the ten individual bark texture QTL were significant above the GW threshold, indicating reproducibility within replicates and sites, between sites and across years. However, after cofactor selection, the LOD maxima greatly varied even within replicates of the same year.

The fifth bark texture QTL cluster was detected on chromosome VI in OR in years 3 and 5 only, within a chromosomal interval from 135-204 cM (before cofactor selection) and had LOD maxima after cofactor selection between 153-162 cM (Table S2; Fig. 3b). Several overlapping QTL, four for diameter (two GW QTL) and two GW QTL for bark thickness, were found within the same chromosomal interval, however the LOD maxima varied among traits after cofactor selection. Six out of the seven individual QTL

for bark texture were above the GW threshold, indicating reproducibility within replicates at the OR site and across the two years. However, the results were not consistent across the two sites, as no QTL was detected at the WV site. Only three out of the nine individual QTL were found significant above the GW threshold, while others were significant at the CW threshold.

The sixth bark texture QTL cluster mapped on chromosome XII and was detected across all years and replicates in OR within the chromosomal interval from 62-113 cM before cofactor selection and consistent LOD maxima between 98-104 cM were detected after cofactor selection at the OR site (Table S2; Fig. 3b). Two individual QTL, one for bark thickness (CW QTL) and one for diameter (CW), were partially overlapping with bark texture QTL, but had LOD maxima separate from the LOD maxima for bark texture. Out of the nine individual QTL, only three were above the GW threshold. The results are reproducible between replicates and within and across the three years in OR. Though no significant QTL were detected in WV, the QTL graph shows a suggestive QTL with increasing LOD score near the same chromosomal interval noted above (Fig. 3b), though still below the CW threshold.

Finally, the seventh bark texture QTL cluster, mapped on chromosome XVIII within the chromosomal interval from 13-35 cM, before cofactor selection, with reproducible LOD maxima at 30 cM after cofactor selection, detected in year 3 and 5 at the OR site (Table S2 and Fig. 3b). A QTL cluster for bark thickness closely overlapped with the same chromosomal interval with LOD maxima further upstream. Four separate QTL for diameter were detected on the same chromosome (two above the GW threshold). In year 5, three out of the seven bark texture QTL were above the GW threshold. The QTL were only reproducible across year 3 and 5 at the OR site, although a suggestive QTL was found for WV (Fig. 3b).

The Kruskal-Wallis rank sum test was subsequently used to confirm significant associations of individual markers linked to the QTL. For all traits, the markers underlying the QTL interval were also significantly associated with the traits ($p \leq 0.005$). The Kruskal-Wallis test provides further confirmation of the marker-trait association, indicating that the results of the QTL analysis were not influenced by segregation distortion or non-normal distribution of certain traits.

Candidate gene identification and characterization

To narrow the QTL position and identify candidate genes within the QTL interval MQM mapping was used. Intervals spanning the genomic regions (physical location by MQM mapping) summarized in Table S2 were used to identify all genes occurring within the seven QTL clusters for bark texture and for the QTL clusters that were associated with bark texture, diameter and bark thickness. The number of genes for each trait in QTL clusters based on MQM mapping with cofactor selection is summarized in Table S3. There were 1869 genes within genome-anchored QTL intervals for bark texture, out of which, 1476 (82%) had annotations based on the InterPro domain and expression profiles (average FPKM) for 22 different tissues and based on Affymetrix microarray expression data for bark and woody tissues (phloem, cambium and xylem), whereas a total of 693 and 789 genes were detected in QTL clusters for diameter and bark thickness, respectively (Table S4).

QTL for all three traits overlap in QTL cluster on chromosome I, where the diameter QTL included 25 genes and the bark thickness QTL encompassed 115 genes. Two additional QTL clusters on chromosome VI and XII were associated with all three traits where diameter and bark thickness QTL had 369 and 209 overlapping genes on chromosome VI and 29 and 14 overlapping genes on chromosome XII, respectively. QTL clusters for bark texture on chromosome II, VIII and XVIII overlapped with bark thickness QTL only, containing 963, 23 and 38 overlapping genes, respectively (Fig. 2, Table S4). As a result of the Salicoid duplication event in the *Populus* genome, nearly every chromosome has a paralogous segment elsewhere in the genome; this is due to the whole-genome duplication between chromosomes resulting in homologous genomic blocks [50]. Each of the seven chromosomes noted above has a Salicoid paralog, yet none of these duplicated genes in paralogous segments co-located with other QTL for the same trait.

Top expressional candidate genes (above the 90th percentile) for mature and developing phloem, cambium, developing xylem and mature xylem in QTL intervals with the highest LOD scores resulted in a compiled list of the top candidate genes for each trait (Table S5). In total, the top candidate genes with putative function in the

27

control of bark texture, diameter and bark thickness are narrowed down to 40, 20 and 46 genes spanning various QTL clusters (Table S5).

Discussion

We have characterized segregating bark features in an interspecific backcross of *Populus*. With the use of QTL mapping, we are able to link the phenotypic traits to their associated polymorphisms in the genome, thus integrating phenotypic and genotypic data to identify putative genetic mechanisms related to phellem development.

While other studies have identified QTL in interspecific *P. trichocarpa* and *P. deltoides* families for many different traits, including leaf size and shape, growth and bud set, diameter, height, stem and root biomass and various wood chemistry phenotypes [25, 35, 51-54], little research has been done on bark features [55-57], despite bark being one of the key-energy-related characteristics of lignocellulosic feedstock [58-60].

We identified several QTL that encompassed both bark traits and stem diameter and found that these intervals mapped consistently across geographic locations, replicates within sites and across years. Interestingly, one study, using the same pedigree, reported several overlapping QTL with our traits [25]. Specifically, bark texture QTL in our study overlapped on chromosome I, II and XVIII with QTL associated with total biomass, C5 and C6 sugars, and height. Additionally, diameter and bark thickness QTL overlapped on chromosome VIII, XII and XVIII with QTL for diameter and biomass traits [25]. It is difficult to determine if there are genes that have pleiotropic effects or whether there are alternate genes within the co-located intervals because of the large size of the interval and lack of expression evidence in the Novaes *et al.* [25] study.

While a few studies have analyzed bark thickness, e.g., *Eucalyptus globulus, Pinus* hybrids and *Boehmeria nivea* [32-34], the genetic basis and causal loci of bark thickness and/or bark texture have not yet been determined. In *Boehmeria nivea,* a perennial herbaceous plant belonging to the Urticaceae Family, several QTL for bark thickness have been mapped and some were identified in the same QTL intervals across two contrasting environments in Changsha, China at varying time throughout the year [34].

Bdeir *et al.* (2016) and Yordanov *et al.* (2010) previously identified genes with a role in bark development [61, 22]. Based on the generation of loss-of-function phenotypes through transgenic plants, Lateral Organ Boundaries Domain (LBD) genes were found to have a crucial role in meristem maintenance and were identified as important regulators of woody perennial growth in poplar, specifically in *Populus tremula* x *Populus alba* (Pta). The overexpression of PtaLBD1 resulted in wide multiseriate rays as compared to uniseriate rays in the wild type [22]. In the regulation of secondary (woody) growth, two genes of the LBD Family (PtaLBD1 and PtaLBD4) were involved in secondary phloem and ray cell development and two genes (PtaLBD15 and PtaLBD18) in secondary xylem formation. Interestingly, one of the PtLBD15 paralogs (Potri.013G156200, total expression 36.30), previously found to have a role in secondary growth, was found in the QTL cluster on chromosome XIII for bark texture with the highest LOD score. In this chromosome region seven out of a total of 10 QTL detected were above the genome-wide threshold (Fig. 2). PtaLBD1 (Potri.008G043900) was detected in the diameter QTL cluster on chromosome VIII with a LOD score of 2.27. Both PtaLBD15 and PtaLBD1 are involved in secondary growth in poplar and are potential candidate genes for diameter growth and bark characteristics. PtaLBD15 was found in *Populus tremula* x *P. alba* to be mainly expressed at the cambium/xylem boundary and thus is likely involved in secondary xylem development. PtaLBD1 was found to regulate secondary phloem and ray development and was highly expressed in the phloem and cambial zone [22]. Therefore, their apparent involvement in secondary growth and development in poplar and their detection in QTL clusters for diameter and bark texture make them candidate genes for these traits. Finally, PtaLBD12 (Potri.008G072800, total expression 26.06) was detected in the QTL cluster on chromosome VIII and overlapped with diameter and bark thickness QTL. PtaLBD12 has been reported to be involved in the development of various lateral organs from the meristem in *Arabidopsis* plants, but its role in secondary growth is unknown [62].

We were able to identify a list of candidate genes underlying the QTL intervals of all three traits using genetic markers anchored to the *Populus trichocarpa* genome. Generally, QTL were highly reproducible among biological replicates and years and even across geographic locations. While some QTL studies obtained good reproducibility

across two different time data set [33, 34, 63], our study has identified significant QTL consistently co-locating across sites, years and replicates especially for bark texture. For instance, QTL in clusters on chromosome I, VIII, XII, XIII and XVIII were consistently identified within ~25 cM.

When comparing all the aspects of the seven bark texture QTL clusters, including significance, reproducibility across replicates, years and sites, along with consistency of the QTL, major QTL were identified on chromosome I, VIII and XIII, where QTL maxima were found within a 5-20 cM interval across most replicates, years and at both sites. Given the environmental contrast between the OR and WV experimental sites, four out of the seven QTL clusters, representing a total of 47 individual QTL, detected for bark texture were remarkably consistent across both sites. Differences in reproducibility for QTL clusters across sites suggest differential environmental effects on gene expression. In comparison to bark texture, QTL clusters for bark thickness and diameter had lower reproducibility across sites.

In QTL clusters on chromosome I, VI and XII, QTL for bark texture, bark thickness and diameter were syntenic. Co-location of QTL for traits can be the result of pleiotropic effects or closely linked genes. These overlapping QTL could be an explanation of different aspects of bark texture and radial growth. Romero [21] proposed that rough bark results in response to the mechanical stresses imposed by a varied radial growth and due to different meristematic activity in the phellogen, a discontinuous periderm. Strong correlation between bark texture and diameter could indicate that bark texture is partly related to diameter growth. Using interval mapping without cofactor selection, QTL for bark texture overlap with QTL for both traits which are related to radial growth (diameter and bark thickness) on three chromosomes (I, VI, XII) and solely with bark thickness on three other chromosomes (II, VIII, XVII). Furthermore, using MQM mapping, QTL for these traits were mapped to different neighboring positions of the same chromosomes (Fig. 2). Consequently, bark texture seems to be only partly related to diameter growth, and other factors such as meristematic activity of the phellogen and cell adhesion are likely to have major effects on bark texture. A higher mapping resolution as obtained in linkage disequilibrium mapping in natural population

samples is needed to narrow down the QTL region to individual genes and to distinguish between pleiotropic effects and close linkage.

Bark features in our study ranged from smooth to deeply furrowed which is characteristic for *P. deltoides*. Variation in shallowly fissured bark which is characteristic for *P. trichocarpa* was not observed among the segregating progeny. Thus, the QTL identified in this progeny set only represent a subset of a larger number of polymorphisms affecting the traits. And in a pseudo-backcross involving multiple *P. deltoides* parents, polymorphisms associated with characteristic bark features of *P. trichocarpa* seem to be largely undetected. Association populations for *P. trichocarpa* will be used to find additional candidate genes associated with bark texture in this species.

Each of the seven QTL clusters detected have a Salicoid paralog, and yet none of these paralogous genes showed up in the QTL analyses as significant. This further supports that the identified chromosomal regions are not artifacts of spurious correlations. Due to large genomic intervals in QTL clusters with partly overlapping QTL intervals the identification of specific candidate genes was difficult. This limitation was evident in our analyses where only two out of the seven QTL clusters on chromosome XIII and XVIII, encompassed less than 70 candidate genes, while the other clusters included from 123 to 963 candidate genes. Nonetheless, several candidate genes within the QTL interval can be identified based on their putative functions. Using the MQM method, we were able to identify informative loci for bark texture and narrow the QTL region to a small chromosomal region with a short and manageable candidate gene list. For example, a total of 11 NAC genes were detected in bark texture QTL clusters one, five, one, three and one paralogs found on chromosome I, II, VI, VIII and XII, respectively. Specifically, the gene PopNAC128 (Potri.001G206900) is a prime candidate gene and was identified in the QTL interval on LG I within the QTL maxima (LOD score 9.88) and with a moderate expression value. PopNAC128 is one of the orthologs of *Arabidopsis* ANAC104 (*Arabidopsis* Nac Domain Containing Protein 104) and XND1 (Xylem NAC Domain 1). In a related study, *Populus* and *Arabidopsis* transgenic plants with overexpression of these genes resulted in severe dwarfing, lacking phloem fibers and a reduction in stem diameter, cell size and number, vessel number, and frequency of rays in

the xylem [64]. While this study did not focus on bark texture, lack of sufficient lignified fiber cells in the mutant affects the development of fiber bundles and ultimately bark texture as result of slowed secondary phloem development.

Another interesting gene, Potri.001G206700, an ortholog of AT4G33430 (BAK1, Bri1-Associated Receptor Kinase; ELONGATED; SERK3), is involved in patterning and growth regulation [65-67] and was found in the QTL interval on chromosome I, also within the QTL maxima (LOD score 9.88) and with a very high expression value above the 90th expression percentile across phloem/xylem and cambium tissues.

Variation in bark texture could be related to cell adhesion, which is essential to form a single periderm resulting in smooth bark, while lack of cell-cell adhesion leads to the development of uneven and discontinued bark or bark splitting causes a peeling and fissured bark appearance [21]. At the molecular level, several QTL and expressional candidate genes with high expression in phloem/ xylem and cambium identified in this study have a role in cell adhesion, including Pinin (Potri.001G208200) and PopFLA or Fasciclin-Like Arabinogalactan (Potri.013G151300, Potri.013G151400 and Potri.013G151500). Interestingly, both genes fall within the QTL interval with the highest LOD scores and are above the 90th gene expression percentile for both xylem and phloem tissues. Many of the studies on Pinin, mainly on animal epithelial cells, revealed a vital role in cell-cell adhesion and cell shape [68, 69]. No studies exploring the Pinin gene in plants were found. The FLA gene is better studied across the plant kingdom, including *Arabidopsis* and *Populus*, and shows specific and high expression during the onset of secondary-wall cellulose synthesis, particularly in stem cells undergoing secondary-wall deposition [70, 71]. Transgenic lines indicate a role in cell-wall architecture and composition. Specifically for PopFLA, a role in tension wood formation in the xylem of mature stems was suggested based on a reduction in transcript levels leading to reduced stem flexural strength by modulation of cellulose and lignin composition in the xylem [70, 72, 73].

Future work

The top QTL and expressional genes reveal additional potential candidates (Table S5), some of which are proteins of unknown or putative function and have never been

studied. These genes represent potential candidate genes for future studies using either functional genomics or forward genetics techniques. Candidate genes within QTL intervals were identified based on Affymetrix Microarray expression profiles obtained from public databases. In the future, qRT-PCR confirmation of candidate genes' expression profiles should be performed in various tissues (phloem, cambium, xylem, phellogen) of the parental clones and part of the mapping pedigree.

Conclusion

In conclusion, the results presented in this study provide a basis for future genomic characterization of genes found within the QTL for bark texture, bark thickness and diameter in order to better understand stem and bark development in *Populus* and other woody perennial plants. Additionally, profiling the expression of the candidate genes (eQTL studies) in the developing bark of the mapping pedigree would allow parsing the list of candidate genes into those genes with high expression profiles in the tissue of interest. Bark texture is a complex trait which can be affected by differences in cell adhesion and radial meristematic growth. In the future, developmental differences between bark texture phenotypes should be analyzed in anatomical sections in representative genotypes and developmental stages.

References

1.   Carlquist S: **Wood, bark, and stem anatomy of Gnetales: a summary.** *International Journal of Plant Sciences* 1996, **157**:S58-S76.
2.   Srivastava LM: **Anatomy, chemistry and physiology of bark.** *International Review of Forestry Research* 1964, **1**:203-277.
3.   Jensen KH, Liesche J, Bohr T, Schulz A: **Universality of phloem transport in seed plants.** *Plant Cell Environ* 2012, **35**(6):1065-1076.
4.   Cernusak LA, Hutley LB, Beringer J, Holtum JA, Turner BL: **Photosynthetic physiology of eucalypts along a sub-continental rainfall gradient in northern Australia.** *Agricultural and Forest Meteorology* 2011, **151**(11):1462-1470.
5.   Pfanz H, Aschan G, Langenfeld-Heyser R, Wittmann C, Loose M: **Ecology and ecophysiology of tree stems: corticular and wood photosynthesis.** *Naturwissenschaften* 2002, **89**(4):147-162.
6.   Scholz FG, Bucci SJ, Goldstein G, Meinzer FC, Franco AC, Miralles-Wilhelm F: **Biophysical properties and functional significance of stem water storage tissues in Neotropical savanna trees.** *Plant, Cell & Environment* 2007, **30**(2):236-248.
7.   Niklas KJ: **The mechanical role of bark.** *American Journal of Botany* 1999, **86**(4):465-469.
8.   Gill A, Ashton D: **The role of bark type in relative tolerance to fire of three central Victorian eucalypts.** *Australian Journal of Botany* 1968, **16**(3):491-498.
9.   Lawes MJ, Adie H, Russell-Smith J, Murphy B, Midgley JJ: **How do small savanna trees avoid stem mortality by fire? The roles of stem diameter, height and bark thickness.** *Ecosphere* 2011, **2**(4):1-13.
10.   Romero C, Bolker BM: **Effects of stem anatomical and structural traits on responses to stem damage: an experimental study in the Bolivian. Amazon**. *Canadian Journal of Forest Research* 2008, **38**(3):611-618.
11.   Romero C, Bolker BM, Edwards CE: **Stem responses to damage: the evolutionary ecology of *Quercus* species in contrasting fire regimes.** *New Phytologist* 2009, **182**(1):261-271.
12.   Fagard M, Höfte H, Vernhettes S: **Cell wall mutants.** *Plant Physiology and Biochemistry* 2000, **38**(1):15-25.
13.   Plomion C, Leprovost G, Stokes A: **Wood formation in trees.** *Plant physiology* 2001, **127**(4):1513-1523.
14.   Whitmore T: **Studies in systematic bark morphology.** . *New Phytologist* 1962, **61**(2):191-207.
15.   Dickmann DI: **An overview of the genus *Populus.*** In: *Poplar culture in north America.* NRC Research Press; 2001: 1-42.
16.   Heyd RL: **Managing beech bark disease in Michigan.** In: *Beech bark disease: proceedings of the beech bark disease symposium June 16–18, 2004; Saranac lake NY Gen Tech Rep NE: 2005.*
17.   Marshall JM, Smith EL, Mech R, Storer AJ: **Estimates of *Agrilus planipennis* infestation rates and potential survival of ash.** *The American Midland Naturalist* 2013, **169**(1):179-193.

18. Hengst GE, Dawson JO: **Bark properties and fire resistance of selected tree species from the central hardwood region of North America.** *Canadian Journal of Forest Research* 1994, **24**(4):688-696.

19. Wang GG, Wangen SR: **Does frequent burning affect longleaf pine (*Pinus palustris*) bark thickness?** *Canadian Journal of Forest Research* 2011, **41**(7):1562-1565.

20. Pereira H: **Cork chemical variability**. *BioResources* 2013, **8**(2):2246-2256.

21. Romero C: **Tree responses to stem damage.** PhD thesis: University of Florida, Gainesville, FL, USA; 2006.

22. Yordanov YS, Regan S, Busov V: **Members of the LATERAL ORGAN BOUNDARIES DOMAIN transcription factor Family are involved in the regulation of secondary growth in *Populus*.** *The Plant Cell* 2010, **22**(11):3662-3677.

23. Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G: **Genetics of gene expression surveyed in maize, mouse and man.** *Nature* 2003, **422**(6929):297-302.

24. Beaulieu J, Doerksen T, Boyle B, Clément S, Deslauriers M, Beauseigle S, Blais S, Poulin P-L, Lenz P, Caron S: **Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression.** *Genetics* 2011, **188**(1):197-214.

25. Kirst M, Myburg AA, De León JP, Kirst ME, Scott J, Sederoff R: **Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus.** *Plant Physiology* 2004, **135**(4):2368-2378.

26. Novaes E, Osorio L, Drost DR, Miles BL, Boaventura-Novaes CR, Benedict C, Dervinis C, Yu Q, Sykes R, Davis M: **Quantitative genetic analysis of biomass and wood chemistry of *Populus* under different nitrogen levels.** *New Phytologist* 2009, **182**(4):878-890.

27. Pot D, Rodrigues J-C, Rozenberg P, Chantre G, Tibbits J, Cahalan C, Pichavant F, Plomion C: **QTLs and candidate genes for wood properties in maritime pine (*Pinus pinaster* Ait.).** *Tree Genetics & Genomes* 2006, **2**(1):10-24.

28. Thumma BR, Southerton SG, Bell JC, Owen JV, Henery ML, Moran GF: **Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus nitens*.** *Tree Genetics & Genomes* 2010, **6**(2):305-317.

29. Wegrzyn JL, Eckert AJ, Choi M, Lee JM, Stanton BJ, Sykes R, Davis MF, Tsai CJ, Neale DB: **Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem.** *New Phytologist* 2010, **188**(2):515-532.

30. Muchero W, Sewell MM, Ranjan P, Gunter LE, Tschaplinski TJ, Yin T, Tuskan GA: **Genome anchored QTLs for biomass productivity in hybrid *Populus* grown under contrasting environments.** *PloS one* 2013, **8**(1):e54468.

31. Muchero W, Guo J, DiFazio SP, Chen J-G, Ranjan P, Slavov GT, Gunter LE, Jawdy S, Bryan AC, Sykes R: **High-resolution genetic mapping of allelic variants associated with cell wall chemistry in *Populus*.** *BMC genomics* 2015, **16**(1):24.

32. Bundock PC, Potts BM, Vaillancourt RE: **Detection and stability of quantitative trait loci (QTL) in** *Eucalyptus globulus. Tree Genetics & Genomes* 2008, **4**(1):85-95.

33. Shepherd M, Cross M, Dieters MJ, Henry R: **Branch architecture QTL for** *Pinus elliottii* **var.** *elliottii* **x** *Pinus caribaea* **var.** *hondurensis* **hybrids.** *Annals of Forest Science* 2002, **59**(5-6):617-625.

34. Liu T, Tang S, Zhu S, Tang Q: **QTL mapping for fiber yield-related traits by constructing the first genetic linkage map in ramie (***Boehmeria nivea* **L. Gaud).** *Molecular Breeding* 2014, **34**(3):883-892.

35. Li J, Burmeister M: **Genetical genomics: combining genetics with gene expression analysis.** *Human Molecular Genetics* 2005, **14**(suppl 2):R163-R169.

36. Drost DR, Puranik S, Novaes E, Novaes CR, Dervinis C, Gailing O, Kirst M: **Genetical genomics of** *Populus* **leaf shape variation.** *BMC Plant Biology* 2015, **15**(1):166.

37. Jansen RC, Nap J-P: **Genetical genomics: the added value from segregation**. *TRENDS in Genetics* 2001, **17**(7):388-391.

38. Yin T, Zhang X, Gunter L, Priya R, Sykes R, Davis M, Wullschleger SD, Tuskan GA: **Differential detection of genetic loci underlying stem and root lignin content in** *Populus***.** *PLoS One* 2010, **5**(11):e14021.

39. Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A: **The genome of black cottonwood,** *Populus trichocarpa* **(Torr. & Gray).** *Science* 2006, **313**(5793):1596-1604.

40. Fitch R: **WinSTAT for Excel**. *The statistics add-in for Microsoft Excel R Fitch Software* 2006.

41. Van Ooijen J, Kyazma B: **MapQTL 6**. *Software for the mapping of quantitative trait loci in experimental populations of diploid species Kyazma BV: Wageningen, Netherlands* 2009.

42. Storey JD, Tibshirani R: **Statistical significance for genomewide studies.** *Proceedings of the National Academy of Sciences* 2003, **100**(16):9440-9445.

43. Van Ooijen J: **MapQTL® 5**. *Software for the mapping of quantitative trait loci in experimental populations Kyazma BV, Wageningen* 2004.

44. Kelleher CT, Chiu R, Shin H, Bosdet IE, Krzywinski MI, Fjell CD, Wilkin J, Yin T, DiFazio SP, Ali J: **A physical map of the highly heterozygous** *Populus* **genome: integration with the genome sequence and genetic map and analysis of haplotype variation.** *The Plant Journal* 2007, **50**(6):1063-1078.

45. Ko JH, Kim HT, Hwang I, Han KH: **Tissue-type-specific transcriptome analysis identifies developing xylem-specific promoters in poplar.** *Plant Biotechnology Journal* 2012, **10**(5):587-596.

46. Bolstad BM, Irizarry RA, Åstrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.** *Bioinformatics* 2003, **19**(2):185-193.

47. Saeed A, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M: **TM4: a free, open-source system for microarray data management and analysis.** *Biotechniques* 2003, **34**(2):374.

48.   Chu VT, Gottardo R, Raftery AE, Bumgarner RE, Yeung KY: **MeV+ R: using MeV as a graphical user interface for Bioconductor applications in microarray analysis.** *Genome Biology* 2008, **9**(7):R118.

49.   Tsai C-J, Ranjan P, DiFazio SP, Tuskan GA, Johnson V, Joshi C: **Poplar genome microarrays.** In: *Genetics, Genomics and Breeding of Poplar.* Enfield, New Hampshire: Science Publishers; 2011: 112-127.

50.   Tuskan GA, DiFazio S, Faivre-Rampant P, Gaudet M, Harfouche A, Jorge V, Labbé JL, Ranjan P, Sabatti M, Slavov G et al.: **The obscure events contributing to the evolution of an incipient sex chromosome in *Populus*: a retrospective working hypothesis**. *Tree genetics & genomes* 2012, **8**(3):559-571.

51.   Rae AM, Tricker PJ, Bunn SM, Taylor G: **Adaptation of tree growth to elevated CO2: quantitative trait loci for biomass in *Populus.*** *New Phytologist* 2007, **175**(1):59-69.

52.   Wu R, Stettler R: **Quantitative genetics of growth and development in *Populus*. II. The partitioning of genotype×environment interaction in stem growth.** *Heredity* 1997, **78**(2):299-310.

53.   Wullschleger SD, Yin T, DiFazio S, Tschaplinski T, Gunter L, Davis M, Tuskan G: **Phenotypic variation in growth and biomass distribution for two advanced-generation pedigrees of hybrid poplar.** *Canadian Journal of Forest Research* 2005, **35**(8):1779-1789.

54.   Zhang D, Zhang Z, Yang K: **QTL analysis of growth and wood chemical content traits in an interspecific backcross Family of white poplar (*Populus tomentosa* × *P. bolleana*) × *P. tomentosa.*** *Canadian Journal of Forest Research* 2006, **36**(8):2015-2023.

55.   Grattapaglia D, Bertolucci FL, Penchel R, Sederoff RR: **Genetic mapping of quantitative trait loci controlling growth and wood quality traits in Eucalyptus grandis using a maternal half-sib Family and RAPD markers.** *Genetics* 1996, **144**(3):1205-1214.

56.   Lima MA, Lavorente GB, da Silva HK, Bragatto J, Rezende CA, Bernardinelli OD, Gomez LD, McQueen-Mason SJ, Labate CA, Polikarpov I: **Effects of pretreatment on morphology, chemical composition and enzymatic digestibility of eucalyptus bark: a potentially valuable source of fermentable sugars for biofuel production–part 1.** *Biotechnology for Biofuels* 2013, **6**(1):1.

57.   Matsushita Y, Yamauchi K, Takabe K, Awano T, Yoshinaga A, Kato M, Kobayashi T, Asada T, Furujyo A, Fukushima K: **Enzymatic saccharification of *Eucalyptus* bark using hydrothermal pre-treatment with carbon dioxide.** *Bioresource Technology* 2010, **101**(13):4936-4939.

58.   Dinus RJ: **Genetic modification of short rotation poplar biomass feedstock for efficient conversion to ethanol.** *US Department of Energy, Office of Fuels Development, Activity No EB* 2000, **52**(03):00.

59.   Dinus RJ: **Genetic improvement of poplar feedstock quality for ethanol production.** *Applied Biochemistry and Biotechnology* 2001, **91**(1-9):23-34.

60.   Sannigrahi P, Ragauskas AJ, Tuskan GA: **Poplar as a feedstock for biofuels: a review of compositional characteristics.** *Biofuels, Bioproducts and Biorefining* 2010, **4**(2):209-226.

61.   Bdeir R, Busov V, Yordanov Y, Gailing O: **Gene dosage effects and signatures of purifying selection in lateral organ boundaries domain (LBD) genes LBD1 and LBD18**. *Plant Systematics and Evolution* 2016, **320**:433-445.

62.   Shuai B, Reynaga-Peña CG, Springer PS: **The lateral organ boundaries gene defines a novel, plant-specific gene Family.** *Plant Physiology* 2002, **129**(2):747-761.

63.   Rae AM, Pinel MP, Bastien C, Sabatti M, Street NR, Tucker J, Dixon C, Marron N, Dillen SY, Taylor G: **QTL for yield in bioenergy *Populus*: identifying GxE interactions from growth at three contrasting sites.** *Tree Genetics & Genomes* 2008, **4**(1):97-112.

64.   Grant EH, Fujino T, Beers EP, Brunner AM: **Characterization of NAC domain transcription factors implicated in control of vascular cell differentiation in *Arabidopsis* and *Populus*.** *Planta* 2010, **232**(2):337-352.

65.   Schwessinger B, Roux M, Kadota Y, Ntoukakis V, Sklenar J, Jones A, Zipfel C: **Phosphorylation-dependent differential regulation of plant growth, cell death, and innate immunity by the regulatory receptor-like kinase BAK1.** *PLoS Genet* 2011, **7**(4):e1002046.

66.   Halliday K, Devlin PF, Whitelam GC, Hanhart C, Koornneef M: **The ELONGATED gene of *Arabidopsis* acts independently of light and gibberellins in the control of elongation growth.** *The Plant Journal* 1996, **9**(3):305-312.

67.   Hecht V, Vielle-Calzada J-P, Hartog MV, Schmidt ED, Boutilier K, Grossniklaus U, de Vries SC: **The *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture.** *Plant Physiology* 2001, **127**(3):803-816.

68.   Shi Y, Tabesh M, Sugrue SP: **Role of cell Adhesion–Associated protein, pinin (DRS/memA), in corneal epithelial migration.** *Investigative Ophthalmology & Visual Science* 2000, **41**(6):1337-1345.

69.   Joo J-H, Alpatov R, Munguba GC, Jackson MR, Hunt ME, Sugrue SP: **Reduction of Pnn by RNAi induces loss of cell-cell adhesion between human corneal epithelial cells.** *Mol Vis* 2005, **11**(1):133-142.

70.   MacMillan CP, Mansfield SD, Stachurski ZH, Evans R, Southerton SG: **Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in *Arabidopsis* and *Eucalyptus*.** *The Plant Journal* 2010, **62**(4):689-703.

71.   Ito S, Suzuki Y, Miyamoto K, Ueda J, Yamaguchi I: **AtFLA11, a fasciclin-like arabinogalactan-protein, specifically localized in screlenchyma cells.** *Bioscience, Biotechnology, and Biochemistry* 2005, **69**(10):1963-1969.

72.   Lafarguette F, Leplé JC, Déjardin A, Laurans F, Costa G, Lesage-Descauses MC, Pilate G: **Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood.** *New Phytologist* 2004, **164**(1):107-121.

73.   Wang H, Jiang C, Wang C, Yang Y, Yang L, Gao X, Zhang H: **Antisense expression of the fasciclin-like arabinogalactan protein FLA6 gene in *Populus* inhibits expression of its homologous genes and alters stem**

biomechanics and cell-wall composition in transgenic trees. *Journal of Experimental Botany* 2014, **66**(5):1291-1302.

Figures and Tables
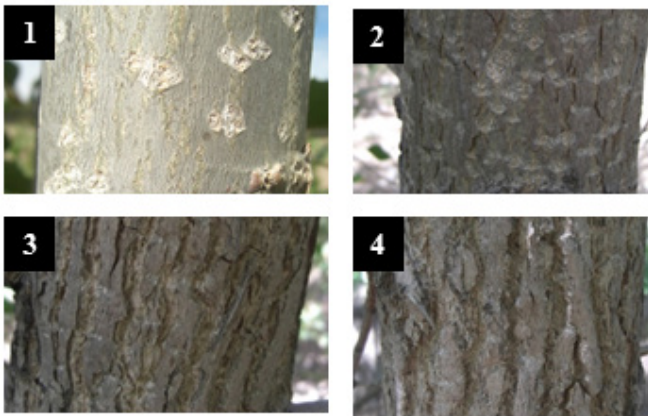
Figure 1 Bark texture scale for *Populus* Family 52-124 offspring. Ranging from smooth (1), medium (2), rough (3) and rough/deeply furrowed bark (4), and level 4 was only found in Oregon.
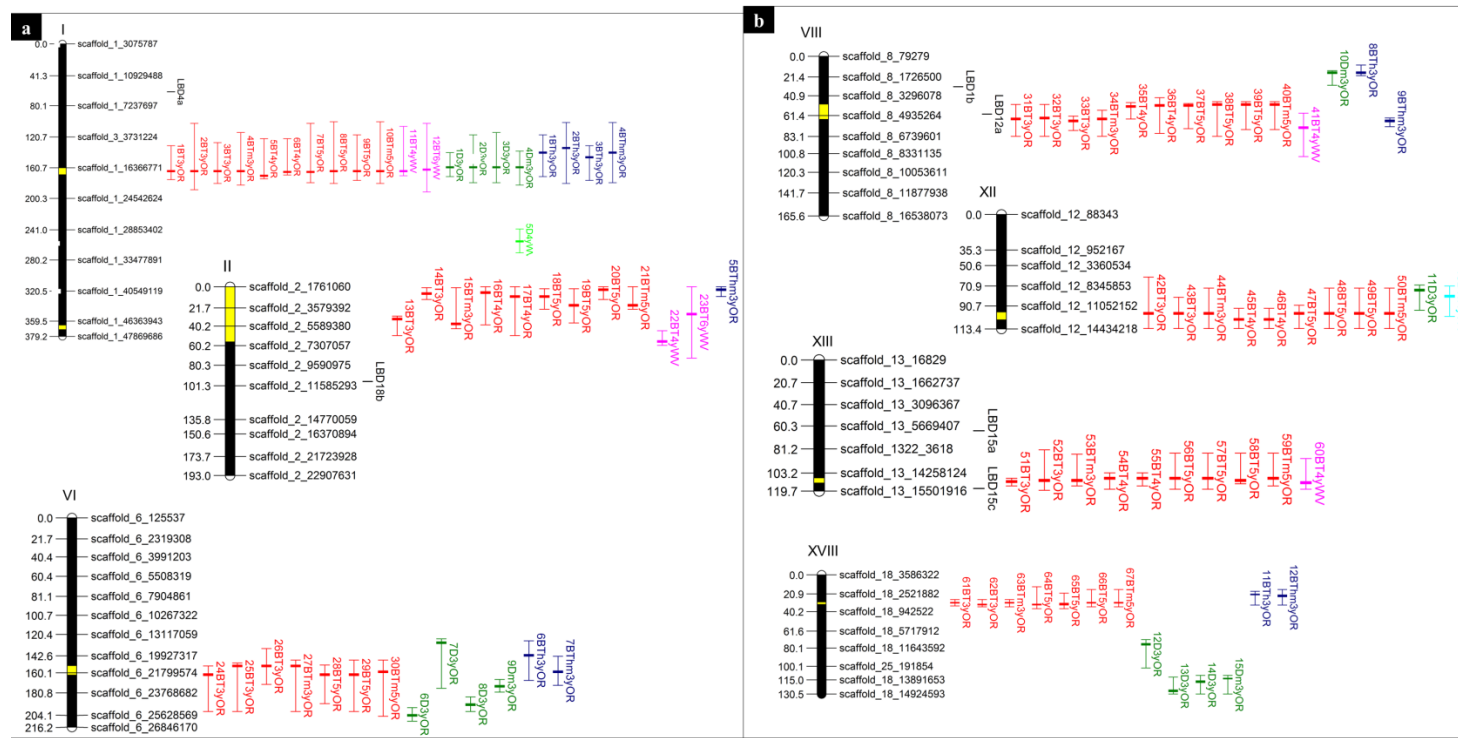
Figure 2 QTL anchored to the genome of *Populus trichocarpa* (V3 assembly, Kelleher *et al.*, 2007). The actual map has a high marker density (average marker spacing: 5 markers per 4 cM). For illustration purposes, for each linkage group an evenly spaced selection of scaffolds is shown (1 marker per 20 cM). The yellow regions on LGs represent LOD score maxima across years and environments. QTL for bark texture (BT), diameter (D), and bark thickness (BTh) are shown in red, green, and blue for Oregon and in pink, light green, and turquois for West Virginia and named according to Table S2. LOD score maxima, genome-wide intervals (solid bars) and chromosome-wide intervals are shown for QTL that were identified in different years and environments (see Table S2). The outer lines of bars are CW thresholds and middle lines are QTL LOD maxima. The exact map and physical locations of QTL are shown in Table S2. Scaffold intervals are represented in Mb. Black vertical lines represent the physical location of LBD genes in the *P. trichocarpa* genome, orthologues are notes by a, b or c.
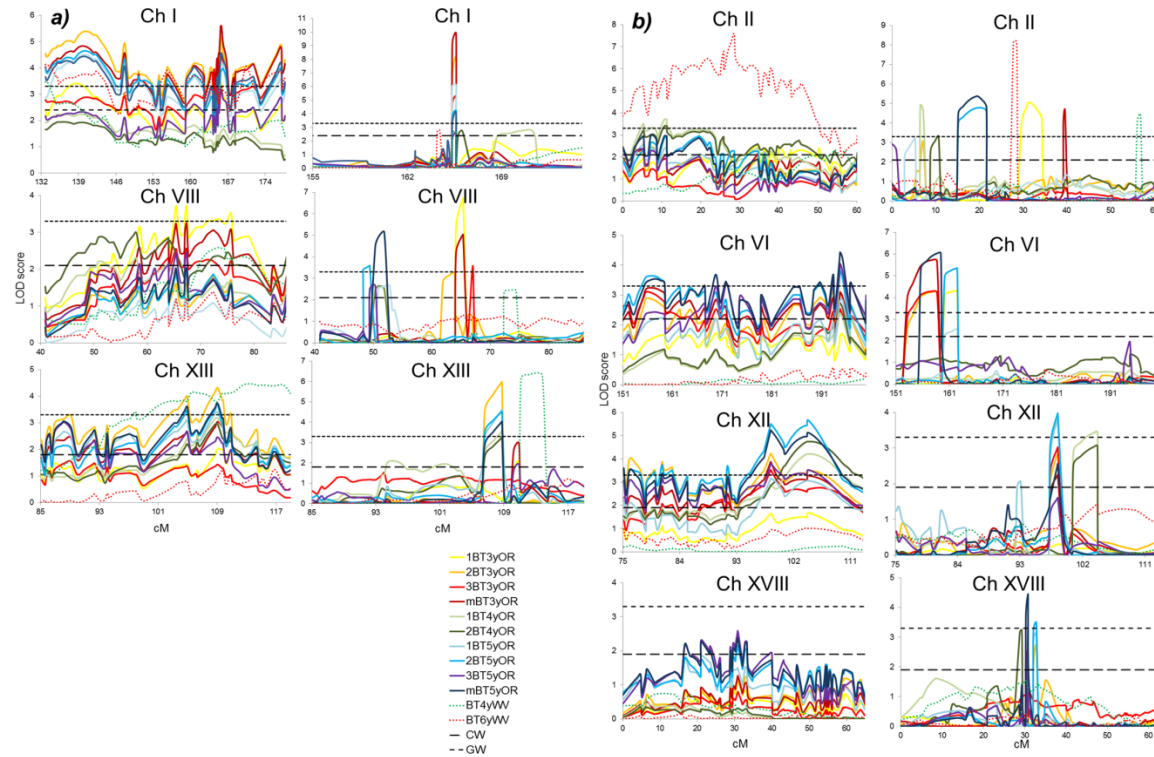
Figure 3 LOD score profiles for bark texture (BT) QTL detected in the *Populus* Family 52-124. Specifically, BT QTL detected on **a)** LG I, VIII and XIII and **b)** LG II, VI, XII and XVIII using interval mapping (left) and the Multiple-QTL Model (MQM) with co-factor selection (right) across all experimental replicates at Oregon (OR) across 3 years and at West Virginia (WV) across 2 years. chromosome-wide (CW) and genome-wide (GW) significance thresholds are shown with dashed lines ($\alpha = 0.05$, 1,000 permutations). Yellow, orange, and red solid lines represent LOD score profiles for 3-year-old replicated samples at Oregon, solid shades of green lines for 4-year-old samples, and shades of blue lines for 5-year-old samples. Dotted red and green lines represent LOD profiles for 4-year-old and 6-year-old samples at the West Virginia site. Two broken horizontal lines represent GW and CW LOD significance threshold after 1,000 permutations at the $p \leq 0.05$ significance level. The exact map and physical locations of QTL are shown in Table S2.
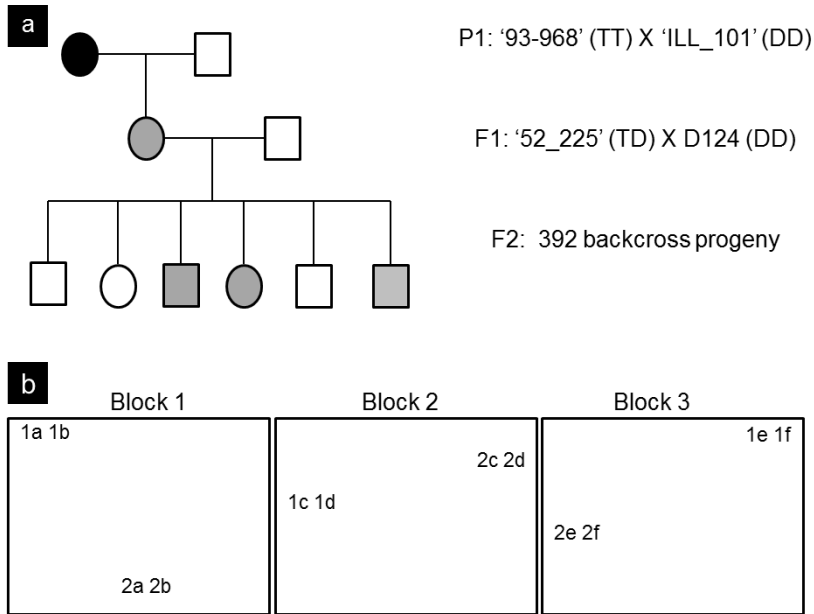
Supplementary Figure 1 Mapping pedigree and block layout for *Populus* Family 52-124 used in this study. a) Family 52-124 is a pseudo-backcross pedigree between clone 52-225, an F1 hybrid derived from *P. trichocarpa* (T), 93–968 X P. deltoides (D) (ILL-101), back-crossed to *P. deltoides* (clone D124 (Novaes et al. 2009). Shading in pedigree: black = TT, gray = TD, White = DD genotype. b) Progeny plantation replicates and block layout in Oregon. Two adjacent replicates in each block and each genotype is represented in three blocks (6 replicates in total). Only two genotypes out of 392 genotypes are shown as an example.
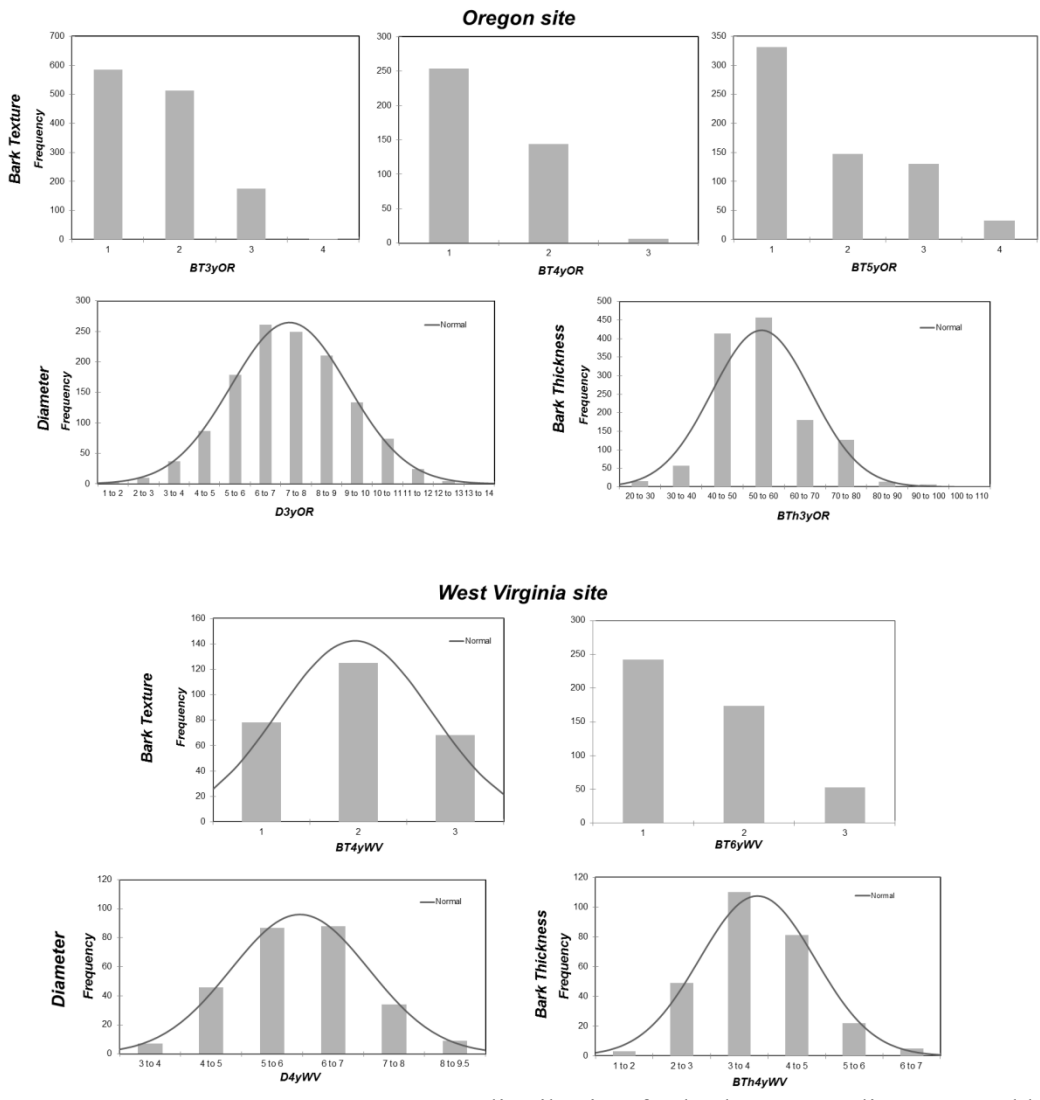
Supplementary Figure 2 Frequency distribution for bark texture, diameter and bark thickness (a, b, and c, respectively) across Oregon and West Virginia sites and various years in *Populus* Family 52-124.

Supplementary Table 1[2] Pair-wise estimates of phenotypic correlations calculated as Pearson correlation coefficients between bark texture (BT), diameter (D) and bark thickness (BTh) phenotypes collected from *Populus* Family 52-124.

Note: Specifically, the analysis was done within and between experimental blocks, across years and across sites. Symbols show p-value significance, where $p \leq 0.05 = *$, $p \leq 0.01 = **$, $p \leq 0.001 = ***$, and $p \leq 0.0001 = ****$. White cells are correlations between same trait across years and sites; red cells are correlation between diameter and bark texture; green cells are correlation between bark thickness and bark texture; blue cells are correlation between bark thickness and diameter.

Supplementary Table 2[2] QTL associated with bark texture, diameter and bark thickness identified in *Populus* Family 52-124 in Oregon and West Virginia.

Note: Chr: chromosome; V2: markers anchored on version 2 of the *P. trichocarpa* genome; V3: version 3 updated physical location; PVE: percent phenotypic variance explained; DD: homozygous for the *P. deltoides* allele, DT: heterozygous for the *P. deltoides* and *P. trichocarpa* alleles. LOD max determined using MQM mapping, value with *: above GW threshold, otherwise above CW threshold. Alternating white or grey shades represent unique QTL found in one or multiple replicates across years and sites. Indexes 1, 2, 3 or m designate: replicate one, two, three or replicates' mean value for OR samples, whereas for WV samples only one replicate was available; BT: bark texture; D: diameter; BTh: bark thickness; 3y, 4y, 5y or 6y: # years old; OR: Oregon, WV: West Virginia. (Example for population: 1BT3yOR: replicate one, bark texture, 3-year-old samples, Oregon site).

---

[2] Supplementary table 1 and 2 are in excel format submitted as digital files.

Supplementary Table 3 Number of candidate genes detected across QTL for all traits.

Trait	QTL on Chromosome	Number of genes
Bark texture	I	148
	VIII	321
	XIII	84
	II	978
	VI	130
	XII	161
	XVIII	47
Diameter	I (~16820000bp)	16
	I (~31760000bp)	23
	VI	362
	VIII	45
	XII	16
	XVIII	231
Bark thickness	I	175
	II	303
	VI	208
	VIII (~157000bp)	16
	VIII (~550000bp)	19
	XII	16
	XVIII	52

The number of genes for each trait in QTL clusters based on MQM mapping with cofactor selection, sorted by significance and reproducibility.

Supplementary Table 4[3] All candidate genes within the ninety-four QTL detected in *Populus* Family 52-124. Physical localization, annotation and expression profile of gene models within each QTL interval for all traits.

Supplementary Table 5 The 90th percentile candidate genes within the ninety-four QTL detected in Populus Family 52-124. Physical localization, annotation and expression profile of gene models in the 90th percentile with high expression within LOD peaks for each QTL interval for all traits.

---

[3] Supplementary table 4 and 5 are in excel format submitted as digital files.

Chapter 3[4]: Genome-Wide Association Studies of bark texture in *Populus trichocarpa*

Abstract

Plants have developed bark as a defense barrier to deal with environmental stresses such as pathogen invasion, drought, fire and UV radiation. Bark is formed when radial expansion pushes the cortex and epidermis outward creating secondary meristem. Mature phellem development requires cell growth, differentiation and cell death as the bark's outer layer is composed of dead differentiated cells. However, the genetic control of this complex phenotype is generally unknown. Here we use association mapping to define the genomic regions associated with natural variation in bark texture in *Populus trichocarpa*. Clonally replicated provenance trials of *P. trichocarpa* were studied for bark texture collected across three sites, multiple years and (two to three) biological replicates per site. A total of 41 genomic intervals containing SNPs significantly associated with bark texture were detected that were highly reproducible across sites, years and replicates. A list of candidate genes within these regions with related putative function was identified. A total of 98 genes were considered candidate genes due to significance and putative function with connection to the bark texture phenotype. Genome-wide association mapping with a high-density SNP array allowed us to detect narrow genomic intervals (1-20 kb) with high reproducibility and shared candidate genes for bark texture variation. For example, a membrane-associated apoptosis protein and a wall-associated receptor kinase (PR5K-like) protein, which both are involved in radial growth and tissue differentiation, lie within highly significant trait-associated regions. Two copies of root hair defective 3 genes, known as transmembrane proteins in *Populus* differentiating xylem and phloem, were also significantly associated with bark texture and co-located with major QTL for bark texture, bark thickness and diameter from our previous study, suggesting an important role in radial growth.

---

[4] The material contained in this chapter has been submitted to Tree Genetics and Genomics.

Introduction

Bark in trees is formed when radial growth pushes the cortex and epidermis outward creating secondary meristem, i.e., the cork cambium [1]. The cork cambium then outwardly produces cork (phellem) cells, the bark's outer layer which is composed of dead differentiated cells [1]. As an adaptation to terrestrial life, bark acts as a defense barrier, including suberins in the phellem layers, to deal with environmental stresses such as pathogen invasion, drought, fire and UV radiation [2, 3]. Despite of its importance, tree bark biology has been less extensively studied than wood biology.

Mature phellem development requires cell growth, differentiation and death. There are common steps in the differentiation of lignified secondary xylem, phloem and phellem cells, including cell division, cell expansion or elongation, secondary cell wall deposition, lignification, and programmed cell death [4]. Carpita et al. [5] estimate that as many as 15% of the genes in a plant genome may play a role in cell wall synthesis, remodeling, or turnover. Secondary cell wall formation is a complex process that requires the coordinate regulation of diverse metabolic pathways [6], however the molecular mechanisms and components that may be essential for cell wall formation in phellem are still unclear. Furthermore, while some defects or variations in secondary cell walls have been characterized, such as irregular xylem [7] or defects in lignin biosynthesis [8], the genetic basis of the phenotypic diversity in bark texture remains completely undefined.

Bark texture varies among species, and even among genotypes within species, and has notable phenotypic diversity ranging from smooth, peeling, fractured, fissured to plated [9]. Differences in bark texture are of ecological importance and have been related to different levels of susceptibility to fungal and insect diseases and to fire [10-13]. For example, tree species with smooth bark, e.g., beech and ash, have been reported to be less susceptible to insect and fungal diseases [12, 13]. Bark thickness and moisture content are reported to be associated with higher fire resistance, where a thicker bark allows for their survival during intense fires [11, 14, 15]. Previously, we have analyzed the bark texture of an interspecific pseudo-backcross pedigree derived from a cross of a female hybrid clone (*Populus trichocarpa* x *P. deltoides*) with a male *P. deltoides* parent through QTL mapping [16]. Highly reproducible QTL were detected encompassing several candidate genes including PopNAC128, which has been reported to play a role in

lignified fiber cell and ray development [17], along with Pinin and Fasciclin (PopFLA) that play a role in cell adhesion, cell shape and migration [18, 19].

The parental species used to derive the pedigree display contrasting bark textures. *P. trichocarpa* (Torr. & Gray) bark is smooth or lightly flaky, while *P. deltoides* (Bartr. ex Marsh) bark is rough and deeply furrowed. Because the backcross was with a *P. deltoides* parent, the QTL pedigree segregated mainly for bark types typical for *P. deltoides*. It is therefore unclear, if and which of the discovered QTL are common or specific for the bark types of the two species.

Here, we took advantage of a large *P. trichocarpa* association population established at three field sites to further investigate the genetic architecture of bark texture through a genome-wide association study (GWAS). We compare the QTL intervals with the location of GWAS candidate genes to identify common genes and at the same time study new candidate genes and loci, not discovered in the QTL study and possibly specific to the development of *P. trichocarpa* bark texture. Our specific objectives are as follows: 1) identification of candidate genes for bark texture, 2) comparison to QTL mapping results and quantification of the number of candidate genes within the QTL intervals, and finally 3) identification of the differences in the genetic control of bark texture in *P. trichocarpa* and *P. deltoides*.

Material and Methods
Plant material

Clonally replicated populations of 1,100 black cottonwood genotypes (*Populus trichocarpa* Torr. & Gray) were collected from the whole species range encompassing ca. 16 native stands stretching from 38.8° to 54.3° N latitude, from California to British Columbia [20] and were established in replicated field trials in Corvallis, OR (44°34′14.81″N 123°16′33.59″W), Clatskanie, OR (46°6′11″N, 123°12′13″W) and Placerville, CA (38°45′14″N, 120°44′25″W) as previously described [21].
Genetic data

A total of 917 genotypes from the population were resequenced to a minimum 15x depth using the Illumina Genome Analyzer, HiSeq 2000 and HiSeq 2500 as previously described for the same population [20]. Briefly, reads were aligned to the *P.*

trichocarpa reference genome version 3 using BWA 0.5.9-r16 with default parameters and SNPs and indels were called using SAMtools mpileup (-E –C 50 –DS –m 2 –F 0.000911 –d 50000) and bcftools (-bcgv –p 0.999089). The resulting SNP and indel dataset is available at http://bioenergycenter.org/besc/gwas/. From this output, we created a GWAS mapping panel of 8.253,066 million SNPs with a minor allele frequency of 0.05.

Phenotypic measurements

Phenotypic data for bark texture (BT) were analyzed in this study for GWAS in all three sites. For the Clatskanie site (CLA), bark texture data were collected in 2012, 2013 and 2015 [i.e., 3-, 4-, and 6-year-old trees] by visual inspection, whereas for the Corvallis (COR) and Placerville sites (PLC), the data were collected in 2013 and 2014 [i.e., 4- and 5-year-old trees], respectively. Bark texture was assigned a qualitative score based on a scale from 1 (smooth) to 3 (flaky with grooves) (Fig. S1). Two replicates of 1,100 genotypes in three blocks, for a total of six ramets per genotype, were analyzed at CLA13 and PLC14 sites, along with three replicates for CLA12 and COR13 and one replicate for CLA15. In addition to the field assessment, the Clatskanie site collection during 2012 had two replicate wood disks with bark collected at breast height per genotype, thus a total of 5 replicates. Differences in age sampling and across different environments were used to evaluate variation in measured phenotypes.

Statistical Analysis

Pair-wise phenotypic correlations were calculated as Spearman's rank correlation coefficients for all data sets across the three different geographic regions using WinSTAT software [22] to assess correlations within blocks and within and between years (Table S1).

Genome-wide association mapping

The data were evaluated for the presence of statistical outliers, and recording errors were corrected or deleted. WinSTAT was used to check for normal distribution of residuals. Transformations were deemed unnecessary (Fig. S2). Association analyses

were performed using the efficient mixed model association (EMMA) algorithm in the EMMAX software with kinship as the correction factor for genetic background effects [23] to compute genotype to phenotype associations using 8.2 million SNP and indel variants with bark texture as the phenotypic variable. On average, a marker is represented every 17 base pairs (bp) across the genome. Significant and reproducible SNPs strongly associated with bark texture trait were detected at a FDR threshold of 0.05. Significance thresholds were set at negative log p-values of 6. False discovery rate (FDR) was determined as described by Storey and Tibshirani [24].

Candidate genes

Based on a linkage disequilibrium (LD) decay below 0.2 within 3-6 kb in this *P. trichocarpa* population [20], a 6-kb upstream and downstream region was examined for candidate genes. The significant genomic regions must have at least one detectable SNP with a significance of -log10 p≥6 and be circumscribed by reproducible SNPs with a significance of -log10 p≥5 (Table S3). The underlying genes were identified based on the *P. trichocarpa* reference genome v3.0 in the Phytozome database using the BioMart tool (https://phytozome.jgi.doe.gov) (Table S3). A complete gene list with InterPro descriptions was generated including loci of putative and/or unknown functions. Expression profiles of the gene models from various tissues: bud, leaf, various parts of root and stem, were obtained from the PhytoMine database (https://phytozome.jgi.doe.gov/phytomine/begin.do).

Results

Phenotyping for bark texture

Phenotypic variation in the association population for bark texture ranged from smooth (score=1) to flaky bark (score=3) (Fig. S1). The mean values and standard errors for Clatskanie, Corvallis and Placerville sites were: CLA12 (mean: 1.53, SE: 0.70), CLA13 (mean: 1.50, SE: 0.69) and CLA15 (mean: 1.89, SE: 0.81), COR13 (mean: 1.71, SE: 0.62) and PLC14 (mean: 1.46, SE: 0.63), respectively. Phenotypic correlations were generally higher within the same environment with varying levels of significance across different environments (Table S1).

Genetic associations with phenotypic variation in bark texture

The GWAS using 8.2 million SNPs uncovered a total of 755 individual SNPs significantly (-log10 p≥5) and reproducibly associated with bark texture distributed across all 19 chromosomes. The 175 genomic regions defined by these SNPs include 380 unique genes (Table S2). Specifically, 472, 255, 4, 28 and 17 SNPs were detected for the data sets Clatskanie 2012 (CLA12), CLA13, CLA15, Corvallis 2013 (COR13) and Placerville 2014 (PLC14), respectively. The Manhattan plots for the five location-by-year data sets with their replicates are shown in Figure 1.

At a significance threshold of -log10 p≥6 the 175 genomic regions were further reduced to 50 genomic regions ranging from 1-20 kb intervals, featuring 324 out of the 755 individual SNPs (Table S3). Nine genomic regions had no predicted genes within 50 kb of significant SNPs; accordingly, only 41 regions containing 98 out of the 380 genes were further used to describe gene annotations based on the interval of the significant genomic region with detectable SNPs (-log10 p≥5).

Characterization of genomic regions

Table 1 summarizes the 41 genomic regions including SNPs with the highest significance value (detailed information in Table S3). These regions were categorized into two groups: reproducible across sites and reproducible across replicates within a site. Twenty-nine regions were reproducible across sites most of them having the exact same markers being significantly associated across multiple replicates/sites. The remaining twelve regions were reproducible across clonal replicates within a site. Comparing the significantly associated regions, they ranged from 1-20 kb with as few as 2 detected SNPs to as many as 53 SNPs. Most trait-associated SNP markers were located within noncoding regions (290) while 34 SNPs were within the coding regions of genes (nonsynonymous=26, synonymous=8). Details on all SNPs' position and features are summarized in Table S3.

Seventeen regions had -log10 p≥7, with the highest significance on LG XV (8.44). Considering replicates, four regions, detected on LG VI, XV, XVI and XVII, were found in three different data sets (regions 1, 8, 11 and 12, Table 1). Four trait-associated

regions had more than 8 SNPs with -log10 p≥6 (region 5, 8, 11 and 31); whereas five regions contained 4-5 individual SNPs above -log10 p≥6 (region 6, 9, 17, 33 and 34, Table S3).

A detailed list of the underlying putative genes within the 41 regions and their interval is shown in Table S3. Some regions contained one gene while others included as many as nine, totaling 98 putative genes of which 75 were reproducible across sites and/or years and 23 were reproducible within sites/years. Analyzing the InterPro annotation and putative function of the significant genes along with the position/feature of SNPs, 32 of the 98 genes had unknown functions. Of the remaining genes several candidate genes had annotations possibly related to bark texture (Table 2). Eight of these genes occurred in QTL detected for bark texture in our previous study [16] (Fig. 2). Specifically, we found on LGI Potri.001G460900 (gene of unknown function), on LGVI the three genes Potri.006G173900 (gene of unknown function), Potri.006G200100 (gene of unknown function) and Potri.006G200200 (ring finger domain, ubiquitin ligase), on LGXII the three genes Potri.012G116800 (gene of unknown function), Potri.012G116900 and Potri.012G117000 (root hair defective 3), and finally on LGXVIII Potri.018G019900 (containing a NAF domain).

Discussion

The molecular machinery leading to alterations in bark texture is unknown. With the use of genome-wide association mapping we were able to link bark texture phenotypes to associated polymorphisms in the genome and identify putative genes related to bark development.

Even though the genes involved in bark texture variation are largely unknown, in the last two decades several studies have closely studied genes involved in related developmental processes. During radial growth mechanical stresses imposed in the phellogen create variation in bark roughness that can either produce a smooth or furrowed/flaky bark texture [25]. From studies on secondary growth from xylem mother cells (wood development), a comprehensive set of regulatory genes was discovered through various methods, including protein arrays linking genomics to proteomics and through EST sequencing providing expression analysis of genes during wood formation

[4, 6, 26-29]. The study by Song et al. (2011) is one of the few which examined both xylem and phloem differentiation and created a subtracted gene list that included receptors, transporters, cell wall formation related and intracellular trafficking proteins [26].

Our results revealed multiple significant associations. Through GWAS, we were able to detect a number of putative genes involved in the control of bark texture, as well as differentiate between the genes responsible for the flaky bark texture in *P. trichocarpa* and the rough furrowed texture in *P. deltoides* in comparison with our previous work [16]. Overall the detected genomic regions involved in the control of bark texture in *P. trichocarpa* and *P. deltoides* vary. Eight genes out of 380 GWAS-identified genes were found to be co-locating with the QTL from the previous study, five of which have an unknown function. Two genes with annotated function, Potri.012G116900 and Potri.012G117000, on LG XII at Clatskanie (2012 and 2013) with a -log10 p=5.12, annotated as root hair defective 3, co-located with major highly reproductive QTL for bark texture [16]. The same QTL region was also associated with bark thickness and diameter, suggesting a role of this QTL in radial growth [16]. These genes putatively function as transmembrane proteins identified from the plasma membrane of *Populus* differentiating xylem and phloem [26].

Genome-wide association studies can suffer from too few observations, low repeatability and insufficient statistical thresholds. The large sample size in the present study (n=917) based on 8.2 million SNPs contributed to minimize the limitation of population size as well as defining the linkage disequilibrium rate between a SNP marker and a functional gene. Thus, most of significant associations were located within relatively small intervals (70% of genomic regions have ≤5 kb intervals). Among the 98 genes underlying the top 41 regions, a list of putative genes based on the highest p-values and reproducibility with possible functional control of bark texture were identified (Table 2).

Radial growth and tissue differentiation

During secondary growth, the vascular cambium consisting of meristematic tissues goes through cell division to radially generate secondary xylem to its inside and

secondary phloem to its outside [30]. The proportions of the cell types in inner bark (e.g., parenchyma, fibers, and sclereids) and the types of sieve elements, their sizes, and arrangements in the phloem vary among species, often forming patterns of taxonomic importance [31]. Vascular differentiation and differentiation of the outer bark involves cell division, orientated cell differentiation, cell expansion, cell wall thickening, and programmed cell death [26, 31]. Discontinuous periderms (textured bark) could result from variable radial meristematic activity in the cork cambium due to mechanical stresses from radial growth [25].

The gene Potri.015G124000, a membrane-associated apoptosis protein (NAP1), with a SNP located in its intron reached the highest significance score (-log10 p=8.44) and reproducibility across three datasets, CLA12, CLA13 and PLC14 (Table 2). It is an orthologous to the *A. thaliana* gene AT2G35110, NCK-associated protein (NAP1), an adaptor protein thought to modulate actin nucleation. NAP1 is a component of the WAVE regulatory protein complex that acts as an activator of the ARP2/3 complex involved in actin nucleation and trichome morphogenesis. Mutants in this gene display distorted trichomes and irregularities in trichome branch positioning and expansion [32-34]. NAP1 is thus involved in plant cell morphogenesis and plays an essential role in cell expansion.

Another gene family, Cytochromes, have been detected within six genomic regions (four having (-log10 p≥7) across all years for Clatskanie and Corvallis sites (Table 2). Cytochromes are brassinosteroids perceived by the plasma membrane-localized leucine-rich repeat-receptor kinase BRI1 with essential roles in developmental processes including cell expansion and vascular differentiation [35]. For instance, loss-of-function of *brl1* causes abnormal phloem:xylem differentiation ratios and vascular defects [36]. Other studies have revealed that Cytochrome P450 plays a functional role in secondary metabolite biosynthesis, specifically in lignin monomer composition and controlling cell elongation and de-etiolation in *Arabidopsis* [37, 38]. While an overexpression of F5H, a cytochrome P450-dependent monooxygenases, in *Arabidopsis* led to the accumulation of lignin indicating a key regulatory role in syringyl lignin biosynthesis [37].

Several additional GWAS candidate genes, including pectate lyase, have been reported to regulate cell wall loosening and expansion [39]. Multicopper oxidase (Potri.019G088500, ortholog AT4G28250), known as an Expansin protein, functioning in abscission and cell wall loosening, has a significant SNP within the coding sequence (-log10 p=7.26) resulting in a nonsynonymous substitution and was detected in the CLA12 dataset [40].

Suberin accumulation

Suberization occurs during the development of the secondary radial meristems, i.e., the cork cambium [3]. Phellem is multilayered dead tissue that is made impervious by the disposition of suberin onto their cell walls, a waxy substance that is highly impermeable to gases and water to provide the internal cells of the plants with extra insulation and protection [2, 3].

The genes Potri.017G034700 and Potri.017G035400, wall-associated receptor kinases (PR5K-like), are found within highly significant genomic regions (-log10 p=7.90, synonymous SNP in coding sequence; -log10 p=5.72, SNP in intron, respectively) and are reproducible across year 2012 and 2013 in Clatskanie (Table 2). They are related to a family of protein-serine/threonine kinases and involved in self-incompatibility and disease resistance and are orthologous to the *Arabidopsis* gene AT3G55550 [41, 42]. A highly significant synonymous SNP (-log10 p=7.13) was found in the coding region of the gene Potri.017G034500, an ortholog to the same *Arabidopsis* gene AT3G55550, however it was only characterized as *Populus* gene with protein kinase domain of unknown family (Table 2). Interestingly, an ortholog of the PR5K-like gene, AT5G38280, was identified as a candidate gene involved in apple russeting [43], which results from the accumulation of suberin on the inner part of the cell wall of the outer epidermal cell layers.

Other regulatory genes for suberin accumulation are the MYB transcription factors of which three Myb-like DNA-binding protein genes were detected in Clatskanie across all years at -log10 p=7.49 (Table 2). The MYB gene family have been reported as key regulators in the biosynthesis of lignin and the regulation of secondary cell wall formation [44-47]. In *Arabidopsis*, MYB46 and MYB83 are direct targets of a group of

NAC domain master regulators controlling secondary wall biosynthesis [47-49]. In addition, in *Populus*, several genes were found to be involved in the regulation of wood formation, including PtrMYB3 and PtrMYB20, while a NAC protein Potri.006G277000, detected in the COR13 dataset, is involved in the biosynthesis and regulation of *Populus* suberin affecting the bark periderm, specifically the cork (phellem) [43, 50-52].

Programmed cell death

Programmed cell death is important in secondary plant development and is involved in the generation of the vascular system [53]. Programmed cell death also contributes to the development of bark and its abscission [54]. Abscission results in the separation of the bark tissue from the trunk without injury to living tissue. Bark abscission creates a protection layer of the newly exposed tissue [54]. Lignification occurs in xylem cells during wood development and in bark [4, 55, 56]. When lignification is completed, vessel elements undergo programmed cell death, which involves the hydrolysis of the protoplast [4, 57, 58]. Nonetheless, the specific effects of cell death on bark has not been studied.

Six genes with ring finger domain and ubiquitin ligase function were detected across all sites and several years, of which Potri.006G200200 co-locates with a QTL cluster on LG VI associated with bark texture from our previous study [16]. The genes Potri.001G450700 and Potri.014G138000 (orthologs AT3G61460 and AT1G63900, respectively) have the highest significant values at -log10 p=7.75 and 7.56, respectively, and their putative functions are characterized as SFC and E3 ubiquitin ligases, respectively (Table 2). E3 ubiquitin ligases are known as positive regulators of cell death and defense across various plants [59]. Studies have suggested their involvement in regulation of cellular elongation in *Arabidopsis* and in the biosynthesis of tension wood in *Populus* [29, 60]. The ubiquitin-dependent pathway of proteolysis is essential for vascular development and overexpression of a mutant form of ubiquitin unable to form polyubiquitin chains in transgenic tobacco resulted in abnormal vascular development [61].

One way for a plant to defend itself is by using molecular switches of plant disease resistance, such as NB-ARC domain containing proteins, and activating a cascade

leading to cell death [62, 63]. Four genes within several genomic regions (-log10 p≤7.13) detected across all three locations and years contain a NB-ARC (nucleotide-binding adaptor shared by APAF-1, R proteins, and CED-4) domain, of which two had SNPs within genes, one in the coding sequence (synonymous substitution) and the other in the genes' intron (Table 2). A recent study by Fossdal et al. examined gene expression of secondary phloem of Norway spruce after it had been wounded and inoculated with the fungus *Ceratocystis polonica* [64]. NB resistance genes showed a significant increase in expression along with five miRNAs with putative NB targets [64].

Although seven *Populus* genes were characterized as kinases with unknown family, two genes, Potri.015G124100 and Potri.017G034100 (-log10 p=8.44 and 7.13, respectively), were identified across three of our test sites and are orthologous to *Arabidopsis* genes are known as cysteine-rich receptor-like protein kinases (CRKs). Induced expression of these CRKs triggered hypersensitive response-like cell death in transgenic plants [65].

Future Work

While the candidate gene list in Table 2 provides many promising putative genes involved in cell wall differentiation they need to be further studied to validate their involvement in bark biosynthesis through transgenic lines and forward genetic methods. Genes of unknown function detected within highly significant regions open new opportunities in the future and should be functionally characterized. Two genes, Potri.006G008100 and Potri.017G035600 (-log10 p=6.91 and 7.73, respectively), are top candidates for functional characterization along with 12 other genes with high significance (-log10 p≥7) (Table S3).

Conclusions

      The combination of QTL and GWAS mapping supports the identification of candidate genes. QTL mapping increases the ability to reduce false positive rates whereas GWAS has a much higher resolution than QTL mapping and can narrow down the genomic region to individual candidate genes [66-68]. In our case, the combination of QTL and GWAS mapping provided insights on the different genes responsible for the variations in bark texture ranging from smooth and furrowed across *Populus* species.

References

1.	Carlquist S: **Wood, bark, and stem anatomy of Gnetales: a summary**. *International Journal of Plant Sciences* 1996, **157**(S6):S58-S76.
2.	Pollard M, Beisson F, Li Y, Ohlrogge JB: **Building lipid barriers: biosynthesis of cutin and suberin**. *Trends in plant science* 2008, **13**(5):236-246.
3.	Soler M, Serra O, Molinas M, Huguet G, Fluch S, Figueras M: **A genomic approach to suberin biosynthesis and cork differentiation**. *Plant Physiology* 2007, **144**(1):419-431.
4.	Mellerowicz EJ, Baucher M, Sundberg B, Boerjan W: **Unravelling cell wall formation in the woody dicot stem**. In: *Plant Cell Walls.* Springer; 2001: 239-274.
5.	Carpita N, Tierney M, Campbell M: **Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics**. *Plant molecular biology* 2001, **47**(1):1-5.
6.	Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR: **Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics**. *The Plant Cell* 2005, **17**(8):2281-2295.
7.	Turner SR, Somerville CR: **Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall**. *The plant cell* 1997, **9**(5):689-701.
8.	Piquemal J, Lapierre C, Myton K, O'connell A, Schuch W, Grima-pettenati J, Boudet AM: **Down-regulation of Cinnamoyl-CoA Reductase induces significant changes of lignin profiles in transgenic tobacco plants**. *The Plant Journal* 1998, **13**(1):71-83.
9.	Whitmore T: **STUDIES IN SYSTEMATIC BARK MORPHOLOGY**. *New Phytologist* 1962, **61**(2):191-207.
10.	Gill A, Ashton D: **The role of bark type in relative tolerance to fire of three central Victorian eucalypts**. *Australian Journal of Botany* 1968, **16**(3):491-498.
11.	Lawes MJ, Adie H, Russell-Smith J, Murphy B, Midgley JJ: **How do small savanna trees avoid stem mortality by fire? The roles of stem diameter, height and bark thickness**. *Ecosphere* 2011, **2**(4):1-13.
12.	Marshall JM, Smith EL, Mech R, Storer AJ: **Estimates of Agrilus planipennis infestation rates and potential survival of ash**. *The American Midland Naturalist* 2013, **169**(1):179-193.
13.	Beaver R, Wilding N, Collins N, Hammond P, Webber J: **Insect-fungus relationships in the bark and ambrosia beetles**. *Insect-fungus interactions* 1989:121-143.
14.	Pausas JG: **Bark thickness and fire regime**. *Functional Ecology* 2015, **29**(3):315-327.
15.	Climent J, Tapias R, Pardos JA, Gil L: **Fire adaptations in the Canary Islands pine (Pinus canariensis)**. *Plant ecology* 2004, **171**(1-2):185-196.
16.	Bdeir R, Muchero W, Yordanov Y, Tuskan GA, Busov V, Gailing O: **Quantitative trait locus mapping of Populus bark features and stem diameter**. *BMC plant biology* 2017, **17**(1):224.

17. Grant EH, Fujino T, Beers EP, Brunner AM: **Characterization of NAC domain transcription factors implicated in control of vascular cell differentiation in Arabidopsis and Populus**. *Planta* 2010, **232**(2):337-352.

18. Lafarguette F, Leplé JC, Déjardin A, Laurans F, Costa G, Lesage-Descauses MC, Pilate G: **Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood**. *New Phytologist* 2004, **164**(1):107-121.

19. Wang H, Jiang C, Wang C, Yang Y, Yang L, Gao X, Zhang H: **Antisense expression of the fasciclin-like arabinogalactan protein FLA6 gene in Populus inhibits expression of its homologous genes and alters stem biomechanics and cell wall composition in transgenic trees**. *Journal of experimental botany* 2014, **66**(5):1291-1302.

20. Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, Lipphardt MF, Pennacchio CP, Hellsten U, Pennacchio LA: **Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree Populus trichocarpa**. *New Phytologist* 2012, **196**(3):713-725.

21. Muchero W, Guo J, DiFazio SP, Chen J-G, Ranjan P, Slavov GT, Gunter LE, Jawdy S, Bryan AC, Sykes R: **High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus**. *BMC genomics* 2015, **16**(1):24.

22. Fitch R: **WinSTAT for Excel**. *The statistics add-in for Microsoft Excel R Fitch Software* 2006.

23. Kang HM, Sul JH, Zaitlen NA, Kong S-y, Freimer NB, Sabatti C, Eskin E: **Variance component model to account for sample structure in genome-wide association studies**. *Nature genetics* 2010, **42**(4):348.

24. Storey JD, Tibshirani R: **Statistical significance for genomewide studies**. *Proceedings of the National Academy of Sciences* 2003, **100**(16):9440-9445.

25. Romero C: **Tree responses to stem damage**. University of Florida; 2006.

26. Song D, Xi W, Shen J, Bi T, Li L: **Characterization of the plasma membrane proteins and receptor-like kinases associated with secondary vascular differentiation in poplar**. *Plant molecular biology* 2011, **76**(1-2):97-115.

27. Li Y, Jin F, Chao Q, Wang B-C: **Proteomics analysis reveals the molecular mechanism underlying the transition from primary to secondary growth of poplar**. *Journal of Plant Physiology* 2017, **213**:1-15.

28. Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R: **Gene discovery in the wood-forming tissues of poplar: analysis of 5,692 expressed sequence tags**. *Proceedings of the National Academy of Sciences* 1998, **95**(22):13330-13335.

29. Andersson-Gunnerås S, Mellerowicz EJ, Love J, Segerman B, Ohmiya Y, Coutinho PM, Nilsson P, Henrissat B, Moritz T, Sundberg B: **Biosynthesis of cellulose-enriched tension wood in Populus: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis**. *The Plant Journal* 2006, **45**(2):144-165.

30. Chaffey N, Cholewa E, Regan S, Sundberg B: **Secondary xylem development in Arabidopsis: a model for wood formation**. *Physiologia plantarum* 2002, **114**(4):594-600.

31. Plomion C, Leprovost G, Stokes A: **Wood formation in trees**. *Plant physiology* 2001, **127**(4):1513-1523.

32. Brembu T, Winge P, Seem M, Bones AM: **NAPP and PIRP encode subunits of a putative wave regulatory protein complex involved in plant cell morphogenesis**. *The Plant Cell* 2004, **16**(9):2335-2349.

33. Deeks MJ, Kaloriti D, Davies B, Malhó R, Hussey PJ: **Arabidopsis NAP1 is essential for Arp2/3-dependent trichome morphogenesis**. *Current Biology* 2004, **14**(15):1410-1414.

34. Zhang X, Dyachok J, Krishnakumar S, Smith LG, Oppenheimer DG: **IRREGULAR TRICHOME BRANCH1 in Arabidopsis encodes a plant homolog of the actin-related protein2/3 complex activator Scar/WAVE that regulates actin and microtubule organization**. *The Plant Cell* 2005, **17**(8):2314-2326.

35. Clouse SD, Sasse JM: **Brassinosteroids: essential regulators of plant growth and development**. *Annual review of plant biology* 1998, **49**(1):427-451.

36. Caño-Delgado A, Yin Y, Yu C, Vafeados D, Mora-García S, Cheng J-C, Nam KH, Li J, Chory J: **BRL1 and BRL3 are novel brassinosteroid receptors that function in vascular differentiation in Arabidopsis**. *Development* 2004, **131**(21):5341-5351.

37. Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C: **Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis**. *Proceedings of the National Academy of Sciences* 1998, **95**(12):6619-6623.

38. Szekeres M, Németh K, Koncz-Kálmán Z, Mathur J, Kauschmann A, Altmann T, Rédei GP, Nagy F, Schell J, Koncz C: **Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in Arabidopsis**. *Cell* 1996, **85**(2):171-182.

39. Xu P, Cai X-T, Wang Y, Xing L, Chen Q, Xiang C-B: **HDG11 upregulates cell-wall-loosening protein genes to promote root elongation in Arabidopsis**. *Journal of experimental botany* 2014, **65**(15):4285-4295.

40. Cosgrove DJ: **Loosening of plant cell walls by expansins**. *Nature* 2000, **407**(6802):321.

41. Wang X, Zafian P, Choudhary M, Lawton M: **The PR5K receptor protein kinase from Arabidopsis thaliana is structurally related to a family of plant defense proteins**. *Proceedings of the National Academy of Sciences* 1996, **93**(6):2598-2602.

42. Bent AF: **Plant disease resistance genes: function meets structure**. *The Plant Cell* 1996, **8**(10):1757.

43. Legay S, Guerriero G, Deleruelle A, Lateur M, Evers D, André CM, Hausman J-F: **Apple russeting as seen through the RNA-seq lens: strong alterations in the exocarp cell wall**. *Plant Molecular Biology* 2015, **88**(1-2):21-40.

44. Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N: **EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis**. *The Plant Journal* 2005, **43**(4):553-567.

45. Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlén M, Teeri TT, Lundeberg J: **A transcriptional roadmap to wood formation**. *Proceedings of the National Academy of Sciences* 2001, **98**(25):14732-14737.

46. Caño-Delgado A, Lee J-Y, Demura T: **Regulatory mechanisms for specification and patterning of plant vascular tissues**. *Annual review of cell and developmental biology* 2010, **26**:605-637.

47. Zhong R, Lee C, Zhou J, McCarthy RL, Ye Z-H: **A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis**. *The Plant Cell* 2008, **20**(10):2763-2782.

48. Mitsuda N, Iwase A, Yamamoto H, Yoshida M, Seki M, Shinozaki K, Ohme-Takagi M: **NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis**. *The Plant Cell* 2007, **19**(1):270-280.

49. Zhou J, Lee C, Zhong R, Ye Z-H: **MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis**. *The Plant Cell* 2009, **21**(1):248-266.

50. McCarthy RL, Zhong R, Fowler S, Lyskowski D, Piyasena H, Carleton K, Spicer C, Ye Z-H: **The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis**. *Plant and Cell Physiology* 2010, **51**(6):1084-1090.

51. Zhao Y, Sun J, Xu P, Zhang R, Li L: **WOOD-ASSOCIATED NAC TRANSCRIPTION FACTOR1B Regulates Cell Wall Thickening during Fiber Development in Populus Species1 [W]**. *Plant Physiol* 2014, **164**.

52. Rains MK, de Silva G, Dilini N, Molina I: **Reconstructing the suberin pathway in poplar by chemical and transcriptomic analysis of bark tissues**. *Tree Physiology* 2017:1-22.

53. Groover A, Jones AM: **Tracheary element differentiation uses a novel mechanism coordinating programmed cell death and secondary cell wall synthesis**. *Plant Physiology* 1999, **119**(2):375-384.

54. Addicott FT: **Abscission**: Univ of California Press; 1982.

55. Jelonek T, Pazdrowski W, Tomczak A, Arasimowicz-Jelonek M: **Lignification Markers of the Tracheid Walls of Scots Pine (Pinus sylvestris (L.)) in Various Forms of Dead Bark**. *BioResources* 2017, **12**(2):3992-4003.

56. Sengupta G, Palit P: **Characterization of a Lignified Secondary Phloem Fibre-deficient Mutant of Jute (Corchorus capsularis)**. *Annals of Botany* 2004, **93**(2):211-220.

57. Asif M, Al-Mansoub MA, Khan MSS, Yehya AHS, Ezzat MO, Oon CE, Atif M, Majid ASA, Majid AMSA: **Molecular mechanisms responsible for programmed cell death-inducing attributes of terpenes from Mesua ferrea stem bark towards human colorectal carcinoma HCT 116 cells**. *Journal of Applied Biomedicine* 2017, **15**(1):71-80.

58. Venkatesan T, Choi Y-W, Mun S-P, Kim Y-K: **Pinus radiata bark extract induces caspase-independent apoptosis-like cell death in MCF-7 human breast cancer cells**. *Cell biology and toxicology* 2016, **32**(5):451-464.

59. Janjusevic R, Abramovitch RB, Martin GB, Stebbins CE: **A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase**. *Science* 2006, **311**(5758):222-226.

60. Franciosini A, Lombardi B, Iafrate S, Pecce V, Mele G, Lupacchini L, Rinaldi G, Kondou Y, Gusmaroli G, Aki S: **The Arabidopsis COP9 SIGNALOSOME INTERACTING F-BOX KELCH 1 protein forms an SCF ubiquitin ligase and regulates hypocotyl elongation**. *Molecular plant* 2013, **6**(5):1616-1629.

61. Bachmair A, Becker F, Masterson R, Schell J: **Perturbation of the ubiquitin system causes leaf curling, vascular tissue alterations and necrotic lesions in a higher plant**. *The EMBO journal* 1990, **9**(13):4543.

62. Van Ooijen G, Mayr G, Kasiem MM, Albrecht M, Cornelissen BJ, Takken FL: **Structure–function analysis of the NB-ARC domain of plant disease resistance proteins**. *Journal of experimental botany* 2008, **59**(6):1383-1397.

63. Kohler A, Rinaldi C, Duplessis S, Baucher M, Geelen D, Duchaussoy F, Meyers BC, Boerjan W, Martin F: **Genome-wide identification of NBS resistance genes in Populus trichocarpa**. *Plant molecular biology* 2008, **66**(6):619-636.

64. Fossdal CG, Yaqoob N, Krokene P, Kvaalen H, Solheim H, Yakovlev IA: **Local and systemic changes in expression of resistance genes, nb-lrr genes and their putative microRNAs in Norway spruce after wounding and inoculation with the pathogen Ceratocystis polonica**. *BMC plant biology* 2012, **12**(1):105.

65. Chen K, Fan B, Du L, Chen Z: **Activation of hypersensitive cell death by pathogen-induced receptor-like protein kinases from Arabidopsis**. *Plant molecular biology* 2004, **56**(2):271-283.

66. Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F: **Linkage and association mapping of Arabidopsis thaliana flowering time in nature**. *PLoS genetics* 2010, **6**(5):e1000940.

67. Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P: **An Arabidopsis example of association mapping in structured samples**. *PLoS genetics* 2007, **3**(1):e4.

68. Manenti G, Galvan A, Pettinicchio A, Trincucci G, Spada E, Zolin A, Milani S, Gonzalez-Neira A, Dragani TA: **Mouse genome-wide association mapping needs linkage analysis to avoid false-positive loci**. *PLoS genetics* 2009, **5**(1):e1000331.

Figures and Tables
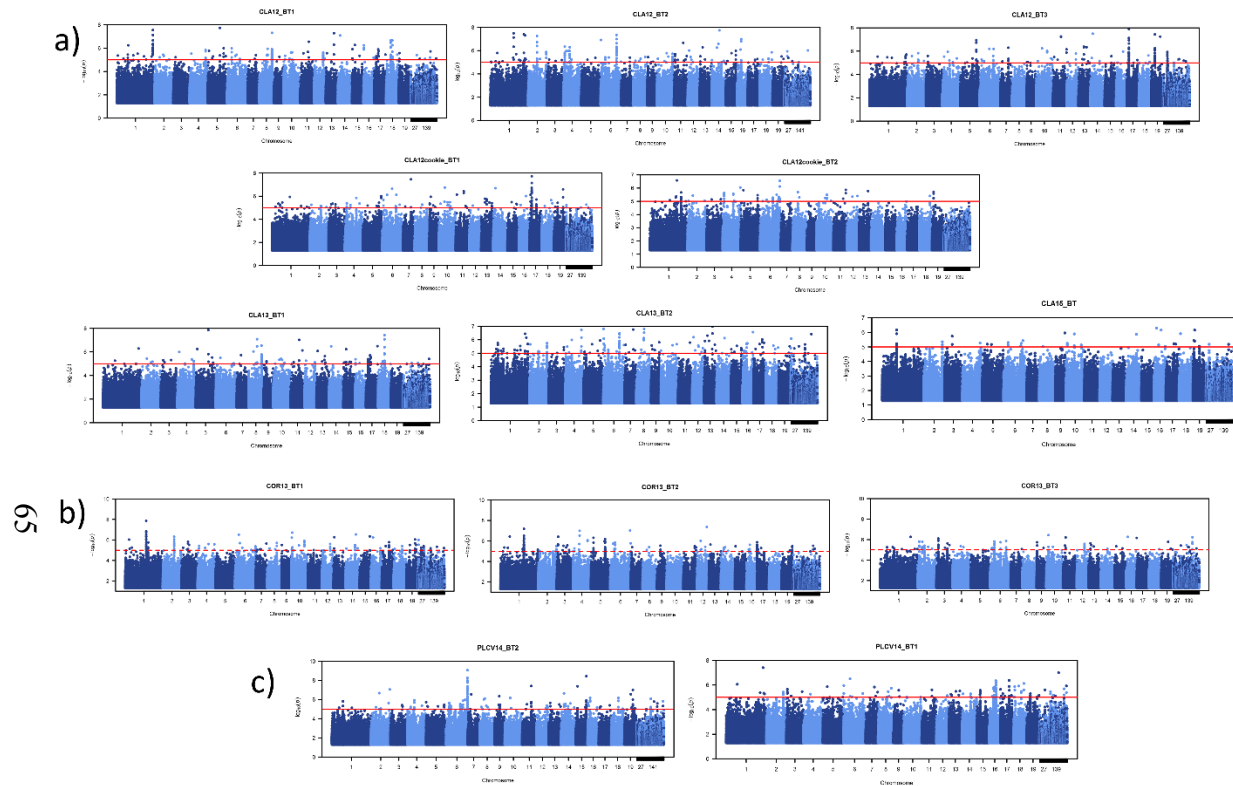
Figure 1 Manhattan plots for GWAS for the *P. trichocarpa* BESC population. Standard −log10 P plots of the study results. The five datasets across three sites: Clatskanie (CLA) for year 2012, 2013 and 2015 (a), Corvallis (COR) for year 2013 (b) and Placerville (PLC) for year 2014 (b); BT1, BT2, BT3: bark texture replicate 1, 2, 3.
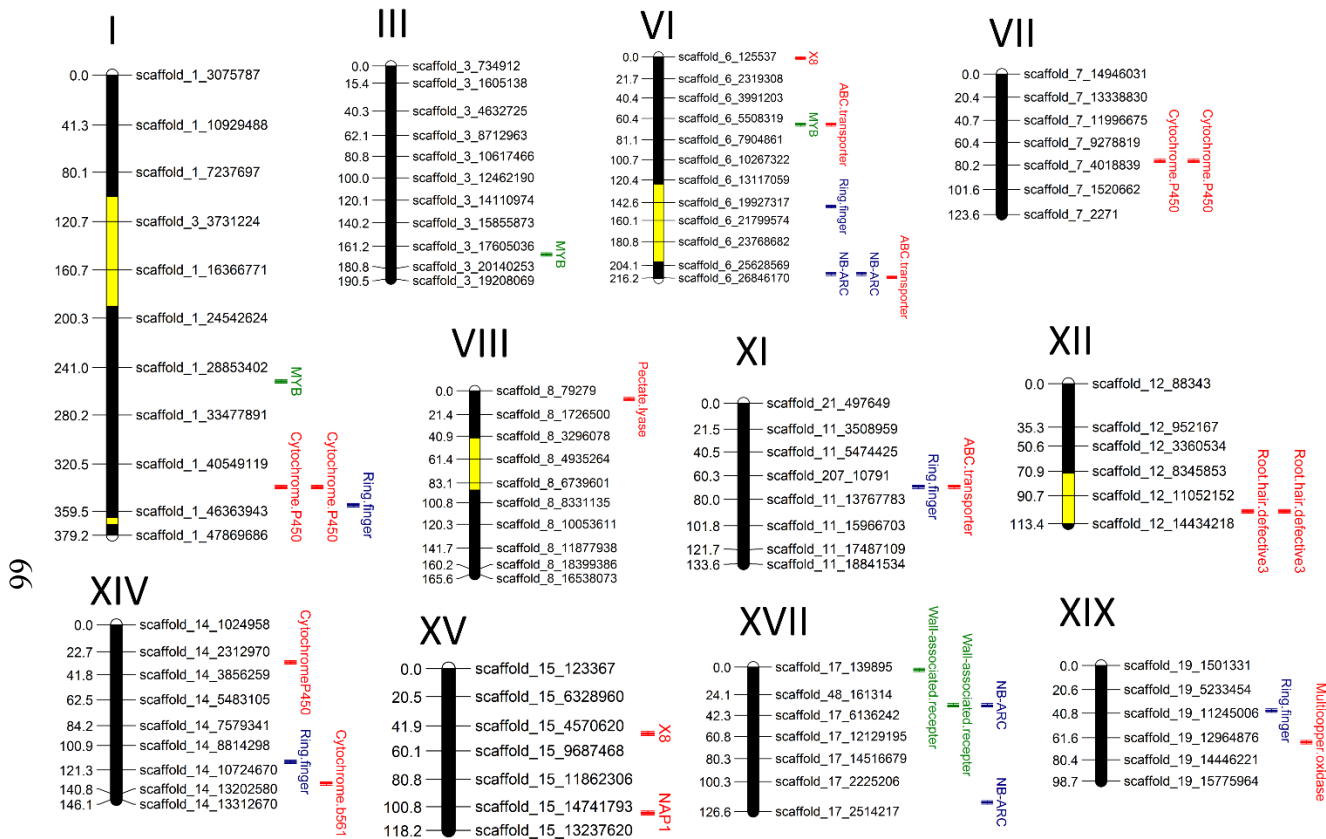
Figure 2 Location of top candidate genes on the linkage map derived from a cross of a female hybrid clone (*Populus trichocarpa* x *P. deltoides*) with *P. deltoides* (Bdeir et al. 2017). The linkage map is anchored to the genome of *Populus trichocarpa* (V3 assembly, Kelleher et al. 2007), see Table 1. QTL intervals for bark texture are shown in yellow based on Bdeir et al. (2017). Vertical lines show the location of top candidate genes on the *P. trichocarpa* genome detected near significant and highly reproducible SNPs in the genome-wide association mapping analysis. Red, green and blue genes are putatively involved in radial growth and tissue differentiation, suberin accumulation and programmed cell death, respectively.

Table 1 The 41 genomic regions with reproducible SNPs most significantly associated with bark texture in the genome-wide association study (GWAS).

Region #	LG	SNP position interval	Interval (kb)	Population	SNP position	Corr -log10 P
1	XV	13756415-13759891	3	PLCV14_BT2	13,759,891	**8.44**
				CLA12_BT2	13,756,415	5.98
				CLA13_BT2	13,756,415	5.22
2	XVII	2930303-2942559	12	CLA12_BT3	2,931,167	**7.89**
				CLA12cookieBT1	2,931,167	**6.70**
				CLA13_BT1	2,941,148	5.07
3	V	16,609,078	1	CLA13_BT1	16,609,078	**7.86**
				CLA12_BT1	16,609,078	**7.74**
4	XIV	10,543,886	1	CLA12_BT2	10,543,886	**7.74**
				CLA13_BT2	10,543,886	**6.12**
5	I	48502337-48513466	11	CLA12_BT1	48,504,159	**7.56**
				CLA12_BT3	48,504,159	5.40
				CLA13_BT2	48,504,159	5.68
6	XIV	3,085,288	1	CLA12_BT3	3,085,288	**7.50**
				CLA12_BT1	3,085,288	**7.09**
				CLA12cookieBT1	3,085,288	**6.69**
				CLA13_BT1	3,085,288	**6.48**
7	I	44916073-44932416	16	CLA12_BT2	44,931,643	**7.40**
				CLA13_BT2	44,931,643	**6.44**
8	XVII	2894584-2915048	20	CLA12_BT3	2,905,659	**7.12**
				CLA12cookieBT1	2,905,659	**6.72**

				CLA13_BT1	2,905,659	5.39
				CLA13_BT2	2,900,788	5.32
				CLA12_BT1	2,905,659	5.30
				PLCV14_BT1	2,915,048	5.30
9	VIII	11209007-11217624	9	CLA13_BT1	11,209,437	**7.08**
				CLA12_BT1	11,209,437	5.95
10	XI	11,521,810	1	CLA13_BT1	11,521,810	**7.02**
				CLA12_BT1	11,521,810	**6.55**
11	VI	21529197-21533794	5	CLA12_BT2	21,530,675	**7.00**
				CLA13_BT2	21,530,675	**6.41**
				CLA15	21,530,272	5.24
12	XVI	5614771-5630821	16	CLA12_BT2	5,626,773	**6.99**
				CLA13_BT2	5,626,773	5.73
				PLCV14_BT1	5,628,967	5.77
13	VI	543,317	1	CLA12_BT2	543,317	**6.91**
				CLA13_BT2	543,317	**6.80**
14	X	8131399-8141525	10	CLA12cookieBT1	8,131,399	**6.75**
				CLA12_BT3	8,131,399	**6.24**
				CLA13_BT1	8,141,525	5.33
15	VII	16608390-16609078	1	CLA13_BT2	11,609,206	**6.76**
				CLA12_BT2	11,609,206	5.40
16	IV	21617052-21619943	3	CLA13_BT2	21,619,943	**6.73**
				CLA12_BT1	21,619,943	5.86
17	V	21642667-21646857	4	CLA12_BT3	21,644,588	**6.73**

				CLA12_BT1	21,644,588	**6.00**
				CLA13_BT1	21,644,588	5.53
				CLA12cookieBT1	21,642,667	5.44
18	XVII	15160420-15163613	3	CLA13_BT1	15,160,420	**6.49**
				CLA12_BT1	15,160,420	5.45
19	XVIII	6040783-6046262	6	CLA12_BT1	6,042,491	**6.47**
				CLA13_BT1	6,042,491	5.01
				COR13_BT2	27,839,980	**6.31**
				COR13_BT1	27,837,548	5.38
21	IV	8874775-8883033	8	CLA12_BT2	8,877,397	**6.29**
				CLA13_BT2	8,877,397	5.99
22	XV	3,872,558	1	CLA12_BT1	3,872,558	**6.27**
				CLA13_BT1	3,872,558	5.14
23	XIV	7,277,970	1	CLA12_BT2	7,277,970	**6.26**
				CLA13_BT2	7,277,970	5.01
24	V	2,047,117	1	CLA13_BT1	2,047,117	**6.25**
				CLA12_BT1	2,047,117	5.77
25	XVI	889,154	1	CLA12_BT1	889,154	**6.25**
				CLA13_BT1	889,154	5.39
26	XII	2,115,668	1	CLA13_BT2	2,115,668	**6.19**
				CLA12_BT2	2,115,668	**6.01**
27	XI	15754964-15758609	4	CLA13_BT1	15,758,609	**6.13**
				CLA12_BT2	15,754,964	5.24
28	XIII	1,649,986	1	CLA13_BT1	1,649,986	**6.08**
				CLA12_BT1	1,649,986	5.57

29	VIII	10028445-10032564	4	CLA13_BT2	10,028,892	**6.00**
				CLA12cookieBT1	10,028,445	5.27
30	I	31275502-31277465	2	CLA12_BT2	31,276,103	**7.49**
31		3001881-3008293	7	CLA12cookieBT1	3,007,903	**7.73**
				CLA12_BT3	3,007,903	**6.66**
32	VII	10,801,084	1	CLA12cookieBT1	10,801,084	**7.46**
				CLA12_BT3	10,801,084	**6.30**
33	XIX	4752983-4762402	10	CLA12_BT3	4,762,402	**7.44**
				CLA12cookieBT1	4,762,402	5.42
34	XI	10996808-10996815	1	CLA12_BT3	10,996,815	**7.26**
				CLA12cookieBT1	10,996,815	**6.43**
35	XIX	12020500-12025588	5	CLA12_BT3	12,025,588	**7.25**
				CLA12cookieBT1	12,025,588	**6.60**
36	VI	13,299,251	1	CLA12cookieBT1	13,299,251	**6.63**
				CLA12_BT3	13,299,251	**6.20**
				CLA12_BT1	13,299,251	5.44
37	XVI	12838532-12843036	5	COR13_BT2	12,838,536	**6.36**
				COR13_BT1	12,843,036	5.19
38	XVI	2627146-2630738	3	COR13_BT3	2,627,146	**6.27**
				COR13_BT2	2,630,738	5.14
39	XI	1,779,913	1	CLA12cookieBT1	1,779,913	**6.14**
				CLA12_BT3	1,779,913	5.49
40	II	15467058-15489644	22	COR13_BT2	15,469,012	**6.13**

				COR13_BT1	15,482,878	5.53
41	XVI	9,570,806	1	CLA12cookieBT1	9,570,806	**6.01**
				CLA12_BT3	9,570,806	5.97

LG: linkage group. Sites: CLA: Clatskanie, COR: Corvallis, PLC: Placerville. Populations: the name indicates site, year and replicate (e.g. CLA12_BT3, Clatskanie, year 2012, bark texture replicate three). The first 29 genomic regions with highly significant SNPs are found reproducible across sites, while the remaining 12 regions are reproducible within a site. For each genomic region and experiment the most significant associations are shown. Highly significant SNPs are in bold (-log10 P ≥ 6). The interval in which significant SNPs were detected at a significance threshold of -log10 P ≥ 5, the populations and replicates in which SNPs were detected in corresponding genomic regions along with the highest corrected p-value are shown. Detailed information on SNPs and their genomic location can be found in Table S4.

Table 2 Genomic location and putative functions SNP containing genes and of genes adjacent to SNPs associated with bark texture identified by genome-wide association study (GWAS).

Region #	Population	Highest - log10 P	SNP within gene feature	Gene Name	Putative function	Arabidopsis ortholog
1	CLA12, CLA13, PLC14	8.44	Intron	Potri.015G124000	**Membrane-associated apoptosis protein, NAP1**	AT2G35110
1	CLA12, CLA13, PLC14	8.44	Intron, Upstream/Downstream	Potri.015G124100	**Protein kinase domain**	AT4G11480
31	CLA12	7.74	CDS, NS, Downstream	Potri.017G035500	**Protein kinase domain**	AT5G38260
8	CLA12, CLA13, PLC14	7.13	CDS, S, Upstream/downstream	Potri.017G034500	**Protein kinase domain**	AT3G55550
8	CLA12, CLA13, PLC14	7.13	Upstream/Downstream	Potri.017G034100	Protein tyrosine kinase	AT4G04490
20	PLC14, COR13	6.34	Downstream	Potri.006G279900	Protein tyrosine kinase	AT4G31170
20	PLC14, COR13	6.34	3' UTR	Potri.006G280000	**Protein tyrosine kinase**	AT2G24360
NA	COR13, CLA15	5.87	Downstream, intergenic	Potri.014G161400	Protein kinase domain	AT5G21222
2	CLA12, CLA13	7.90	CDS, S, Upstream/Downstream	Potri.017G034700	**Wall-associated receptor kinase galacturonan-binding**	None
NA	CLA12, CLA13	5.73	Intron	Potri.017G035400	**Wall-associated receptor kinase galacturonan-binding**	AT3G55550
3	CLA12, CLA13	7.86	Intron	Potri.005G162300	**VQ motif**	AT1G78310
4	CLA12, CLA13	7.75	Upstream, intergenic	Potri.014G138000	Zinc finger, C3HC4 type (RING finger), ubiquitin ligase	AT1G63900

5	CLA12, CLA13	7.56	Upstream, intergenic	Potri.001G450700	Ring finger domain, ubiquitin ligase	None
10	CLA12, CLA13	7.03	Upstream, intergenic	Potri.011G094800	Ring finger domain	AT5G55970
11	CLA12, CLA13, CLA15	7.00	Downstream, intergenic	Potri.006G200200	Ring finger domain, ubiquitin ligase	AT5G22000
NA	PLC14, COR13	5.78	Intron	Potri.019G060900	Zinc finger, C3HC4 type (RING finger), DNA helicase	AT2G40770
NA	CLA12, PLC14	5.26	Upstream, intergenic	Potri.009G080500	RING-type zinc-finger, ubiquitin ligase	None
5	CLA12, CLA13	7.56	Intron	Potri.001G450800	**Ribosomal protein L11, RNA binding domain**	AT1G32990
6	CLA12, CLA13	7.51	Upstream, intergenic	Potri.014G037800	Cytochrome P450	AT5G25180
32	CLA12	7.47	Upstream, intergenic	Potri.007G083300	Cytochrome P450	AT2G24180
7	CLA12, CLA13	7.41	Upstream, intergenic	Potri.001G424300	Cytochrome P450	None
7	CLA12, CLA13	7.41	Downstream, intergenic	Potri.001G424400	Cytochrome P450	None
NA	COR13, CLA15	5.87	Upstream, intergenic	Potri.014G161700	Eukaryotic cytochrome b561	None
NA	CLA12	5.35	Downstream, intergenic	Potri.007G083700	Cytochrome P450	AT2G24180
30	CLA12	7.49	Downstream, intergenic	Potri.001G309100	Myb/SANT-like DNA-binding domain	AT1G76880
NA	CLA12, CLA13	5.45	Downstream, intergenic	Potri.003G200000	Myb/SANT-like DNA-binding domain	AT3G14180
NA	CLA12	5.98	Downstream, intergenic	Potri.006G085900	Myb-like DNA-binding domain	AT2G37630

|						
7	CLA12, CLA13	7.41	Intron	Potri.001G424600	**EAP30/Vps36 family, RNA polymerase II transcription factor complex**	AT4G27040
7	CLA12, CLA13	7.41	5' UTR	Potri.001G424700	**Kua-ubiquitin conjugating enzyme hybrid localisation domain,**	AT4G27030
35	CLA12	7.26	CDS, NS	Potri.019G088500	**Multicopper oxidase**	AT5G58910
8	CLA12, CLA13, PLC14	7.13	Upstream/Downstream	Potri.017G034400	NB-ARC domain, Leucine rich repeat	None
18	CLA12, CLA13, COR13	6.50	CDS, NS	Potri.017G144100	**NB-ARC domain, Leucine rich repeat**	None
NA	CLA12, CLA13	5.61	Downstream	Potri.006G273900	NB-ARC domain, Leucine rich repeat	None
NA	CLA12, CLA13	5.61	Intron	Potri.006G274000	**NB-ARC domain, Leucine rich repeat**	AT5G06040
8	CLA12, CLA13, PLC14	7.13	CDS, S and NS, Upstream/Downstream	Potri.017G034200	**Leucine rich repeat**	None
19	CLA12, CLA13	6.47	Upstream, intergenic	Potri.018G057100	Leucine rich repeat N-terminal domain	AT5G25930
NA	CLA12, CLA13	5.73	Downstream, intergenic	Potri.017G035300	Leucine rich repeat, Apoptotic ATPase	None
NA	CLA12, CLA13	5.30	Downstream, intergenic	Potri.003G027200	Leucine Rich Repeat	None
NA	CLA12, CLA13	5.06	Upstream, intergenic	Potri.002G019900	Leucine rich repeat N-terminal domain	AT1G75820
8	CLA12, CLA13, PLC14	7.13	CDS, S and Downstream	Potri.017G033900	**BED zinc finger, Leucine rich repeat, Apoptotic ATPase**	None

8	CLA12, CLA13, PLC14	7.13	Downstream	Potri. 017G033800	Potential DNA-binding domain	None
10	CLA12, CLA13	7.03	Downstream, intergenic	Potri.011G094700	ABC transporter	AT5G60740
NA	CLA12	5.98	Upstream, intergenic	Potri.006G086000	ABC transporter	AT5G02270
NA	COR13	5.69	Downstream, intergenic	Potri.006G276600	ABC1 family	AT5G24810
12	CLA12, CLA13, PLC14	6.99	Downstream, intergenic	Potri.016G074200	TCP family transcription factor	None
12	CLA12, CLA13, PLC14	6.99	Downstream, intergenic	Potri.016G074400	PROTEASE S28 PRO-X CARBOXYPEPTIDASE-RELATED	None
13	CLA12, CLA13	6.92	Upstream	Potri.006G008200	X8 domain	None
16	CLA12, CLA13	6.73	Intron and 3' UTR	Potri.004G207300	**EF-hand domain pair**	None
14	CLA12, CLA13	6.75	Upstream/Downstream, intergenic	Potri.010G050900	WD domain, G-beta repeat	AT1G24530
36	CLA12	6.64	Intron	Potri.006G151800	**WD domain, G-beta repeat**	AT2G19520
22	CLA12, CLA13	6.28	Upstream, intergenic	Potri.015G041200	WD domain, G-beta repeat	AT3G18060
16	CLA12, CLA13	6.73	5' UTR and Intron	Potri.004G207400	**PEPTIDASE M20 FAMILY MEMBER**	AT4G38225
20	PLC14, COR13	6.34	CDA, NS and Intron	Potri.006G280200	**VPRBP PROTEIN-RELATED**	AT4G31160
20	PLC14, COR13	6.34	CDS, NS	Potri.006G280100	**Amino acid kinase family**	AT3G57560

23	CLA12, CLA13	6.27	Intron	Potri.014G093300	**Phosphatidylinositol-4-phosphate 5-Kinase**	AT2G26420
39	CLA12	6.14	Intron	Potri.011G020700	**Ankyrin repeats (3 copies), 26S proteasome regulatory complex, subunit PSMD10**	None
27	CLA12, CLA13	6.13	Downstream, intergenic	Potri.011G132700	Breast carcinoma amplified sequence 3	AT3G20720
27	CLA12, CLA13	6.13	Upstream, intergenic	Potri.011G132800	N-terminal region of Chorein, a TM vesicle-mediated sorter	AT5G54730
29	CLA12, CLA13	6.01	Upstream, intergenic	Potri.008G148800	Pectate lyase	AT3G01270

Genes are sorted according to highest -log10 values within family/description. Sites: CLA: Clatskanie, COR: Corvallis, PLC: Placerville. The putative function of SNP containing genes is shown in bold. Populations: the name indicates site and year. CDS: coding sequences, S: synonymous, NS: nonsynonymous. NA: not applicable since these reproducible regions (-log10 of $\geq$ 5) are not part of the 41 highly significant regions (-log10 of $\geq$ 6) but these genes from the same putative family/description are added. The genomic location of genes is given by the gene name (e.g., Potri.015G124000 is located on linkage group 15). Detailed information can be found in Table S4.

Supplementary Figure 1 Bark texture scale from smooth (1), medium flaky (2) to rough flaky (3) for the *P. trichocarpa* populations.
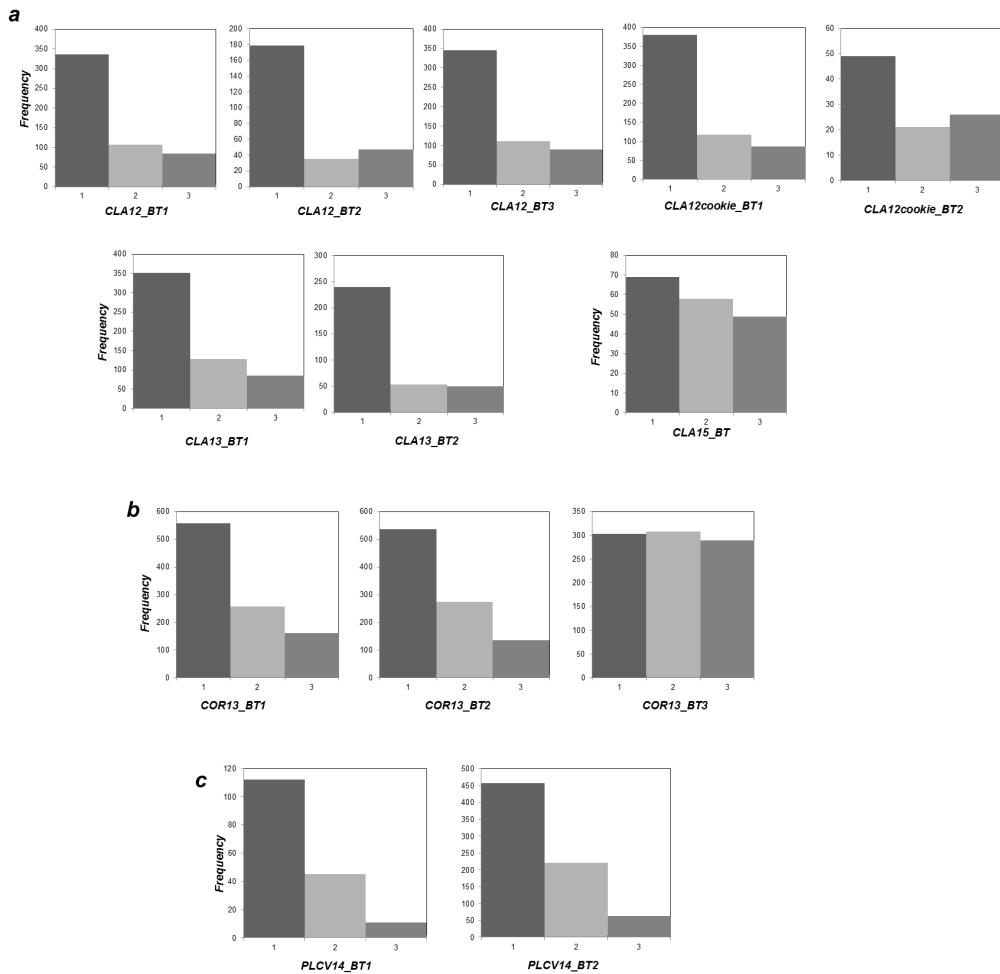
Supplementary Figure 2 Frequency distribution for bark texture across Clatskanie (CLA), Corvallis (COR) and Placerville (PLC) sites and various years (a, b, c, respectively).

Supplementary Table 1 Spearman's rank correlation coefficients for bark texture phenotypes across five datasets.

		COR13			CLA12			CLA13		CLA15	PLC14
	Replicate	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**1**	**1**
COR13	1	x									
	2	**0.3523**	x								
	3	**0.4441**	**0.4116**	x							
CLA12	1	**0.4048**	**0.3740**	**0.4893**	x						
	2	**0.4080**	**0.4172**	**0.4383**	**0.9051**	x					
	3	**0.3804**	**0.3213**	**0.4302**	**0.7213**	0.8826	x				
CLA13	1	**0.3483**	**0.3541**	**0.4956**	**0.9581**	**0.6973**	**0.6255**	x			
	2	**0.3842**	**0.3672**	**0.4025**	**0.7822**	**0.9783**	**0.6670**	**0.5583**	x		
CLA15	1	**-0.3775**	**-0.4129**	**-0.5250**	-0.4485*	0.0ns	-0.2522*	**-0.6271**	-0.4553*	x	
PLC14	1	0.0431ns	**0.2897**	0.1122ns	0.3098*	0.2953*	0.3423*	0.2821*	0.2456*	0.1398ns	x
	2	**0.2248**	**0.2067**	**0.2160**	**0.2569**	**0.3126**	**0.2698**	**0.2316**	**0.2555**	-0.0425ns	**0.3634**

Three different sites including Corvallis (COR), Clatskanie (CLA) and Placerville (PLC), of which CLA datasets were collected across three years. Statistically significant correlations across different environments are shown in bold ($p \leq 0.0001$), * $p \leq 0.05$, ns = not significant.

Supplementary Table 2[5] All 1825 SNPs associated with bark texture and the underlying genes within a 6 kb interval. The population in which these SNP were detected in, their position, linkage group and significance score along with the physical localization, annotation and expression profiles of the underlying genes are summarized.

Supplementary Table 3[5] The 50 genomic regions most strongly associated with bark texture and the underlying genes. The population in which these 129 individual SNPs were detected, position, linkage group and significance score as well as the physical localization and annotation of the underlying genes are summarized. SNPs are categorized into three groups: reproducible across sites, reproducible within a site and reproducible but no putative gene detected within 50 kb interval. Scores of -log10 P ≥ 6 are in bold, genes with SNPs within its genomic sequence are highlighted in yellow.

---

[5] Supplementary Table 2 and 3 are in excel format submitted as digital files.

Chapter 4[6]: QTL mapping of stomata density in hybrid *Populus*

Abstract

Stomatal development is known to be regulated by various genetic and environmental factors to optimize gas exchange. Stomata in the epidermal tissues of leaves are valves that passes CO_2, and thus they influence the global carbon cycle, and during leaf development, react to the fluctuations on temperature, humidity and other climate changes. Genotypic variation in stomatal density has been reported but little is known of the genetic mechanisms behind these leaf traits. With the genome sequenced for the model tree *Populus trichocarpa*, identifying quantitative trait loci (QTL) for stomatal density will pinpoint polymorphisms linked to genes and provide a foundation for future work. In the present study, a hybrid pseudo-backcross poplar family was used to map QTL for stomatal density across two different environments and three years. A total of eighteen QTL were detected in eleven linkage groups, which explained 4.6% to 15.6 % of the phenotypic variance. Four major genomic regions, with reproducible and consistent QTL clusters associated with stomatal density, were found on linkage groups II, III, XVI and XIX. Further study is needed to fully understand the mechanism underlying the observed genetic association and to elucidate the function of the QTLs involved.

---

[6] The material contained in this chapter is to be submitted to Tree Genetics and Genomics.

Introduction

In plants, stomata, bounded by a pair of guard cells, control gas exchange such as water vapor and CO_2 between the plant and its environment, providing an active contribution to photosynthesis and transpiration [1]. By sensing different exogenous signals, the stomatal complex balances CO_2 uptake versus water loss under continually changing environmental conditions [2, 3]. Breeders are continuously working to improve crop productivity and stress tolerance. Knowledge of the underlying molecular genetic basis of traits related to plant productivity and adaptation, such as stomatal development, can assist breeding efforts [4].

Fossil leaves records of the stomatal density (SD) spanning the past 400 Myr supported the predicted changes in atsmopheric CO_2, as such provides key information in light of predicted changes in precipitation and temperature, as well as the CO_2 environment, during leaf development [3, 5, 6], which can impact biomass production [7]. SD varies among individual leaves, among individuals within and among populations and among plant species [8]. Moreover, the regulation of stomatal numbers is controlled by environmental signals including CO_2 concentration, light intensity, temperature and water deficit [9-11]. For instance, increased concentrations of atmospheric CO_2 result in the reduction of stomatal density [12-14]. So far, a handful of genes have been found to be involved in stomatal development [15-19]. Thus, the knowledge of signaling pathways by which environmental signals control stomatal development remains incomplete. Alternatively, some authors suggest that stomatal development is controlled by more complex scale-free networks [3, 20].

As scale-free networks are composed of many interconnected nodes, they are robust and exhibit a high degree of tolerance to node removal but are also fragile to hub removal [21]. Both cases are found in stomatal development for such its signaling is suggested to be organized as a scale-free network [3]. For example, recent data revealed that guard cells continue to develop regardless of the loss of certain signaling components, such as mutations in the blue light receptor genes resulting in stomata fail to respond to blue light [22, 23], as such exhibiting high tolerance to nodal removal. However, other essential components, including ABA, extracellular calcium ion, hydrogen peroxide, CO_2 and IAA signaling, can have a direct effect and cause a

significant increase in guard cell number [2, 24-29]. Thus, many genes and pathways are involved in stomatal development. Through QTL (Quantitative Trait Locus) mapping, the underlying genomic regions that contribute to complex traits such as stomatal number can be identified, and their effect quantified. With the availability of molecular markers, sequence data and genetic linkage maps, QTL analysis has been critical to uncover the genetic architecture of complex, adaptive traits.

The genus *Populus* L. has been adopted as a model system in forest genetics and tree physiology [30-32]. Due to the variation in adaptive traits within various poplar species, countless opportunities are available to study and isolate underlying genes associated with ecologically and/or economically important traits [32, 33]. QTL mapping has been applied to analyze the genetic variation underlying stomatal density for several species, including *Arabidopsis thaliana*, *Quercus robur*, *Oryza sativa* and *Vicia faba* [6, 34-36]. In this study, we apply the QTL mapping method to identify chromosomal regions controlling stomatal density in a hybrid pseudo-backcross poplar family for the purpose of identifying underlying candidate genes for major QTL associated with stomatal density.

Materials and methods

Mapping population and plant material

We utilized a hybrid poplar pedigree (Family 52-124) composed of 396 genotypes was generated by crossing the hybrid female clone 52-225 (TD), an F_1 hybrid derived from a cross between *P. trichocarpa* (TT, clone 93-968) and *P. deltoides* (DD, clone ILL-101), with *P. deltoides* (DD, clone D124) [37, 38].

One vegetatively propagated rooted cutting from each genotype of the mapping population was established in one-gallon pots in a greenhouse at Michigan Technological University in the fall of 2009 (Drost et al. 2015). Pots were moved outside the greenhouse in the fall of 2010. On June 2, 2011, 197 genotypes were planted at a field site (2 m x 2 m spacing) at the Ford Forestry Center of Michigan Technological University in Alberta village (Michigan, 46°38′37″N, 88°28′46″W). Two additional biological replicates of each genotype in the mapping pedigree were established between

February 22 and 26, 2016 in a greenhouse at Michigan Technological University, in one-gallon pots.

For stomata measurements, three leaves per plant grown in the greenhouse were analyzed in August 2010 from 90 genotypes. In 2016, leaves collected from two replicates of 150 genotypes grown in greenhouse were analyzed. Finally, in 2017, leaves collected from one replicate of 150 genotype grown in the greenhouse was analyzed.

Stomatal imprints and density measurements

Healthy leaves of the first seasonal flush were collected, pressed and dried. The leaves were covered with a thin layer of clear nail polish on the lower epidermis between the second and third vain, covering an oval spot, and then were left to dry. A leaf imprint was obtained with a strip of clear tape and the imprint was transferred to a microscope slide. Stomatal density replicas were counted on microphotographs obtained with the aid of a light confocal microscope at 20X magnification from a 10 µm in diameter sections of the leaf imprint.

To assess correlations among replicates and between years, pair-wise phenotypic correlations were calculated as Pearson correlation coefficients for stomatal density across the three years using WinSTAT [39] (Table S1).

Genetic linkage map

We used the genetic map of the mapping family 52-124 comprised of 3,568 SNP markers with known genomic positions for QTL identification. SNP genotyping, marker curation and genetic map construction were previously described by Muchero et al. 2015 and Bdeir et al. 2017 [37, 40].

QTL analysis

The data had no outliers, recording errors were corrected or deleted. Each dataset was checked for normal distribution using WinSTAT [39]. Transformations were deemed unnecessary (Figure S1). The data were analyzed for the presence of QTL using the MapQTL6 software [41]. To map QTL intervals on the genetic linkage map and to test for reproducibility across years and environments, the interval mapping method was

used. Furthermore, composite interval mapping with the Multiple-QTL Model (MQM) method was applied to further refine the QTL regions. The putative QTL were subjected to 1,000 genome-wide (GW) and chromosome-wide (CW) permutations to determine LOD significance thresholds at the 0.05 significance level [42]. These parameters were used for declaring the existence of a significant QTL.

Candidate genes

Since QTL regions are genome-anchored, the underlying genes in the intervals were identified from the *Populus* genome assembly V3.0 [43] in the Phytozome database using the BioMart tool (https://phytozome.jgi.doe.gov). Because of large QTL intervals on some of the linkage groups, underlying genes within intervals defined by MQM mapping were identified for the eighteen QTL. Intervals spanning the genomic regions (physical location by MQM mapping) summarized in Table 1 were used to identify the putative candidate genes. The Gene Ontology (GO) of the genes underlying the LOD maxima were found using the *A. thaliana* gene annotation and AgriGO with $p \leq 0.05$ [44].

Results

Repeatability of stomatal density measurements

Correlations among replicates and years are shown in Table S1. Stomatal density showed a high correlation between replicates and across sites and years. The phenotypic correlations between replicates within the same year are as high as r=0.98 for different leaves of the same plant in samples of 2010 and around r=0.96 for clonal replicates in 2016 ($p \leq 0.0001$ for all comparisons). High correlations were also observed for comparisons among years (r=0.81 to 0.96, $p \leq 0.0001$).

Frequency distributions of stomatal densities in all years are shown in Figure S1. All datasets follow the pattern of normally distributed data.

QTL for stomatal density

Twenty-four individual QTL were detected for stomatal density on twelve linkage groups (Table 1), of which eight had LOD scores above the GW threshold. For all datasets, QTL were anchored to the *Populus* genome assembly (Fig. 1, Table 1). For QTL

found on linkage group II, VI, VII, XII, XV, XVIII and XIX, the *P. deltoides* genotype DD was associated with the higher value for stomatal density, while for QTL on linkage group III, VIII, X, XIII, XIV and XVI, the DT genotype was associated with the higher value of stomatal density (Table 1). The percentage of phenotypic variance explained (PVE) by individual QTL ranged from 4.6 to 15.6%. Figure 2 shows a graphical outline of the LOD score profiles for all individual QTL mapped on linkage groups after cofactor selection. Specific genomic regions with significant QTL including LG II, III, VIII, XII, XIV, XVI and XIX also displayed a suggestive QTL with low LOD score not reaching the CW threshold representing different replicates/years.

Based on reproducibility across replicates and years, four QTL clusters stood out on LG II, III, XVI and XIX representing thirteen individual QTL that were reproducible across at least two replicates and/or years with overlapping intervals and LOD maxima within proximity (Fig. 2, Table 1). In the first cluster on LG II, two QTL with LOD maxima between 154-173 cM were detected for one replicate in 2010 and for one clonal replicate in 2016. These QTL explain a comparatively high percentage of the total phenotypic variance (4.6% and 14.5%, respectively) and one of these two QTL was significant above the GW threshold. A suggestive peak was also detected in 2017. Five QTL found on LG III with LOD maxima within an interval between 75-97 cM were detected across all three years, the two clonal replicates in 2016, and in one replicate and their average in 2010, while a suggestive QTL below the CW threshold (LOD score=1.8) was detected between 73-95 cM for one replicate (leaf a) in 2010 (Fig. 2). The phenotypic variance explained by these QTL ranged between 5.5 and 12.7% and the QTL for replicates a in 2016 and replicate m in 2010 were significant at the GW level. The third QTL cluster consisting of three QTL were detected on LG XVI in 2010 and 2016 with LOD maxima at 20 cM. For 2016, one replicate (leaf a) showed a QTL significant at the CW threshold (PVE 6.2%), while for 2010, all replicates revealed a peak within same genomic regions. Specifically, replicates a and m had QTL significant at the GW threshold (PVE 12.3% and 11%) while the other two replicates of year 2010 (leaf b and c) showed a suggestive QTL within the same interval. Lastly, in the fourth QTL cluster on LG XIX an overlapping QTL interval was found between 0-40 cM for two replicates (leaf a, c and m) in 2010, while the third replicate (leaf b) showed a suggestive QTL

within the same interval (Fig. 2). Though only one replicate was significant at the GW threshold, both QTL explain a high percentage of the total phenotypic variance (11.0% and 15.6%).

On LG XIII and XVIII, they both showed reproducibility across 2010 replicates, but the four QTL were significant at the CW threshold. The remaining 7 individual QTL were significant but not reproducible across replicates or years (Fig. 2, Table 1). QTL detected in year 2016 on LG VIII and another QTL on LG XIV, while both had high LOD score and PVE values (5.3 and 8.9%, respectively), their LOD maxima were at different positions (around 140 cM and 1 cM, respectively, Table 1). They both also showed one suggestive QTL within their interval across different replicates/years (Fig. 2). The third significant QTL (11SD10b, Table 1) on LG XII having high LOD and PVE scores (4.52 and 15.2%, respectively) overlapped with three suggestive QTL from replicates in years 2010, 2016 and 2017 (Fig. 2).

Candidate genes

The interval after MQM mapping with cofactor selection was used to identify putative genes and the number of underlying each QTL is listed in Table 2. A total of 869 genes within genome-anchored QTL intervals for stomatal density were detected, out of which 617 (71%) had annotations based on the InterPro domain; the function of the remaining genes is unknown. A detailed list of the putative genes with their annotations is given in Table S2.

Significantly overexpressed GO terms

To enable a functional categorization and to identify biological pathways of the genes detected, Gene Ontology (GO) term enrichment analyses were conducted. Figure S2 is a visual representation of all the underlying genes found within the eighteen QTL intervals and of significantly overrepresented GO terms in biological process categories. Figure S3 lists the significant levels of the represented GO terms.

Discussion

With the aid of QTL mapping, we can pinpoint the genomic regions associated with a quantitative trait and the polymorphisms that are linked to functional genes. Several reviews stated the advantages of using QTL analysis to reveal the genetic architecture especially for complex traits [45-47]. With the availability of the *Populus* genome, candidate genes in QTL intervals can be identified. In this study, we analyzed stomatal density in a hybrid *Populus* progeny using QTL analysis and identified major QTL responsible for stomatal density variation. It has been generally reported that stomatal density is controlled by both environmental and genetic factors, suggesting the involvement of many genes [48]. The signaling pathways and mechanisms that regulate guard cell function have been intensively studied (reviewed in [49, 50]). Furthermore, environmental variables can modulate basal stomatal development pathways to adjust to the environmental conditions [48]. For example, the concentration of atmospheric CO2 directly influences stomatal development and has been used an indicator of paleoatmospheric CO2 levels and in the assessment of the ecological consequences of global change [3, 12-14]. Forty percent of the atmospheric CO2 passes through stomata [51], thus any variation in stomatal density ultimately influences photosynthesis and atmospheric CO2 concentration. Our results however, based on strong inter annual correlations among genotypes and the correlations among replicates within years, suggest that genetics plays a strong role in determining stomatal density in hybrid *Populus*.

It has been suggested that guard cell signaling reflects the organization of a scale-free network rather than a collection of linear pathways [3]. The estimated number and effect of genes involved in such a complex trait most likely linked to a scale-free network can be obtained through QTL mapping. As a quantitative trait, it was genetically analyzed using QTL mapping in both poplar and oaks [6, 52]. Estimates for the effect of genes controlling stomatal density have been obtained in the QTL study by Rae et al. (2006) using a third-generation hybrid population derived from two *Populus trichocarpa* and *P. deltoides* hybrids. Specifically, the hybridization of *P. trichocarpa* (clone 93-968, same clone used in our study) and *P. deltoides* (clone ILL-129) generated two hybrids, 53-246 and 53-242, which were crossed to generate an F_2 full-sib family (family 331). Rae et al (2006) performed QTL analyses under elevated and ambient CO_2 conditions

and found fourteen QTL for adaxial and abaxial stomatal density distributed over ten linkage groups with genetically determined total phenotypic variances for elevated and ambient CO_2 conditions between 19.5-22.1% and 9.3-19.7%, respectively.

In another study, using a hybrid poplar family of *Populus trichocarpa* × *P. deltoides* saplings, stomatal development was monitored under varied CO_2 concentrations, vapor pressure deficit (D), and irradiance (Q) around the young leaves [53]. Results indicated that changes in CO_2 concentration, light and humidity causing an increase in mature leaf stomatal conductance had a regulatory effect on stomatal development of expanding leaves [53]. A plastic response to changing environments has been shown in stomatal density in maize, *Arabidopsis* and poplar for which systemic signals from mature leaves regulate stomatal development of expanding leaves [20, 53, 54]. However, the nature of these signals generated in the mature leaves and transmitted to developing leaves is largely unknown and it is unclear how many signals are elicited [55]. Our study further suggests high genetic variation for and control of stomatal density in a single interspecific hybrid progeny, but also variation among years and replicates.

QTL for stomatal density

We detected twelve genomic regions on eleven linkage groups with a range of QTL effects explained from 4.6% to 15.6 % of the phenotypic variance. Four major QTL clusters with high reproducibility and consistency associated with stomatal density were found on linkage group II, III and consistently identified across all three years, while the clusters on linkage group XVI were found reproducible across two years. Finally, QTL clusters on linkage group XIX were consistently identified across all three replicates in year 2010. Since the identified genomic intervals are large and contain hundreds of candidate genes, it is still difficult to identify the genes in question. Thus, we seek to analyze the overrepresented GO terms in the QTL intervals.

GO analysis showed the functional profile of the underlying genes and their involvement in biological processes such as developmental processes, including cell morphogenesis and differentiation and post-embryonic development, and cellular processes, such as metabolic and cellular biosynthesis regulations (Figure S2). It has been demonstrated that mature leaves both detect CO_2 concentration levels and accordingly

signal to expanding leaves thus inducing stomatal development in young leaves [13, 20, 56]. Furthermore, numerous studies showed a relationship between stomatal traits (density, index and conductance) with leaf development (expansion, length and area) [38, 52, 57, 58].

Future Work

The natural variation of stomatal density is linked to plant fitness and adaptation to their environment, thus directly impacting speciation and evolutionary change [3]. The availability of the full annotated poplar sequence and associated genomic resources including microarrays will be essential to further study and enable research to be focused on understanding stomatal development and the genetic control of this trait. Future work on nucleotide variation in candidate genes controlling such complex adaptive traits as stomatal density under different environmental conditions such as water deficit and different CO_2 concentrations will allow a better understanding of the genetic and environmental control of this trait.

References

1. Willmer C, Fricker M: **Stomata.**, vol. 2: Springer Science & Business Media; 1996.
2. Schroeder JI, Allen GJ, Hugouvieux V, Kwak JM, Waner D: **Guard cell signal transduction.** *Annual Review of Plant Biology* 2001, **52**(1):627-658.
3. Hetherington AM, Woodward FI: **The role of stomata in sensing and driving environmental change.** *Nature* 2003, **424**(6951):901.
4. Cattivelli L, Rizza F, Badeck F-W, Mazzucotelli E, Mastrangelo AM, Francia E, Mare C, Tondelli A, Stanca AM: **Drought tolerance improvement in crop plants: an integrated view from breeding to genomics.** *Field Crops Research* 2008, **105**(1-2):1-14.
5. Beerling D, Woodward F: **Changes in land plant function over the Phanerozoic: reconstructions based on the fossil record.** *Botanical Journal of the Linnean Society* 1997, **124**(2):137-153.
6. Gailing O, Langnfeld-heyser R, Polle A, Finkeldey R: **Quantitative trait loci affecting stomatal density and growth in a *Quercus robur* progeny: implications for the adaptation to changing environments.** *Global Change Biology* 2008, **14**(8):1934-1946.
7. Al Afas N, Marron N, Ceulemans R: **Clonal variation in stomatal characteristics related to biomass production of 12 poplar (*Populus*) clones in a short rotation coppice culture.** *Environmental and Experimental Botany* 2006, **58**(1-3):279-286.
8. Schlüter U, Muschak M, Berger D, Altmann T: **Photosynthetic performance of an *Arabidopsis* mutant with elevated stomatal density (sdd1-1) under different light regimes.** *Journal of Experimental Botany* 2003, **54**(383):867-874.
9. Beerling DJ, Chaloner WG: **The impact of atmospheric CO_2 and temperature changes on stomatal density: observation from *Quercus robur* lammas leaves.** *Annals of Botany* 1993, **71**(3):231-235.
10. Gay A, Hurd R: **The influence of light on stomatal density in the tomato.** *New Phytologist* 1975, **75**(1):37-46.
11. Xu Z, Zhou G: **Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass.** *Journal of Experimental Botany* 2008, **59**(12):3317-3325.
12. Royer D: **Stomatal density and stomatal index as indicators of paleoatmospheric CO_2 concentration.** *Review of Palaeobotany and Palynology* 2001, **114**(1-2):1-28.
13. Lake J, Quick W, Beerling DJ, Woodward FI: **Plant development: signals from mature to new leaves.** *Nature* 2001, **411**(6834):154.
14. Woodward FI: **Stomatal numbers are sensitive to increases in CO_2 from pre-industrial levels.** *Nature* 1987, **327**(6123):617.
15. Doheny-Adams T, Hunt L, Franks PJ, Beerling DJ, Gray JE: **Genetic manipulation of stomatal density influences stomatal size, plant growth and tolerance to restricted water supply across a growth carbon dioxide gradient.** *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 2012, **367**(1588):547-555.
16. Yu H, Chen X, Hong Y-Y, Wang Y, Xu P, Ke S-D, Liu H-Y, Zhu J-K, Oliver DJ, Xiang C-B: **Activated expression of an *Arabidopsis* HD-START protein confers

drought tolerance with improved root system and reduced stomatal density. *The Plant Cell* 2008, **20**(4):1134-1151.

17.   Tripathi P, Rabara RC, Reese RN, Miller MA, Rohila JS, Subramanian S, Shen QJ, Morandi D, Bücking H, Shulaev V: **A toolbox of genes, proteins, metabolites and promoters for improving drought tolerance in soybean includes the metabolite coumestrol and stomatal development genes.** *BMC Genomics* 2016, **17**(1):102.

18.   Engineer CB, Ghassemian M, Anderson JC, Peck SC, Hu H, Schroeder JI: **Carbonic anhydrases, EPF2 and a novel protease mediate CO$_2$ control of stomatal development.** *Nature* 2014, **513**(7517):246.

19.   Lin G, Zhang L, Han Z, Yang X, Liu W, Li E, Chang J, Qi Y, Shpak ED, Chai J: **A receptor-like protein acts as a specificity switch for the regulation of stomatal development.** *Genes & Development* 2017.

20.   Coupe S, Palmer B, Lake J, Overy S, Oxborough K, Woodward F, Gray J, Quick W: **Systemic signalling of environmental cues in *Arabidopsis* leaves.** *Journal of Experimental Botany* 2005, **57**(2):329-341.

21.   Friedman A, Perrimon N: **Genetic screening for signal transduction in the era of network biology.** *Cell* 2007, **128**(2):225-231.

22.   Takemiya A, Shimazaki K-i: **Arabidopsis phot1 and phot2 phosphorylate BLUS1 kinase with different efficiencies in stomatal opening**. *Journal of plant research* 2016, **129**(2):167-174.

23.   Horrer D, Flütsch S, Pazmino D, Matthews JS, Thalmann M, Nigro A, Leonhardt N, Lawson T, Santelia D: **Blue light induces a distinct starch degradation pathway in guard cells for stomatal opening**. *Current Biology* 2016, **26**(3):362-370.

24.   Hetherington AM: **Guard cell signaling.** *Cell* 2001, **107**(6):711-714.

25.   Kinoshita T, Doi M, Suetsugu N, Kagawa T, Wada M, Shimazaki K-i: **Phot1 and phot2 mediate blue light regulation of stomatal opening.** *Nature* 2001, **414**(6864):656.

26.   Hunt L, Mills LN, Pical C, Leckie CP, Aitken FL, Kopka J, Mueller-Roeber B, McAinsh MR, Hetherington AM, Gray JE: **Phospholipase C is required for the control of stomatal aperture by ABA.** *The Plant Journal* 2003, **34**(1):47-55.

27.   Kim T-H, Böhmer M, Hu H, Nishimura N, Schroeder JI: **Guard cell signal transduction network: advances in understanding abscisic acid, CO$_2$, and Ca^{2+} signaling.** *Annual review of plant biology* 2010, **61**:561-591.

28.   Jiang K, Sorefan K, Deeks MJ, Bevan MW, Hussey PJ, Hetherington AM: **The ARP2/3 complex mediates guard cell actin reorganization and stomatal movement in *Arabidopsis*.** *The Plant Cell* 2012, **24**:2031–2040.

29.   Nagy R, Grob H, Weder B, Green P, Klein M, Frelet A, Schjoerring JK, Brearley CA, Martinoia E: **The *Arabidopsis* ATP-binding cassette protein ATMRP5/ATABCC5 is a high-affinity inositol hexakisphosphate transporter involved in guard cell signaling and phytate storage.** *Journal of Biological Chemistry* 2009:jbc. M109. 030247.

30.   Tuskan G, Wullschleger S, Bradshaw H, Dalhman R: **Sequencing the *Populus* genome: applications to the energy-related missions of DOE.** In: *Abstr Plant Anim Microbe Genomes Conf: 2002*.

31. Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A: **The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray).** *Science* 2006, **313**(5793):1596-1604.

32. Bradshaw H, Ceulemans R, Davis J, Stettler R: **Emerging model systems in plant biology: poplar (*Populus*) as a model forest tree.** *Journal of Plant Growth Regulation* 2000, **19**(3):306-313.

33. Tuskan G, DiFazio S, Teichmann T: **Poplar genomics is getting popular: the impact of the poplar genome project on tree research.** *Plant Biology* 2004, **6**(1):2-4.

34. Juenger TE, Mckay JK, Hausmann N, Keurentjes JJ, Sen S, Stowe KA, Dawson TE, Simms EL, Richards JH: **Identification and characterization of QTL underlying whole-plant physiology in *Arabidopsis thaliana*: δ13C, stomatal conductance and transpiration efficiency.** *Plant, Cell & Environment* 2005, **28**(6):697-708.

35. Laza MRC, Kondo M, Ideta O, Barlaan E, Imbe T: **Quantitative trait loci for stomatal density and size in lowland rice.** *Euphytica* 2010, **172**(2):149-158.

36. Khazaei H, O'Sullivan DM, Sillanpää MJ, Stoddard FL: **Use of synteny to identify candidate genes underlying QTL controlling stomatal traits in faba bean (*Vicia faba* L.).** *Theoretical and Applied Genetics* 2014, **127**(11):2371-2385.

37. Bdeir R, Muchero W, Yordanov Y, Tuskan GA, Busov V, Gailing O: **Quantitative trait locus mapping of *Populus* bark features and stem diameter.** *BMC Plant Biology* 2017, **17**(1):224.

38. Drost DR, Puranik S, Novaes E, Novaes CR, Dervinis C, Gailing O, Kirst M: **Genetical genomics of *Populus* leaf shape variation.** *BMC Plant Biology* 2015, **15**(1):166.

39. Fitch R: **WinSTAT for Excel**. *The statistics add-in for Microsoft Excel R Fitch Software* 2006.

40. Muchero W, Guo J, DiFazio SP, Chen J-G, Ranjan P, Slavov GT, Gunter LE, Jawdy S, Bryan AC, Sykes R: **High-resolution genetic mapping of allelic variants associated with cell wall chemistry in *Populus*.** *BMC Genomics* 2015, **16**(1):24.

41. Ooijen J, Kyazma B: **MapQTL 6.0, software for the mapping of quantitative trait loci in experimental populations of dihaploid species**. *Netherlands: Wageningen* 2009.

42. Storey JD, Tibshirani R: **Statistical significance for genomewide studies.** *Proceedings of the National Academy of Sciences* 2003, **100**(16):9440-9445.

43. Kelleher CT, Chiu R, Shin H, Bosdet IE, Krzywinski MI, Fjell CD, Wilkin J, Yin T, DiFazio SP, Ali J: **A physical map of the highly heterozygous Populus genome: integration with the genome sequence and genetic map and analysis of haplotype variation**. *The Plant Journal* 2007, **50**(6):1063-1078.

44. Du Z, Zhou X, Ling Y, Zhang Z, Su Z: **agriGO: a GO analysis toolkit for the agricultural community**. *Nucleic acids research* 2010, **38**(suppl_2):W64-W70.

45. Wayne ML, McIntyre LM: **Combining mapping and arraying: an approach to candidate gene identification.** *Proceedings of the National Academy of Sciences* 2002, **99**(23):14903-14906.

46. Asins M: **Present and future of quantitative trait locus analysis in plant breeding.** *Plant breeding* 2002, **121**(4):281-291.

47. Jansen RC, Nap J-P: **Genetical genomics: the added value from segregation.** *Trends in Genetics* 2001, **17**(7):388-391.

48. Casson SA, Hetherington AM: **Environmental regulation of stomatal development.** *Current opinion in plant biology* 2010, **13**(1):90-95.

49. Israelsson M, Siegel RS, Young J, Hashimoto M, Iba K, Schroeder JI: **Guard cell ABA and CO_2 signaling network updates and Ca^{2+} sensor priming hypothesis.** *Current opinion in plant biology* 2006, **9**(6):654-663.

50. Hetherington AM, Brownlee C: **The generation of Ca^{2+} signals in plants.** *Annu Rev Plant Biol* 2004, **55**:401-427.

51. Ciais P, Tans PP, Denning AS, Francey RJ, Trolier M, Meijer HA, White JW, Berry JA, Randall DA, Collatz GJ: **A three-dimensional synthesis study of δ18O in atmospheric CO_2: 2. Simulations with the TM2 transport model.** *Journal of Geophysical Research: Atmospheres* 1997, **102**(D5):5873-5883.

52. Rae A, Ferris R, Tallis M, Taylor G: **Elucidating genomic regions determining enhanced leaf growth and delayed senescence in elevated CO_2.** *Plant, Cell & Environment* 2006, **29**(9):1730-1741.

53. Miyazawa S-I, Livingston NJ, Turpin DH: **Stomatal development in new leaves is related to the stomatal conductance of mature leaves in poplar (*Populus trichocarpa× P. deltoides*).** *Journal of Experimental Botany* 2005, **57**(2):373-380.

54. Driscoll S, Prins A, Olmos E, Kunert K, Foyer C: **Specification of adaxial and abaxial stomata, epidermal structure and photosynthesis to CO_2 enrichment in maize leaves.** *Journal of Experimental Botany* 2005, **57**(2):381-390.

55. Bergmann DC, Sack FD: **Stomatal development.** *Annu Rev Plant Biol* 2007, **58**:163-181.

56. Lake JA, Woodward FI, Quick WP: **Long-distance CO_2 signalling in plants.** *Journal of Experimental Botany* 2002, **53**(367):183-193.

57. Ferris R, Long L, Bunn S, Robinson K, Bradshaw H, Rae A, Taylor G: **Leaf stomatal and epidermal cell development: identification of putative quantitative trait loci in relation to elevated carbon dioxide concentration in poplar.** *Tree Physiology* 2002, **22**(9):633-640.

58. Gailing O, Bodénès C, Finkeldey R, Kremer A, Plomion C: **Genetic mapping of EST-derived simple sequence repeats (EST-SSRs) to identify QTL for leaf morphological characters in a *Quercus robur* full-sib family**. *Tree genetics & genomes* 2013, **9**(5):1361-1367.
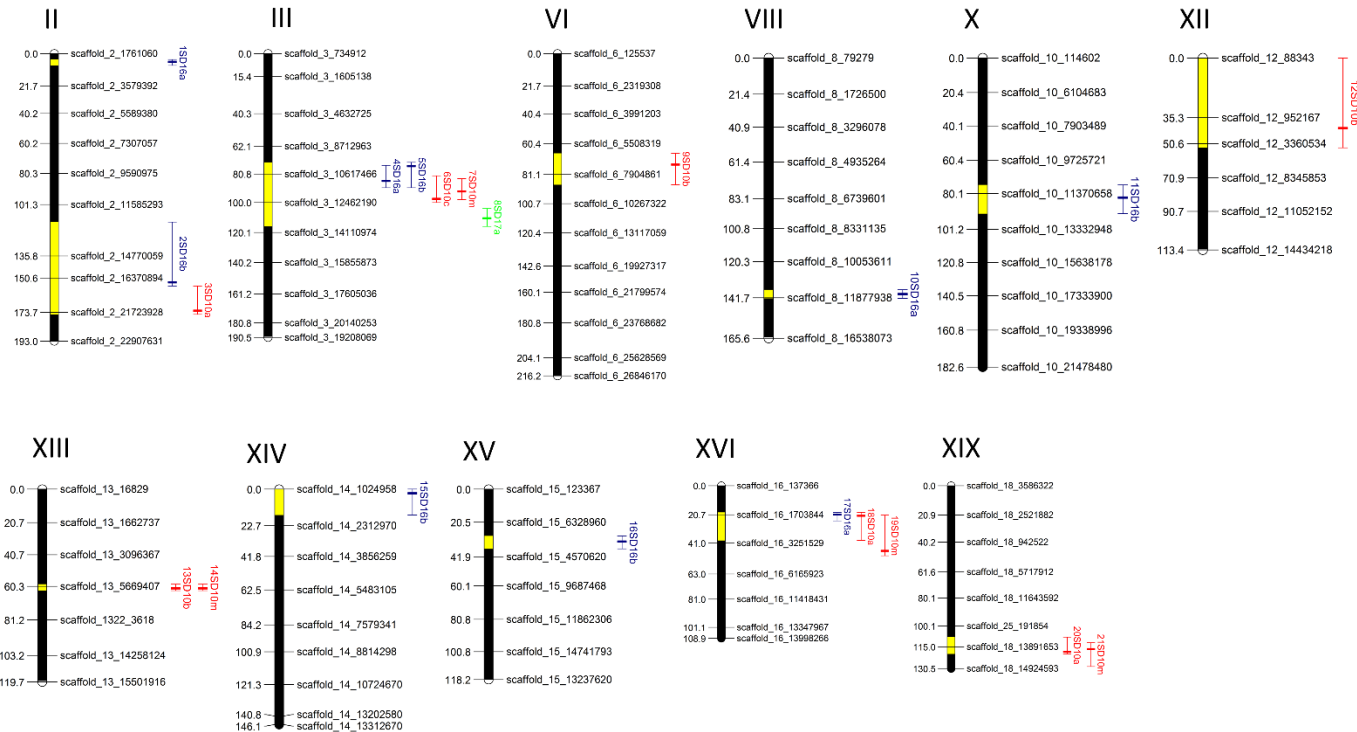
Figure 1 QTL anchored to the genome of *Populus trichocarpa* (V3 assembly, Kelleher et al. 2007). The actual map has a high marker density (average marker spacing: 5 markers per 4 cM). For illustration purposes, for each linkage group an evenly spaced selection of scaffolds is shown (one marker per 20 cM). The yellow regions on LGs represent QTL intervals using interval mapping (before co-factor selection). QTL for stomatal density (SD) are shown in red, green and blue for datasets collected in 2010, 2016 and 2017, respectively. The outer lines of bars are CW thresholds and the middle lines are LOD maxima (see Table 1). Scaffold intervals are represented in Mb.
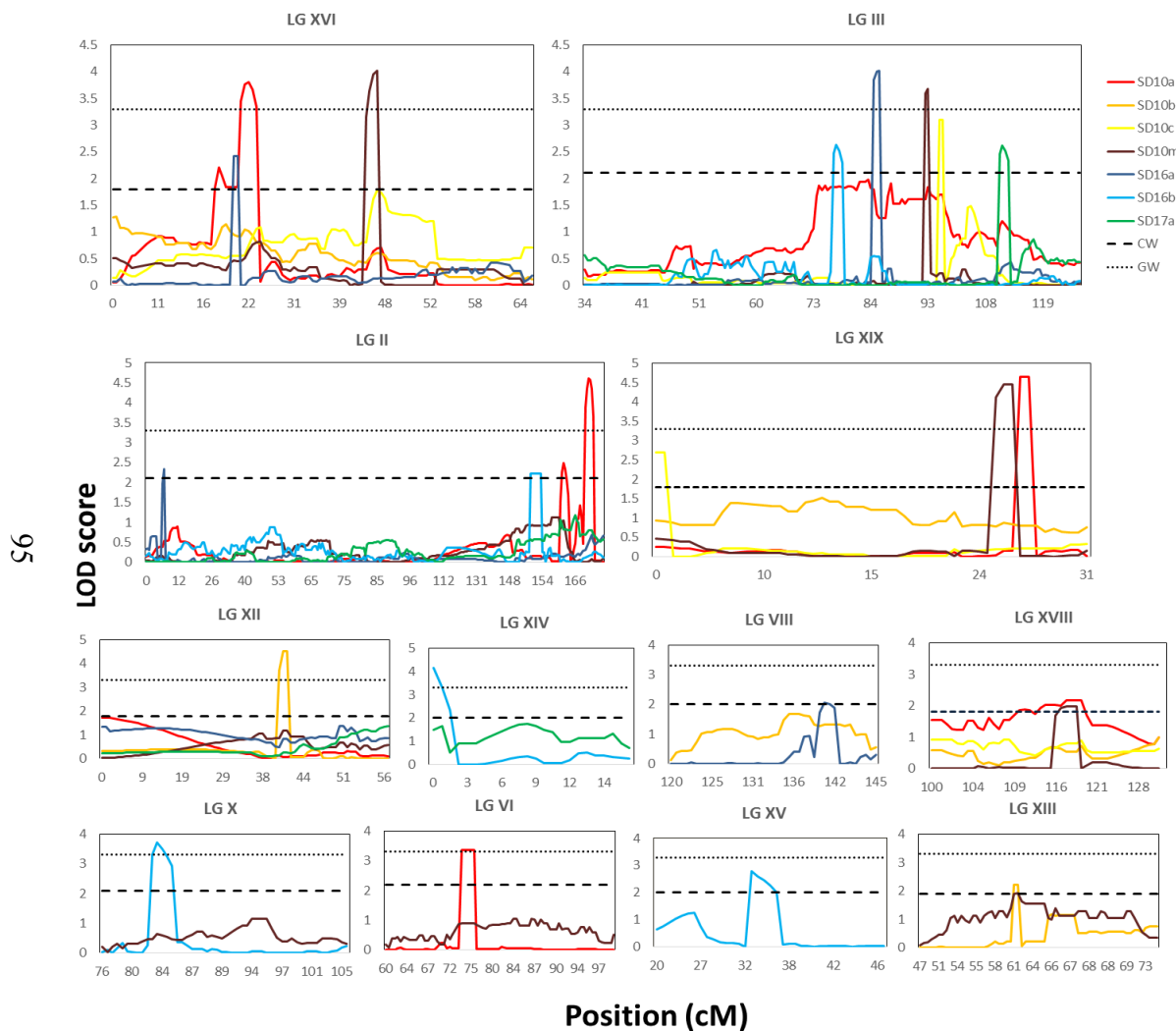
Figure 2 LOD score profiles for stomatal density QTL on LG II, III, VI, VIII, X, XII, XIII, XIV, XV, XVI and XIX using the Multiple-QTL Model (MQM) with co-factor selection across all experimental replicates. Chromosome-wide (CW) and genome-wide (GW) significance thresholds are shown with dashed lines (α =0.05, 1,000 permutations). Profiles for stomatal density are shown in shades of red, blue and green for data sets collected in 2010, 2016 and 2017, respectively. Some suggestive QTL with lower LOD score than the CW threshold are shown as well.

Table 1. QTL associated with stomatal density identified in the poplar pseudo-backcross pedigree 52-124 for years 2010, 2016 and 2017.

QTL #, population (lg)	Interval Mapping		MQM Mapping	LOD max				Origin of positive allele	
	Map location (V2, cM)	Physical location (V3)	Physical location (V3)	LOD value	Location (cM)	Physical location	% PVE	DD	DT
1SD16a (II)	6.321-7.943	2390513..2532397	2390513..2532397	2.34	7.146	2494005	6	**31**	27
2SD16b (II)	113.997-155.482	12992182..19970547	16904784..18010025	2.23	154.032	17984858	4.6	**25**	22
3SD10a (II)	156.015-175.197	20910935..23483297	20910935..23483297	4.35*	173.734	23433435	14.5	**38**	29
4SD16b (III)	74.564-89.295	11627271..12900652	11654461..11956032	2.63	75.849	11758002	5.5	22	**25**
5SD16a (III)	75.849-89.295	11758002..12900652	12307868..12635780	4.01*	85.512	12577722	10.5	26	**31**
6SD10c (III)	82.228-100.001	12307868..14031477	13706345..13918404	3.09	97.3	13882750	12.7	32	**39**
7SD10m (III)	84.343-97.647	12485960..13918404	13457418..13606110	3.68*	93.335	13508478	10.1	33	**37**
8SD17a (III)	104.001-116.113	14514569..15336641	14830986..15157363	2.62	110.959	14981850	6.3	22	**25**
9SD10b (VI)	67.482-88.672	6540985..8945894	7010101..7418569	3.37*	74.575	7286559	10.9	**37**	31
10SD16a (VIII)	137.888-141.713	11733076..12070763	11733076..12070763	2.06	139.505	11896024	5.3	27	**31**
11SD16b (X)	75.586-92.973	11655403..13247359	12121301..12589424	3.72*	82.619	12329381	7.9	21	**25**
12SD10b (XII)	0.877-53.703	407164..4446679	3037452..3341456	4.52*	41.339	3311474	15.1	**38**	31
13SD10b (XIII)	60.322-61.958	6076968..6442032	6076968..6442032	2.22	61.416	6283774	7	32	**37**
14SD10m (XIII)	60.322-61.958	6076968..6442032	6076968..6442032	1.91	61.416	6283774	9.1	32	**36**
15D16b (XIV)	0-16.565	0..1979129	0..68230	4.17*	0	0	8.9	21	**26**
16SD16b (XV)	30.834-36.579	2608259..3099304	2608259..3099304	2.58	32.888	2711938	5.4	**25**	22
17SD16a (XVI)	19.291-24.68	1768332..2141944	1768332..1955790	2.41	20.738	1916966	6.2	27	**31**
18SD10a (XVI)	19.291-39.166	1588548..3292119	1588548..2141944	3.76*	21.477	1975247	12.3	30	**37**

19SD10m (XVI)	21.477-49.048	1975247..4163931	3582922..3913475	4.01*	46.355	3913475	11	32	**38**
20SD10a (XVIII)	108.468-119.783	15087827..16199806	15910227..16065443	2.17	119.283	16199806	10.5	**35**	29
21SD10m (XVIII)	112.31-119.611	15301185..16216454	15910227..16065443	1.97	117.395	16037719	5.2	**36**	33
22SD10c (XIX)	0-33.215	0..8436001	0..389959	2.7	0.128	350563	11	**39**	32
23SD10a (XIX)	21.904-40.875	4152016..10085840	4554508..4983998	4.65*	28.132	4870964	15.6	**37**	30
24SD10m (XIX)	0-40.875	0..10085840	4152016..4870964	4.45*	26.312	4541228	12.4	**38**	32

lg: linkage group; V2: markers anchored on version 2 of the *P. trichocarpa* genome; V3: version 3 updated physical location; PVE: percent phenotypic variance explained; DD: homozygous for the *P. deltoides* allele, DT: heterozygous for the *P. deltoides* and *P. trichocarpa* alleles. LOD max determined using MQM mapping, value with *: above GW threshold, otherwise above CW threshold (Example: 1SD16a: QTL number one, stomata density, year 2016, replicate a). The number of genotypes is 90 for year 2010 and 150 for both 2016 and 2017.

Table 2 Number of candidate genes detected across the twenty QTL significantly associated with stomatal density.

QTL #, dataset (lg)	Interval (V3, bp)	Number of genes
1SD16a (II)	2390513..2532397	22
2SD16b (II)	16904784..18010025	45
3SD10a (II)	20910935..23483297	114
4SD16b (III)	11654461..11956032	34
5SD16a (III)	12307868..12635780	44
6SD10c (III)	13706345..13918404	27
7SD10m (III)	13457418..13606110	14
8SD17a (III)	14830986..15157363	42
9SD10b (VI)	7010101..7418569	52
10SD16a (VIII)	11733076..12070763	50
11SD16b (X)	12121301..12589424	66
12SD10b (XII)	3037452..3341456	35
13SD10b (XIII)	6076968..6442032	28
14SD10m (XIII)	6076968..6442032	28
15D16b (XIV)	0..68230	6
16SD16b (XV)	2608259..3099304	34
17SD16a (XVI)	1768332..1955790	24
18SD10a (XVI)	1588548..2141944	82
19SD10m (XVI)	3582922..3913475	34
20SD10a (XVIII)	15910227..16065443	15
21SD10m (XVIII)	15910227..16065443	15
22SD10c (XIX)	0..389959	38
23SD10a (XIX)	4554508..4983998	30
24SD10m (XIX)	4152016..4870964	54

Number genes found within the nineteen QTL based on MQM mapping with cofactor selection. lg: linkage group.
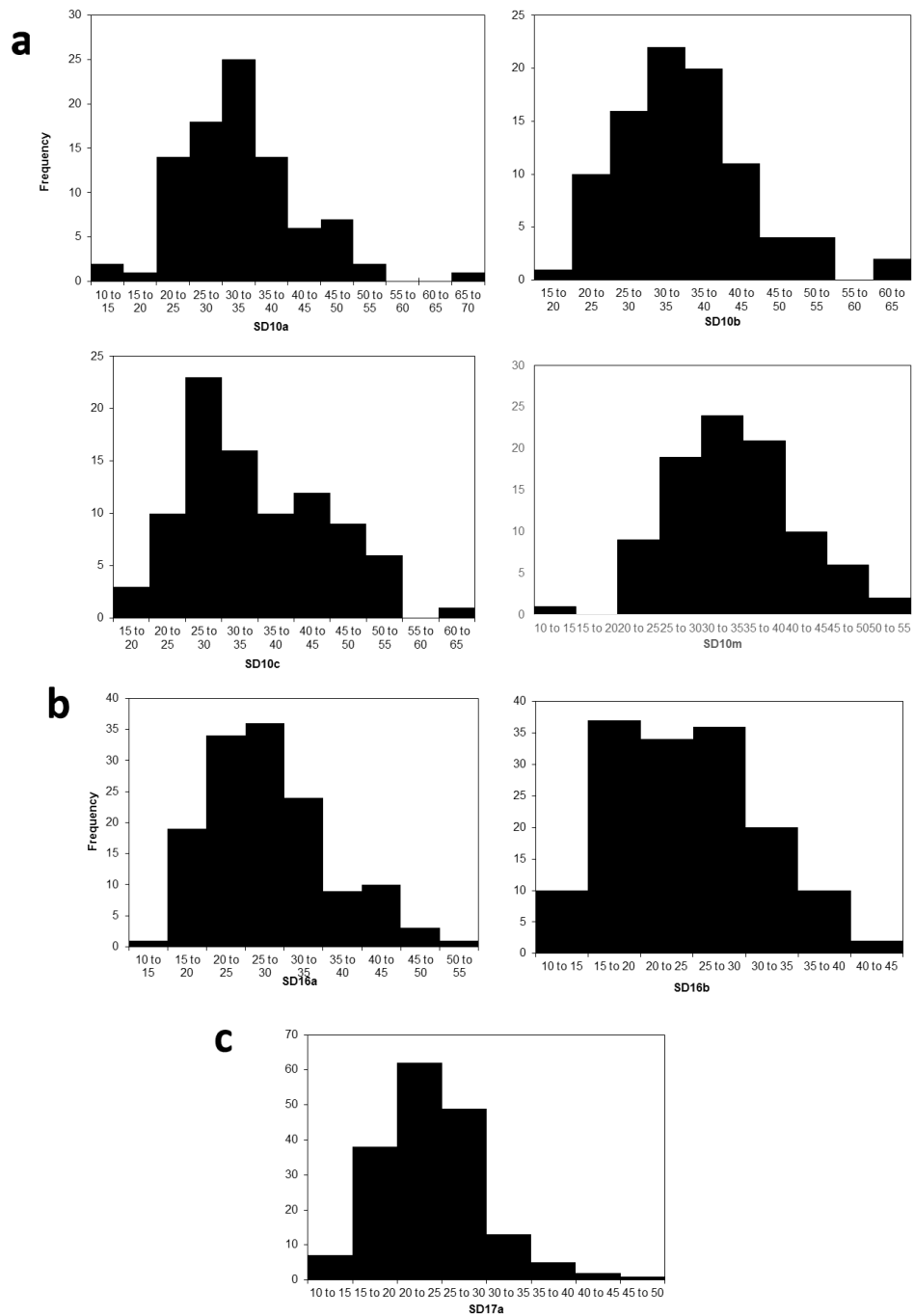
Supplementary Figure 1 Frequency distribution for stomatal density collected in year 2010, 2016 and 2017 (a, b and c, respectively). Three leaves per genotype were tested for stomatal density in 2010, while leaves taken from two individual plants per genotype were sampled in 2016.
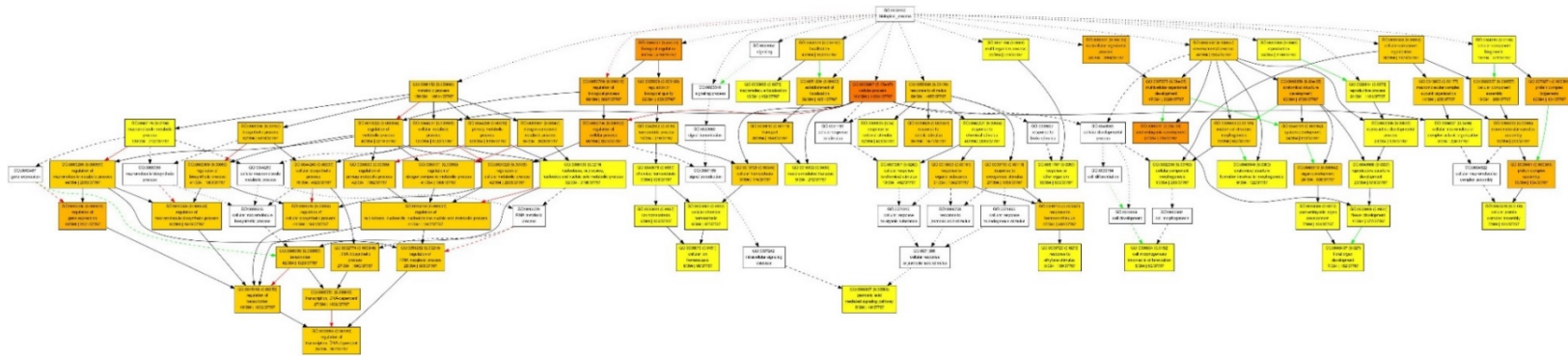
Supplementary Figure 2 Hierarchical tree graph of overrepresented GO terms in biological process categories for all underlying genes within the twenty trait-associated QTL. The non-significant terms are shown as white boxes, and significant terms (P ≤ 0.05) are marked with color (the degree of color saturation of boxes is associated to the enrichment level of the term). Red indicates the highest enrichment level.
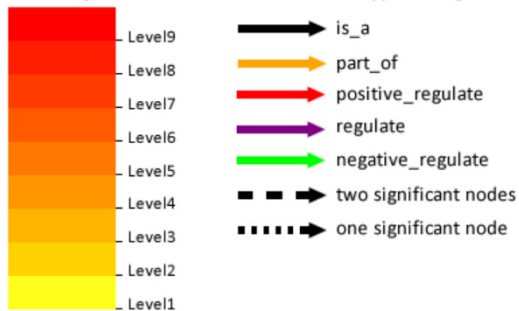
Supplementary Figure 3 The significant levels of the overrepresented GO terms colored from shades of yellow, orange and red as well as the relationships between GO terms indicated by different arrows are shown. Red indicates the highest enrichment level.

Supplementary Table 1 Pair-wise estimates of phenotypic correlations calculated as Pearson correlation coefficients for stomatal density collected for years 2010, 2016 and 2017.

	SD10a	SD10b	SD10c	SD10m	SD16a	SD16b	SD17a
SD10a	X						
SD10b	0.985032	X					
SD10c	0.988059	0.981475	X				
SD10m	0.976836	0.98128	0.985206	X			
SD16a	0.92854	0.926599	0.944615	0.960826	X		
SD16b	0.949868	0.9464	0.967087	0.966766	0.980137	X	
SD17a	0.91097	0.898656	0.924961	0.930541	0.939645	0.961409	X

The analysis for stomatal density (SD) was done between replicates and across years. All correlations had a p-value < 0.0001. For 2010, replicates a, b and c represent leaves taken from the same plant and replicate m is their mean; whereas for 2016, replicates a and b are clonal replicates taken from two individual plants per genotype.

Supplementary Table 2[7] All candidate genes within the twenty QTL detected for stomatal density variation in *Populus*. Physical localization and annotation of gene models within each QTL interval are listed.

---

[7] Supplementary Table 2 are in excel format submitted as digital files.

Chapter 5: Summary and Conclusion

The present thesis investigates the genus *Populus* using population genomics approaches to better understand the genetic system controlling the inheritance of bark features, diameter growth and stomatal density and to associate these traits with respective genes.

For Quantitative Trait Locus (QTL) analysis, the experimental material consisted of an interspecific hybrid poplar pseudo-backcross pedigree composed of 396 genotypes. The mapping population was planted at Oregon (OR) and West Virginia (WV) in a three-block replication for a total of six ramets per cloned genotype. Overall, bark texture, bark thickness and diameter growth showed high correlations among replicates and years, however, across sites, correlations were only significant for bark texture.

A total of 94 individual QTL detected for the three traits, bark texture, bark thickness and stem diameter, across various chromosomes were successfully anchored to the *Populus* genome assembly. Specifically, five QTL clusters for diameter and seven QTL clusters for bark thickness were observed. The most significant and reproducible results were found associated with bark texture for which seven QTL clusters were detected. Given the environmental contrast between the OR and WV experimental sites, four out of the seven QTL clusters, representing a total of 47 individual QTL detected for bark texture, were remarkably consistent across both sites. Differences in reproducibility for QTL clusters across sites suggest differential environmental effects on gene expression. In comparison to bark texture, QTL clusters for bark thickness and diameter had lower reproducibility across sites.

QTL clusters on chromosome I, VI and XII were associated with all three traits. Co-location of QTL for traits can be the result of pleiotropic effects or closely linked genes. These overlapping QTL could be an explanation of different aspects of bark texture, bark thickness and radial growth. Romero [1] proposed that rough bark results in response to the mechanical stresses imposed by a varied radial growth and due to different meristematic activity in the phellogen, a discontinuous periderm. Strong correlations between bark texture and diameter could indicate that bark texture is partly related to diameter growth. Furthermore, using MQM mapping, QTL for these traits were mapped to different neighboring positions of the same chromosomes (Fig. 2).

Consequently, bark texture seems to be only partly related to diameter growth, and other factors such as meristematic activity of the phellogen and cell adhesion are likely to have major effects on bark texture. A higher mapping resolution as obtained in linkage disequilibrium mapping in natural population samples is needed to narrow down QTL regions to individual genes and to distinguish between pleiotropic effects and close linkage.

This limitation was evident in our QTL analyses while looking at the candidate genes list, where most clusters included from 123 to 963 candidate genes. Nonetheless, several candidate genes within the QTL interval can be identified based on their putative functions including PopNAC128 (Potri.001G206900), one of the orthologs of *Arabidopsis* ANAC104 (*Arabidopsis* Nac Domain Containing Protein 104) and XND1 (Xylem NAC Domain 1), and Potri.001G206700, an ortholog of AT4G33430 (BAK1, Bri1-Associated Receptor Kinase; ELONGATED; SERK3). The first gene was shown to be involved in the development of phloem fibers and its overexpression resulted in a slowed secondary phloem development [2], while the second gene is found to be involved in patterning and growth regulation [3, 4].

Variation in bark texture could be related to cell adhesion which is essential to form a single periderm resulting in smooth bark. The lack of cell-cell adhesion leads to the development of uneven and discontinued bark or bark splitting causes a peeling and fissured bark appearance. At the molecular level, several expressional candidate genes identified in this study have a role in cell adhesion, including Pinin (Potri.001G208200) and PopFLA or Fasciclin-Like Arabinogalactan (Potri.013G151300, Potri.013G151400 and Potri.013G151500)[5-7]. They all fell within QTL intervals with the highest LOD scores and had high expression in phloem and xylem tissues and the cambium.

Bark features in our study ranged from smooth to deeply furrowed which is characteristic for *P. deltoides*. Variation in shallowly fissured bark which is characteristic for *P. trichocarpa* was not observed in the segregating progeny. Thus, the QTL identified in this progeny set only represent a subset of a larger number of polymorphisms affecting the traits. And in our pseudo-backcross pedigree (DD X DT) involving multiple *P. deltoides* alleles, polymorphisms associated with characteristic bark features of *P. trichocarpa* seem to be largely undetected. Association populations for *P. trichocarpa*

will be used to find additional candidate genes associated with bark texture in this species.

Therefore, a Genome-Wide Association Study (GWAS) analysis was carried out using an association population consisting of 1,100 black cottonwood genotypes (*Populus trichocarpa* Torr. & Gray). They were planted at Oregon (OR) and California (CA) where genotypes were planted in a three-block replication for a total of six ramets per genotype. Our results revealed multiple significant associations. Through GWAS, we were able to detect several putative genes involved in the control of bark texture, as well as differentiate between the genes responsible for the flaky bark texture in *P. trichocarpa* and the rough furrowed texture in *P. deltoides* in comparison to the QTL study.

Overall, the detected genomic regions involved in the control of bark texture in *P. trichocarpa* and *P. deltoides* vary, only eight genes out of 380 GWAS-identified genes were found to be co-locating with the QTL study.

A candidate gene list based on the highest p-values and reproducibility with possible functional control of bark texture was identified. Their putative function fell within three different categories: 1) radial growth and tissue differentiation, 2) suberin accumulation and 3) programmed cell death. Vascular differentiation and differentiation of the outer bark involves several steps including cell division, orientated cell differentiation, cell expansion, cell wall thickening, and programmed cell death. Discontinuous periderms (textured bark) could result from variable radial meristematic activity in the cork cambium due to mechanical stresses from radial growth. Phellem is multilayered dead tissue that is made impervious by the disposition of suberin onto their cell walls while suberization occurs during the development of the secondary radial meristems (the cork cambium). Finally, programmed cell death is an essential component during the plant's secondary development and is involved in the generation of the vascular system contributing to the development of bark and its abscission.

All three processes are essential to secondary development and can influence the development of bark. Thus, a better understanding of the putative genes involved in these developments can provide insights into bark tissue development. Several genes of unknown function detected within highly significant regions associated with bark texture open new opportunities for future studies and should be functionally characterized.

The combination of QTL and GWAS mapping supports the identification of candidate genes. Eight genes out of 380 GWAS-identified genes were found to be co-locating with the QTL from the previous study, five of which have an unknown function. Two genes with annotated function as root hair defective 3, Potri.012G116900 and Potri.012G117000, co-located with major highly reproductive QTL for bark texture. The same QTL region was also associated with bark thickness and diameter, suggesting a role of this QTL in radial growth. These genes putatively function as transmembrane proteins identified from the plasma membrane of *Populus* differentiating xylem and phloem [8]. QTL mapping increases the ability to reduce false positive rates whereas GWAS has a much higher resolution than QTL mapping and can narrow down the genomic region to individual candidate genes. In our case, the combination of QTL and GWAS mapping provided insights into the different genes responsible for variations in bark texture ranging from smooth and furrowed across *Populus* species.

Looking at the stomatal density results using QTL mapping, four QTL clusters were detected along with several individual QTL. It has been generally reported that stomatal density is controlled by both environmental and genetic factors, suggesting the involvement of many genes. Any variation in stomatal density ultimately influences photosynthesis and atmospheric CO_2 concentration. Our results however, based on strong inter annual correlations among genotypes and the correlations among replicates within years, suggest that genetics plays a strong role in determining stomatal density in *Populus*.

Systemic signals from mature leaves regulate stomatal development of expanding leaves, however, the nature of these signals generated in the mature leaves and transmitted to developing leaves is largely unknown and it is unclear how many signals are elicited. Our study further suggests high genetic variation for the control of stomatal density in a single interspecific hybrid progeny, but also variation among years and replicates. We detected twelve genomic regions on eleven linkage groups with a range of QTL effects explaining from 4.6% to 15.6 % of the phenotypic variance. Four major QTL clusters with high reproducibility and consistency associated with stomatal density were found on linkage groups (LGs) II, III, XVI and XIX. Clusters on LG II, III were consistently identified across all three years, while the clusters on LG XVI were found

reproducible across two years. Since the identified genomic intervals are large and contain hundreds of candidate genes, it is still difficult to identify the genes in question. Thus, we seek to analyze the overrepresented GO terms in the QTL intervals. GO analysis showed the functional profile of the underlying genes and their involvement in biological processes such as developmental processes, including cell morphogenesis and differentiation, post-embryonic development, and cellular processes, such as metabolic and cellular biosynthesis regulations. The natural variation of stomatal density is linked to plant fitness and adaptation to their environment, thus directly impacting speciation and evolutionary change [9].

QTL and GWAS mapping provide a baseline, a list of candidate genes associated with the trait of interest. Future work through transgenic lines and assessment of nucleotide variation in these candidate genes controlling such complex adaptive traits under different environmental conditions will allow a better understanding of the genetic and environmental control of these traits.

References

1. Romero C: **Tree responses to stem damage.**: University of Florida; 2006.
2. Grant EH, Fujino T, Beers EP, Brunner AM: **Characterization of NAC domain transcription factors implicated in control of vascular cell differentiation in *Arabidopsis* and *Populus**. Planta* 2010, **232**(2):337-352.
3. Schwessinger B, Roux M, Kadota Y, Ntoukakis V, Sklenar J, Jones A, Zipfel C: **Phosphorylation-dependent differential regulation of plant growth, cell death, and innate immunity by the regulatory receptor-like kinase BAK1.** *PLoS Genetics* 2011, **7**(4):e1002046.
4. Halliday K, Devlin PF, Whitelam GC, Hanhart C, Koornneef M: **The ELONGATED gene of *Arabidopsis* acts independently of light and gibberellins in the control of elongation growth.** *The Plant Journal* 1996, **9**(3):305-312.
5. MacMillan CP, Mansfield SD, Stachurski ZH, Evans R, Southerton SG: **Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in *Arabidopsis* and *Eucalyptus.** The Plant Journal* 2010, **62**(4):689-703.
6. Wang H, Jiang C, Wang C, Yang Y, Yang L, Gao X, Zhang H: **Antisense expression of the fasciclin-like arabinogalactan protein FLA6 gene in *Populus* inhibits expression of its homologous genes and alters stem biomechanics and cell wall composition in transgenic trees.** *Journal of Experimental Botany* 2014, **66**(5):1291-1302.
7. Shi Y, Tabesh M, Sugrue SP: **Role of cell Adhesion–Associated protein, pinin (DRS/memA), in corneal epithelial migration.** *Investigative ophthalmology & visual science* 2000, **41**(6):1337-1345.
8. Song D, Xi W, Shen J, Bi T, Li L: **Characterization of the plasma membrane proteins and receptor-like kinases associated with secondary vascular differentiation in poplar.** *Plant Molecular Biology* 2011, **76**(1-2):97-115.
9. Hetherington AM, Woodward FI: **The role of stomata in sensing and driving environmental change.** *Nature* 2003, **424**(6951):901.