# Genome-wide association studies of bark texture in *Populustrichocarpa*

Roba Bdeir[1] • Wellington Muchero[2] • Yordan Yordanov[3] • Gerald A. Tuskan[2] • Victor Busov[1] • Oliver Gailing[1,4]

* Oliver Gailing
ogailing@mtu.edu; ogailin@gwdg.de

[1] School of Forest Resources and Environmental Science, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931, USA
[2] Biosciences Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA
[3] Departement of Biology, Eastern Illinois University, 600 Lincoln Ave, Charleston, IL 61920, USA
[4] Present address: Forest Genetics and Forest Tree Breeding, Faculty of Forest Sciences and Forest Ecology, University of Göttingen, Büsgenweg 2, 37077 Göttingen, Germany

### Abstract

Plants have developed bark as a defense barrier to deal with environmental stresses, such as pathogen invasion, drought, and UV radiation. Bark is composed of dead differentiated cells and is formed when radial expansion pushes the cortex and epidermis outward creating secondary meristem. 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 (BT) in *Populus trichocarpa*. Clonally replicated provenance trials of *P. trichocarpa* were studied for BT collected across three sites, multiple years, and (2–3) biological replicates per site. Forty-one genomic intervals containing SNPs significantly associated with BT 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 phenotype. Association mapping using low coverage sequencing allowed us to detect narrow genomic intervals (1–20 kb) with high reproducibility and shared candidate genes. For example, a membrane-associated apo- ptosis protein and a wall-associated receptor kinase (PR5K-like) protein, which both are involved in radial growth and tissue differentiation, lie within significant trait-associated region. Two copies of root hair defective3 genes, transmembrane protein in *Populus* differentiating xylem and phloem, were also significantly associated with BTand co-located with major quantitative trait locus (QTL) for BT, bark thickness, and diameter from our previous study suggesting an important role in radial growth.

### Keywords

Bark texture; Genome-wide association studies (GWAS); Association mapping; *Populus trichocarpa*

### Funding information

Support for the phenotyping and graduate support for Roba Bdeir was provided by the USDA-NIFA grant (# 2012-67013- 19389) to VB, YY, and OG. Additional partial support for Roba Bdeir was provided by the National Science Foundation grant to OG (# 1230803) as part of the BSustainable Forest-Based Biofuel Pathways to Hydrocarbon Transportation Fuels^ project at Michigan Tech. The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This work was supported by the U.S. Department of Energy under Contract to Oak Ridge National Laboratory. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the US Department of Energy under contract number DE-AC05-00OR22725. We also acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Funds of Göttingen University.

## Introduction

Bark in trees is formed when radial growth pushes the cortex and epidermis outward creating secondary meristem, i.e., the cork cambium (Carlquist 1996). The cork cambium then out- wardly produces cork (phellem) cells; the bark's outer layer which is composed of dead differentiated cells (Carlquist 1996). As an adaptation to terrestrial life, the 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 (Pollard et al. 2008; Soler et al. 2007). Despite of its importance, tree bark development has been less extensively studied than wood development.

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 (Mellerowicz et al. 2001). Carpita et al. (2001) 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 coordinated regulation of diverse metabolic path- ways (Brown et al. 2005); however, the molecular mecha- nisms 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 (Turner and Somerville 1997) or defects in lignin biosynthesis (Piquemal et al. 1998), 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, and to plated (Whitmore 1962). Differences in bark texture are of eco- logical importance and have been related to different levels of susceptibility to fungal and insect diseases and to fire (Beaver et al. 1989; Gill and Ashton 1968; Lawes et al. 2011; Marshall et al. 2013). For example, tree species with smooth bark, e.g., beech and ash, have been reported to be less susceptible to insect and fungal diseases (Beaver et al. 1989; Marshall et al. 2013). 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 (Climent et al. 2004; Lawes et al. 2011; Pausas 2015). Previously, we have analyzed the bark texture of an interspecific pseudo-backcross pedigree derived from a cross of a female hybrid clone (*Populus trichocarpa* × *P. deltoides*) with a male *Populus deltoides* parent through quantitative trait locus (QTL) mapping (Bdeir et al. 2017). 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 (Grant et al. 2010), along with Pinin and Fasciclin (PopFLA) that play a role in cell adhesion, cell shape, and migration (Lafarguette et al. 2004; Wang et al. 2014).

The parental species used to derive the pedigree display contrasting bark textures. *P. trichocarpa* (Torr. and 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, (3) and finally, 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 1100 black cottonwood genotypes (*P. trichocarpa* Torr. and A.Gray ex Hook.) 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 (Slavov et al. 2012) 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 (Muchero et al. 2015).

### *Genetic Data*

A total of 917 genotypes from the population were resequenced to a minimum 15× depth using the Illumina Genome Analyzer, HiSeq 2000 and HiSeq 2500 as previously described for the same population (Evans et al. 2014; Zhang et al. 2018). 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 50,000) 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), BT 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. BT was assigned a qualitative score based on a scale from 1 (smooth) to 3 (flaky with grooves) (Online Resource 1). Two replicates of 1100 genotypes in three blocks, for a total of six ramets per genotype, were an- alyzed at CLA13 and PLC14 sites, along with three replicates for CLA12 and COR13 and one replicate for CLA15. In ad- dition to the field assessment, the CLA collection during 2012 had two replicate wood disks with bark collected at breast height per genotype, thus a total of five 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 (Fitch 2006) to assess correlations within blocks and within and between years (Online Recourse 2).

### *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 (Online Resource 3). Association analyses were performed as previously described by Zhang et al. (2018), using the efficient mixed model association (EMMA) algorithm in the EMMAX software with kinship as the correction factor for genetic background effects (Kang et al. 2010) to compute genotype to phenotype associations using 8.2 million SNP and indel variants with BT as the phenotypic variable. On average, a marker is represented every 17 base pairs (bp) across the genome. Significant and reproducible SNPs strong- ly associated with BT 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 (Storey and Tibshirani 2003).

### *Candidate Genes*

Based on a linkage disequilibrium (LD) decay below 0.2 with- in 3–6 kb in this *P. trichocarpa* population (Slavov et al. 2012), a 6-kb upstream and downstream region was examined for candidate genes. The significant genomic regions have to have at least one detectable SNP with a significance of $-\log10\ p \geq 6$ and be circumscribed by reproducible SNPs with a significance of $-\log10\ p \geq 5$ (Online Resource 5). The underlying genes were identified based on the *P. trichocarpa* reference genome v3.0 in the Phytozome database using the Bio M art t ool (https://phytozome.jgi.doe.gov) (Online Resource 5). 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, e.g., 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 BT ranged from smooth (score = 1) to flaky bark (score = 3) (Online Resource 1). The mean values and standard errors for Clatskanie, Corvallis, and Placerville sites were as follows: CLA12 (mean, 1.53; SE, 0.70); CLA13 (mean, 1.50; SE, 0.69); CLA15 (mean, 1.89; SE, 0.81); COR13 (mean, 1.71; SE, 0.62) PLC14 (mean, 1.46; SE, 0.63). Phenotypic correlations were generally higher within the same environment with varying levels of significance across different environments (Online Resource 2).

### *Genetic Association with Phenotypic Variation in Bark Texture*

The GWAS using 8.2 million SNPs uncovered a total of 755 individual SNPs significantly ($-\log10\ p \geq 5$) and reproducibly associated with bark texture distributed across all 19 chromosomes. The 175 genomic regions defined by these SNPs include 380 unique genes (Online Resource 4). 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 Fig. 1.

At a significance threshold of $-\log10\ p \geq 6$, the 175 genomic regions were further reduced to 50 genomic regions ranging from 1 to 20 kb intervals, featuring 324 out of the 755 individual SNPs (Online Resource 5). 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 detect- able SNPs ($-\log10\ p \geq 5$).

### *Characterization of Genomic Regions*

Table 1 summarizes the 41 genomic regions including SNPs with the highest significance value (detailed information in Online Resource 5). 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 12 regions were reproducible across clonal replicates within a site. Comparing the significantly associated regions, they ranged from 1 to 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' positions and features are summarized in Table Online Resource 5.

Seventeen regions had $-\log10\ p \geq 7$, with the highest sig- nificance on LG XV (8.44). Considering replicates, four re- gions, 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 $-\log10\ p \geq 6$ (region 5, 8, 11, and 31); whereas, five regions contained 4–5 individual SNPs above $-\log10\ p \geq 6$ (region 6, 9, 17, 33, and 34, Online Resource 5).

A detailed list of the underlying putative genes within the 41 regions and their interval is shown in Online Resource 5. 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 BT (Table 2). Eight of these genes occurred in QTL detected for BT in our previous study (Bdeir et al. 2017) (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, o n 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 BT phenotypes to associated polymorphisms in the genome and identify putative genes related to bark development.

Even though the genes involved in BT 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 BT (Romero 2006). 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 (Andersson-Gunnerås et al. 2006; Brown et al. 2005; Li et al. 2017; Mellerowicz et al. 2001; Song et al. 2011; Sterky et al. 1998). 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 intra- cellular trafficking proteins.

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 (Bdeir et al. 2017). 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 $-\log 10$ 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 (Bdeir et al. 2017). These genes putatively function as transmembrane proteins identified from the plasma membrane of Populus differentiating xylem and phloem (Song et al. 2011).

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, the majority 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 p values and reproducibility with possible functional control of BT 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 (Chaffey et al. 2002). 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 (Plomion et al. 2001). Vascular differentiation and differentiation of the outer bark involves cell division, orientated cell differentiation, cell expansion, cell wall thickening, and programmed cell death (Plomion et al. 2001; Song et al. 2011). Discontinuous periderms (textured bark) could result from variable radial meristematic activity in the cork cambium due to mechanical stresses from radial growth (Romero 2006).

The gene Potri.015G124000, a membrane-associated apoptosis protein (NAP1), with a SNP located in its intron reached the highest significance score ($-\log 10$ $p$ = 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 (Brembu et al. 2004; Deeks et al. 2004; Zhang et al. 2005). NAP1 is thus involved in plant cell morphogenesis and plays an essential role in cell expansion.

Another gene family, cytochromes P450, have been detected within six genomic regions (four having ($-\log10\ p \geq 7$) across all years for Clatskanie and Corvallis sites (Table 2). Cytochromes P450 are enzymes involved in the biosynthesis of brassinosteroids which are perceived by the plasma membrane-localized leucine-rich repeat-receptor kinase BRI1. Brassinosteroids play essential roles in developmental processes including cell expansion and vascular differentiation (Clouse and Sasse 1998). For instance, loss-of-function of *brl1* causes abnormal phloem:xylem differentiation ratios and vascular defects (Caño-Delgado et al. 2004). 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* (Meyer et al. 1998; Szekeres et al. 1996). While an overexpression of F5H, a cytochrome P450-dependent monooxygenase, in *Arabidopsis* led to the accumulation of lignin indicating a key regulatory role in syringyl lignin biosynthesis (Meyer et al. 1998).

Several additional GWAS candidate genes, including pectate lyase, have been reported to regulate cell wall loosening and expansion (Xu et al. 2014). 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 (Cosgrove 2000).

### *Suberin Accumulation*

Suberization occurs during the development of the secondary radial meristems, i.e., the cork cambium (Soler et al. 2007). 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 (Pollard et al. 2008; Soler et al. 2007).

The genes Potri.017G034700 and Potri.017G035400, wall-associated receptor kinases (PR5K-like), are found with- in highly significant genomic regions ($-\log10\ p = 7.90$, syn- onymous 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 (Bent 1996; Wang et al. 1996). 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 (Legay et al. 2015), 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 members have been reported as key regulators in the biosynthesis of lignin and the regulation of secondary cell wall formation (Caño-Delgado et al. 2010; Goicoechea et al. 2005; Hertzberg et al. 2001; Zhong et al. 2008). In Arabidopsis, MYB46 and MYB83 are direct targets of a group of NAC domain master regulators controlling secondary wall biosynthesis (Mitsuda et al. 2007; Zhong et al. 2008; Zhou et al. 2009). 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 in- volved in the biosynthesis and regulation of Populus suberin affecting the bark periderm, specifically the cork (phellem) (Legay et al. 2015; McCarthy et al. 2010; Rains et al. 2017; Zhao et al. 2014).

*Program Cell Death*

Programmed cell death is important in secondary plant development and is involved in the generation of the vascular system (Groover and Jones 1999). Programmed cell death also contributes to the development of bark and its abscission (Addicott 1982). 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 ex- posed tissue (Addicott 1982). Lignification occurs in xylem cells during wood development and in bark (Jelonek et al. 2017; Mellerowicz et al. 2001; Sengupta and Palit 2004). When lignification is completed, vessel elements undergo programmed cell death, which involves the hydrolysis of the protoplast (Asif et al. 2017; Mellerowicz et al. 2001; Venkatesan et al. 2016). Nonetheless, the specific effects of cell death on bark have 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 (Bdeir et al. 2017). 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 (Janjusevic et al. 2006). Studies have suggested their involvement in regulation of cellular elongation in Arabidopsis and in the biosynthesis of tension wood in *Populus* (Andersson-Gunnerås et al. 2006; Franciosini et al. 2013). 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 (Bachmair et al. 1990).

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 (Kohler et al. 2008; Van Ooijen et al. 2008). Four genes within several genomic regions ($-\log10\ p \leq 7.13$) 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. (2012) examined gene expression of secondary phloem of Norway spruce after it had been wounded and inoculated with the fungus *Ceratocystis polonica* (Fossdal et al. 2012). NB resistance genes showed a significant increase in expression along with five miRNAs with putative NB tar- gets (Fossdal et al. 2012).

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 known as cysteine-rich receptor-like protein kinases (CRKs). Induced expression of these CRKs triggered hypersensitive response-like cell death in transgenic plants (Chen et al. 2004).

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 an d Potri.017G035600 ($-\log10\ p = 6.91$ and 7.73, respectively), are top candidates for functional characterization along with 12 other genes with high significance ($-\log10\ p \geq 7$) (Online Resource 5).

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 (Brachi et al. 2010; Manenti et al. 2009; Zhao et al. 2007). In our case, the combination of QTL and GWAS mapping provided insights on the different genes responsible for the variations in BT ranging from smooth and furrowed across *Populus* species.

## Data Archiving Statement

Whole-genome resequencing, SNP/indel calling, and SnpEff analysis for the 544 individuals of this Populus GWAS population was previously described by Evans et al. (2014). In this study, we used the same sequencing and analytical pipelines to incorporate an additional 373 accessions (altogether 917 accessions). 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/. The resulting SNP and indel dataset is available at http://bioenergycenter.org/besc/gwas/. To assess genetic control, we used the EMMA algorithm in the EMMAX software with kinship as the correction factor for genetic background effects (Zhou and Stephens 2012) to compute genotype to phenotype associations using 8.253,066 SNP variants with minor allele frequencies $> 0.05$ identified from whole-genome resequencing. Linkage disequilibrium (LD) was determined using HAPLOVIEW v.4.2 (Barrett et al. 2005). To account for multiple testing, we used the Bonferroni correction and considered as significant only those SNPs for which $P < 0.05/8253066 = 6.06910{-}9$. This correction is known to be conservative and thus Bovercorrected^ the raw P values (McIntyre et al. 2000).

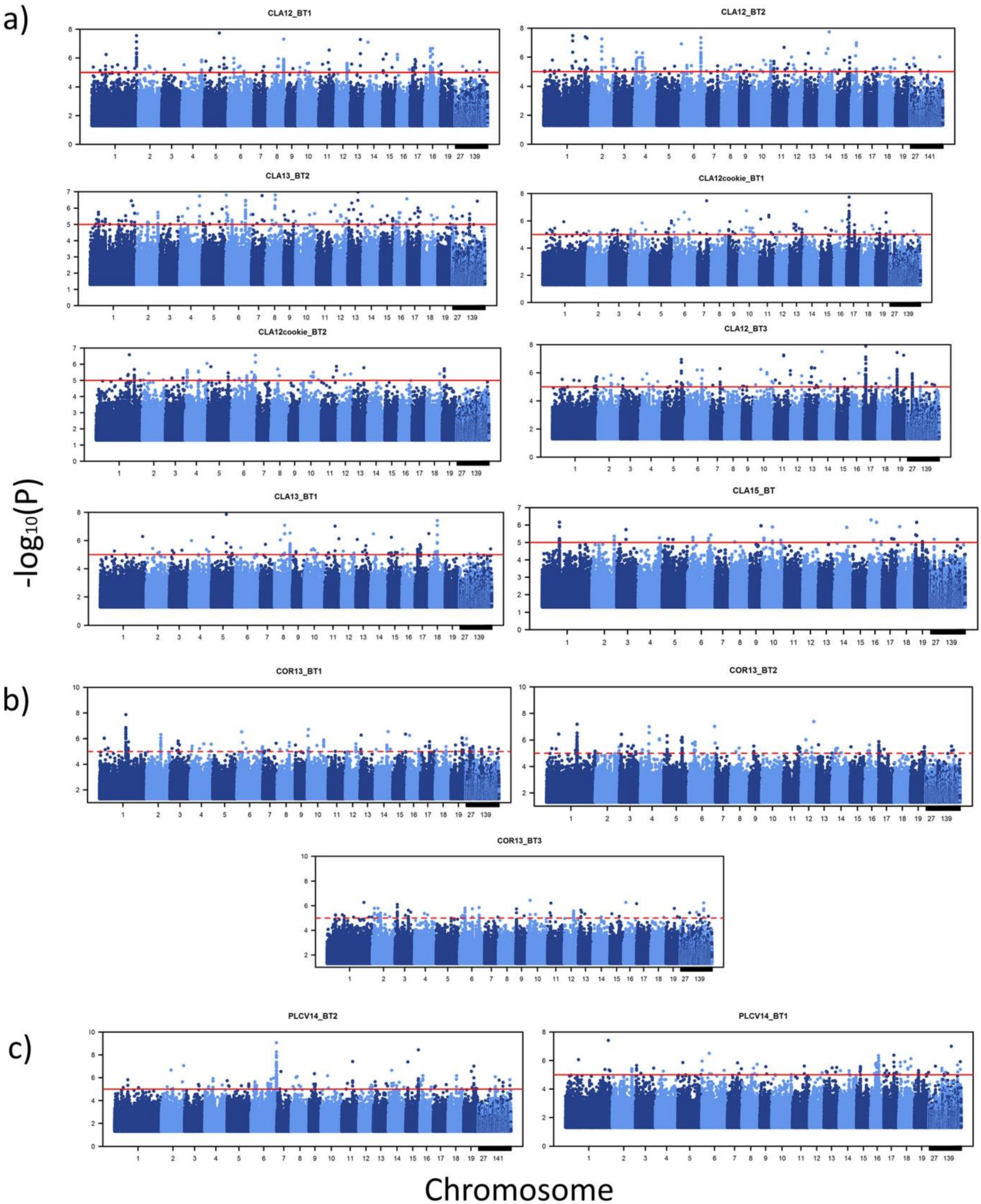


Fig. 1 Manhattan plots for GWAS for the *P. trichocarpa* BESC population. Standard − log10 P plots of the study results. The five datasets across three sites are as follows: Clatskanie (CLA) for year 2012, 2013, and 2015 (a); Corvallis (COR) for year 2013 (b), and Placerville (PLC) for year 2014 (c); BT1, BT2, BT3: bark texture replicates 1, 2, 3

**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 | 13,756,415–13,759,891 | 3 | PLCV14_BT2 | 13,759,891 | *8.44* |
| | | | | CLA12_BT2 | 13,756,415 | 5.98 |
| | | | | CLA13_BT2 | 13,756,415 | 5.22 |
| 2 | XVII | 2,930,303–2,942,559 | 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 | 48,502,337–48,513,466 | 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 | 44,916,073–44,932,416 | 16 | CLA12_BT2 | 44,931,643 | *7.40* |
| | | | | CLA13_BT2 | 44,931,643 | *6.44* |
| 8 | XVII | 2,894,584–2,915,048 | 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 | 11,209,007–11,217,624 | 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 | 21,529,197–21,533,794 | 5 | CLA12_BT2 | 21,530,675 | *7.00* |
| | | | | CLA13_BT2 | 21,530,675 | *6.41* |
| | | | | CLA15 | 21,530,272 | 5.24 |
| 12 | XVI | 5,614,771–5,630,821 | 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 | 8,131,399–8,141,525 | 10 | CLA12cookieBT1 | 8,131,399 | *6.75* |
| | | | | CLA12_BT3 | 8,131,399 | *6.24* |
| | | | | CLA13_BT1 | 8,141,525 | 5.33 |
| 15 | VII | 16,608,390–16,609,078 | 1 | CLA13_BT2 | 11,609,206 | *6.76* |
| | | | | CLA12_BT2 | 11,609,206 | 5.40 |
| 16 | IV | 21,617,052–21,619,943 | 3 | CLA13_BT2 | 21,619,943 | *6.73* |
| | | | | CLA12_BT1 | 21,619,943 | 5.86 |
| 17 | V | 21,642,667–21,646,857 | 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 | 15,160,420–15,163,613 | 3 | CLA13_BT1 | 15,160,420 | *6.49* |
| | | | | CLA12_BT1 | 15,160,420 | 5.45 |
| 19 | XVIII | 6,040,783–6,046,262 | 6 | CLA12_BT1 | 6,042,491 | *6.47* |
| | | | | CLA13_BT1 | 6,042,491 | 5.01 |
| 20 | VI | 27,823,761–27,839,980 | 16 | PLCV14_BT2 | 27,823,761 | *6.33* |
| | | | | COR13_BT2 | 27,839,980 | *6.31* |
| | | | | COR13_BT1 | 27,837,548 | 5.38 |
| 21 | IV | 8,874,775–8,883,033 | 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* |

Table 1 (continued)

| Region # | LG | SNP position interval | Interval (kb) | Population | SNP position | Corr. −log10 *p* |
|---|---|---|---|---|---|---|
| | | | | 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 | 15,754,964–15,758,609 | 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 | 10,028,445–10,032,564 | 4 | CLA13_BT2 | 10,028,892 | *6.00* |
| | | | | CLA12cookieBT1 | 10,028,445 | 5.27 |
| 30 | I | 31,275,502–31,277,465 | 2 | CLA12_BT2 | 31,276,103 | *7.49* |
| 31 | XVII | 3,001,881–3,008,293 | 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 | 4,752,983–4,762,402 | 10 | CLA12_BT3 | 4,762,402 | *7.44* |
| | | | | CLA12cookieBT1 | 4,762,402 | 5.42 |
| 34 | XI | 10,996,808–10,996,815 | 1 | CLA12_BT3 | 10,996,815 | *7.26* |
| | | | | CLA12cookieBT1 | 10,996,815 | *6.43* |
| 35 | XIX | 12,020,500–12,025,588 | 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 | 12,838,532–12,843,036 | 5 | COR13_BT2 | 12,838,536 | *6.36* |
| | | | | COR13_BT1 | 12,843,036 | 5.19 |
| 38 | XVI | 2,627,146–2,630,738 | 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 | 15,467,058–15,489,644 | 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 italics ($-\log10\ p \geq 6$). The interval in which significant SNPs were detected at a significance threshold of $-\log10\ p \geq 5$, 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 of 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 |

Table 2 (continued)

| Region # | Population | Highest, log10 *p* | SNP within gene feature | Gene name | Putative function | Arabidopsis ortholog |
|---|---|---|---|---|---|---|
| 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 | Nonr |
| 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 (three 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 italics. Populations, the name indicates site and year; CDS, coding sequences; S, synonymous; NS, nonsynonymous; NA, not applicable since these reproducible regions (− log10 of ≥ 5) are not part of the 41 highly significant regions (− log10 of ≥ 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

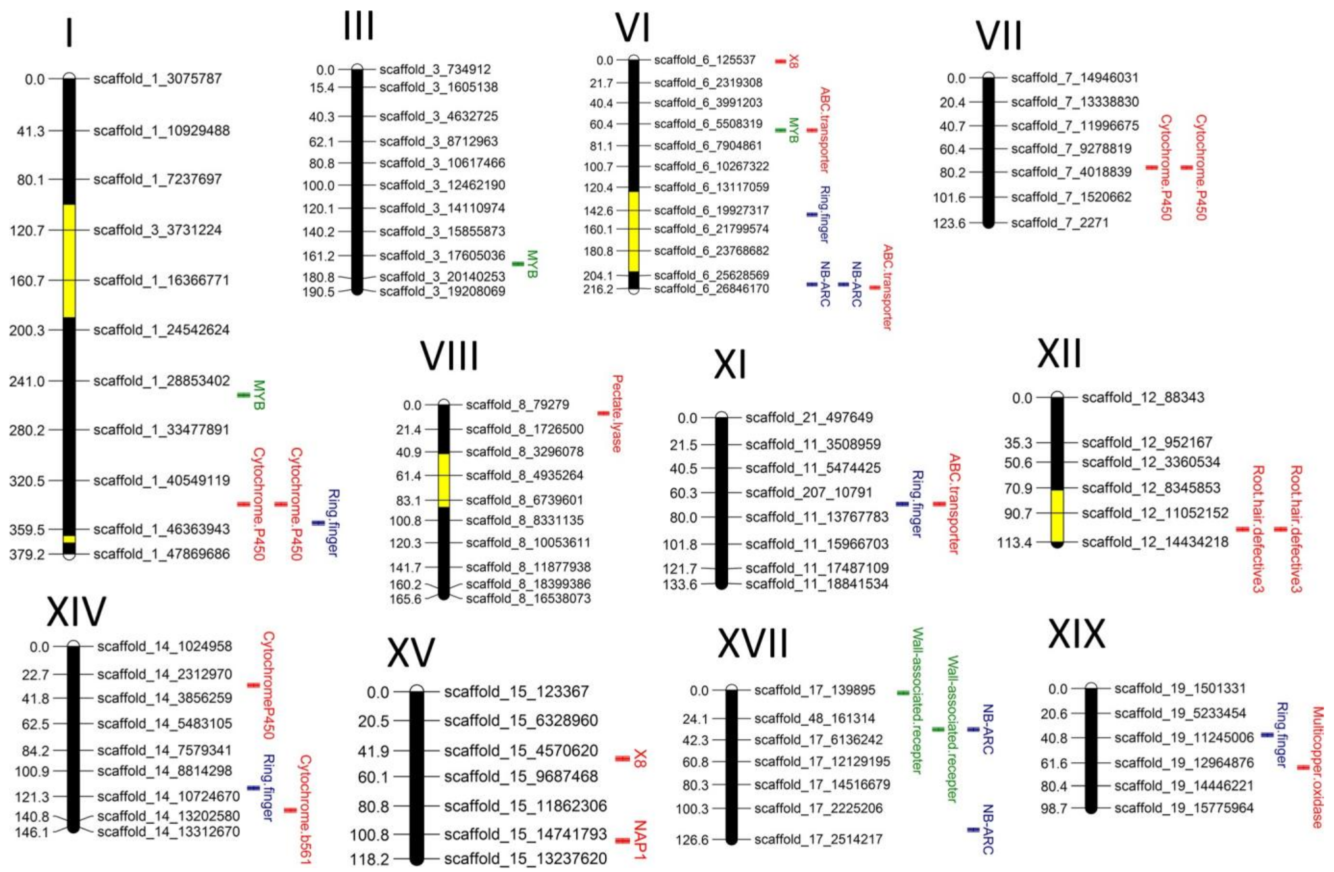


Fig. 2 Location of top candidate genes on the linkage map derived from across of a female hybrid clone (*Populus trichocarpa* × *P. deltoides*) with *P. deltoides* (Bdeir et al. 2017). The linkage map is anchored to the genome of *Populus trichocarpa* (V3 assembly, Kelleher et al. 2007; see, Table 1). 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 andtissue differentiation, suberin accumulation, and programmed cell death,respectively

## References


Addicott FT (1982) Abscission. Univ of California Press, Berkeley

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

Asif M, al-Mansoub MA, Khan MDSS, Yehya AHS, Ezzat MO, Oon CE, Atif M, Abdul Majid AS, Abdul Majid AMS (2017) Molecular mechanisms responsible for programmed cell death-inducing attri- butes of terpenes from *Mesua ferrea* stem bark towards human co- lorectal carcinoma HCT 116 cells. J Appl Biomed 15:71–80

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

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265

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

Beaver R, Wilding N, Collins N, Hammond P, Webber J (1989) Insect- fungus relationships in the bark and ambrosia beetles. Insect-fungus interactions:121–143

Bent AF (1996) Plant disease resistance genes: function meets structure. Plant Cell 8:1757–1771

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

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

Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17:2281–2295

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

Caño-Delgado A, Lee J-Y, Demura T (2010) Regulatory mechanisms for specification and patterning of plant vascular tissues. Annu Rev Cell Dev Biol 26:605–637

Carlquist S (1996) Wood, bark, and stem anatomy of Gnetales: a summa- ry. Int J Plant Sci 157:S58–S76

Carpita N, Tierney M, Campbell M (2001) Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. Plant Mol Biol 47:1–5

Chaffey N, Cholewa E, Regan S, Sundberg B (2002) Secondary xylem development in *Arabidopsis*: a model for wood formation. Physiol Plant 114:594–600

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

Climent J, Tapias R, Pardos JA, Gil L (2004) Fire adaptations in the Canary Islands pine (*Pinus canariensis*). Plant Ecol 171:185–196

Clouse SD, Sasse JM (1998) Brassinosteroids: essential regulators of plant growth and development. Annu Rev Plant Biol 49:427–451

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

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

Evans LM, Slavov GT, Rodgers-Melnick E, Martin J, Ranjan P, Muchero W, Brunner AM, Schackwitz W, Gunter L, Chen JG, Tuskan GA, DiFazio SP (2014) Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. Nat Genet 46:1089–1096

Fitch R (2006) WinSTAT for Excel. The statistics add-in for Microsoft Excel R Fitch Software

Fossdal CG, Yaqoob N, Krokene P, Kvaalen H, Solheim H, Yakovlev IA (2012) 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 Biol 12:105

Franciosini A, Lombardi B, Iafrate S, Pecce V, Mele G, Lupacchini L, Rinaldi G, Kondou Y, Gusmaroli G, Aki S, Tsuge T, Deng XW, Matsui M, Vittorioso P, Costantino P, Serino G (2013) The *Arabidopsis* COP9 SIGNALOSOME INTERACTING F-BOX KELCH 1 protein forms an SCF ubiquitin ligase and regulates hypocotyl elongation. Mol Plant 6:1616–1629

Gill A, Ashton D (1968) The role of bark type in relative tolerance to fire of three central Victorian eucalypts. Aust J Bot 16:491–498

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

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

Groover A, Jones AM (1999) Tracheary element differentiation uses a novel mechanism coordinating programmed cell death and second- ary cell wall synthesis. Plant Physiol 119:375–384

Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlen M, Teeri TT, Lundeberg J, Sundberg B, Nilsson P, Sandberg G (2001) A transcriptional roadmap to wood formation. Proc Natl Acad Sci 98:14732−14737

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

Jelonek T, Pazdrowski W, Tomczak A, Arasimowicz-Jelonek M (2017) Lignification markers of the tracheid walls of scots pine (*Pinus sylvestris* (L.)) in various forms of dead bark. BioResources 12: 3992−4003

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

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

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

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

Lawes MJ, Adie H, Russell-Smith J, Murphy B, Midgley JJ (2011) How do small savanna trees avoid stem mortality by fire? The roles of stem diameter, height and bark thickness. Ecosphere 2:1−13

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

Li Y, Jin F, Chao Q, Wang B-C (2017) Proteomics analysis reveals the molecular mechanism underlying the transition from primary to sec- ondary growth of poplar. J Plant Physiol 213:1−15

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

Marshall JM, Smith EL, Mech R, Storer AJ (2013) Estimates of *Agrilus planipennis* infestation rates and potential survival of ash. Am Midl Nat 169:179−193

McCarthy RL et al (2010) The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of second- ary wall biosynthesis. Plant Cell Physiol 51:1084−1090

McIntyre LM, Martin ER, Simonsen KL, Kaplan NL (2000) Circumventing multiple testing: a multilocus Monte Carlo approach to testing for association. Genet Epidemiol 19:18−29

Mellerowicz EJ, Baucher M, Sundberg B, Boerjan W (2001) Unravelling cell wall formation in the woody dicot stem. Plant Mol Biol Report 47:239−274

Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. Proc Natl Acad Sci 95:6619−6623

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

Muchero W, Guo J, DiFazio SP, Chen JG, Ranjan P, Slavov GT, Gunter LE, Jawdy S, Bryan AC, Sykes R, Ziebell A, Klápště J, Porth I, Skyba O, Unda F, el-Kassaby YA, Douglas CJ, Mansfield SD, Martin J, Schackwitz W, Evans LM, Czarnecki O, Tuskan GA (2015) High-resolution genetic mapping of allelic variants associat- ed with cell wall chemistry in *Populus*. BMC Genomics 16:24

Pausas JG (2015) Bark thickness and fire regime. Funct Ecol 29:315−327

Piquemal J, Lapierre C, Myton K, O'connell A, Schuch W, Grima- pettenati J, Boudet AM (1998) Down-regulation of cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. Plant J 13:71−83

Plomion C, Leprovost G, Stokes A (2001) Wood formation in trees. Plant Physiol 127:1513−1523

Pollard M, Beisson F, Li Y, Ohlrogge JB (2008) Building lipid barriers: biosynthesis of cutin and suberin. Trends Plant Sci 13:236−246

Rains MK, de Silva G, Dilini N, Molina I (2018) Reconstructing the suberin pathway in poplar by chemical and transcriptomic analysis of bark tissues. Tree Physiol 38:340−361

Romero C (2006) Tree responses to stem damage. PhD thesis. University of Florida, Gainsville, USA

Sengupta G, Palit P (2004) Characterization of a lignified secondary phloem fibre-deficient mutant of jute (*Corchorus capsularis*). Ann Bot 93:211−220

Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers- Melnick E, Lipphardt MF, Pennacchio CP, Hellsten U, Pennacchio LA, Gunter LE, Ranjan P, Vining K, Pomraning KR, Wilhelm LJ, Pellegrini M, Mockler TC, Freitag M, Geraldes A, el-Kassaby YA, Mansfield SD, Cronk QCB, Douglas CJ, Strauss SH, Rokhsar D, Tuskan GA (2012) Genome resequencing reveals multiscale geo- graphic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. New Phytol 196:713−725

Soler M, Serra O, Molinas M, Huguet G, Fluch S, Figueras M (2007) A genomic approach to suberin biosynthesis and cork differentiation. Plant Physiol 144:419−431

Song D, Xi W, Shen J, Bi T, Li L (2011) Characterization of the plasma membrane proteins and receptor-like kinases associated with sec- ondary vascular differentiation in poplar. Plant Mol Biol 76:97−115 Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, van Montagu M, Sandberg G, Olsson O, Teeri

TT, Boerjan W, Gustafsson P, Uhlen M, Sundberg B, Lundeberg J (1998) Gene discovery in the wood- forming tissues of poplar: analysis of 5,692 expressed sequence tags. Proc Natl Acad Sci 95:13330−13335

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci 100:9440–9445

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

Turner SR, Somerville CR (1997) Collapsed xylem phenotype of *Arabidopsis* identifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell 9:689–701

Van Ooijen G, Mayr G, Kasiem MM, Albrecht M, Cornelissen BJ, Takken FL (2008) Structure–function analysis of the NB-ARC do- main of plant disease resistance proteins. J Exp Bot 59:1383–1397 Venkatesan T, Choi Y-W, Mun S-P, Kim Y-K (2016) *Pinus radiata* bark extract induces caspase-independent apoptosis-like cell death in MCF-7 human breast cancer cells. Cell Biol Toxicol 32:451–464