Functional characterization of PtMYB115, a regulator of condensed tannin synthesis in poplar

by

Amy Midori Franklin
B.Sc., University of Victoria, 2011

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science

in the Department of Biology

Supervisory Committee

Functional characterization of PtMYB115, a regulator of condensed tannin synthesis in poplar

by

Amy Midori Franklin
B.Sc., University of Victoria, 2011

Supervisory Committee

Dr. C. Peter Constabel, Department of Biology
Supervisor

Dr. Jürgen Ehlting, Department of Biology
Departmental Member

Dr. Armand Séguin, Department of Biology
Departmental Member

Dr. Christopher Nelson, Department of Biochemistry and Microbiology
Outside Member

Abstract

Supervisory Committee
Dr. C. Peter Constabel, Department of Biology
Supervisor
Dr. Jürgen Ehlting, Department of Biology
Departmental Member
Dr. Armand Séguin, Department of Biology
Departmental Member
Dr. Christopher Nelson, Department of Biochemistry and Microbiology
Oustide Member

Condensed tannins are wide-spread polyphenols with diverse ecological functions, including defense against herbivores and microbial pathogens. In poplar, condensed tannin synthesis is induced by a variety of biotic and abiotic stresses. The objective of this study was to determine the function of the R2R3 MYB transcription factor MYB115 in the regulation of condensed tannin synthesis. MYB115 was shown to be induced by wounding along with tannin biosynthetic genes and shows sequence similarity to characterized regulators of tannin synthesis in grape and persimmon suggesting that it functions in the regulation of condensed tannin synthesis. To analyze the function of MYB115, transgenic plants overexpressing MYB115 were generated and showed enhanced accumulation of condensed tannins and higher expression of flavonoid biosynthetic genes involved in condensed tannin biosynthesis compared to wild-type control plants. In promoter activation assays, MYB115 activated the promoter of a tannin-specific biosynthetic enzyme, anthocyanidin reductase. This suggests that MYB115 acts as a regulator of condensed tannin synthesis. MYB115 overexpressors showed additional changes to phenolic metabolism, including changes in levels of phenolic glycosides and hydroxycinnamic acids. These results indicate an important role of MYB115 in the regulation of the condensed tannin pathway in poplar.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank my supervisor, Dr. C. Peter Constabel for giving me the opportunity to work on this exciting project and for his valuable advice and support throughout the project. I would also like to thank my committee members Drs. Jürgen Ehlting, Chris Nelson, and Armand Séguin for their helpful advice. Thank you to the members of Armand Séguin's lab for their assistance with analyzing the microarray data. Thank you to Dr. Michael Reichelt at the Max Planck Institute for Chemical Ecology in Jena, Germany for performing LC/MS analysis. Thank you to Jane Guo for performing *in silico* co-expression analysis for MYB115. Thank you to Brad Binges for technical assistance in the greenhouse. Thank you to Dr. Andreas Gesell, Lan Tran, and Dr. Kazuko Yoshida for construction of vectors and also for teaching and mentoring me throughout my degree. Thank you to Dr. Vincent Walker for his guidance with HPLC analysis. I would also like thank to all past and present members of the Constabel lab for guidance, support and technical assistance.

1. Chapter One: General introduction

1.1 Poplar as a model system for the study of the regulation of defense responses in trees

The genus *Populus* includes approximately 30 species which include the poplars, cottonwoods and aspens (referred to simply as poplar here). Poplars are eudicots that belong to the family Salicaceae, which includes *Populus* spp., *Salix* spp. and *Chosenia arbutifolia*. Poplar species are widespread across the Northern Hemisphere. In North America, *Populus tremuloides* is the most geographically wide-spread native tree (Brunner et al., 2004). Poplars are considered to be foundation species due to their ability to rapidly colonize disturbed sites and generate stable conditions for colonization by other species (Bradshaw et al., 2000; Whitham et al., 2006). They have historically been an economically important crop for pulp and paper production. They also have potential uses in biofuel production.

Poplar is an ideal model system for studying tree biology. While *Arabidopsis* remains a key resource for studying basic plant biology due to the abundant availability of genetic resources, other aspects of plant biology that are specific to trees are better studied in poplar. For example, poplar is a better model for studying extensive secondary xylem formation and differences in development due to a perennial life cycle (Bradshaw et al., 2000).

Poplar plants are also a useful experimental system as they can be vegetatively propagated, are easily grown in tissue culture and are amenable to genetic transformation via infection with *Agrobacterium tumefaciens*, which is a naturally occurring plant pathogen that can transfer a segment of foreign DNA into the plant where it is then stably integrated into the plant genome. In fact, poplar were the first trees to be genetically transformed and regenerated (Fillatti et al., 1987).

Due to its many desirable traits including fast growth and susceptibility to genetic transformation, poplar is an ideal target for genetic engineering. Plants with increased resistance to harmful pathogens and pests can be developed through

manipulation of metabolic pathways. Furthermore, the future of biofuels could benefit from the bioengineering of plants with improved characteristics, for example high biomass yield and reduced lignin content (Sannigrahi et al., 2010).

Many genetic resources for poplar are available including extensive EST collections and the complete sequenced genome of *Populus trichocarpa*. The first draft version of the sequenced genome of *P. trichocarpa* was released in 2001 by the DOE Joint Genome Institute. An assembly of the poplar genome was released in 2006 (Tuskan et al., 2006). In August 2012, a second assembly was released (Phytozome, v3.0 assembly). The availability of comprehensive EST libraries and the sequencing of *the P. trichocarpa* genome facilitated the development of whole genome microarrays. Comprehensive *in silico* expression data is available through tools such as the Poplar Expression Browser from The Bio-Analytic Resource for Plant Biology (BAR). Furthermore, next generation sequencing techniques, such as RNA-seq, allow for even more sensitive analysis of changes in gene expression that could not be achieved through microarray analysis (Wang et al., 2009).

1.2 Phenylpropanoid pathway

Phenylpropanoids are a diverse group of compounds derived from phenylalanine, a product of the shikimate pathway (Figure 1-1). Phenylpropanoids have a wide variety of functions and structures. They can be constitutively synthesized such as the secondary cell wall component lignin, and can also be produced in response to stress. The large diversity of compounds in the phenylpropanoid pathway can be attributed to the modification of a set of core structures by various classes of enzymes that function as transferases, oxygenases, and ligases (Vogt, 2010).

Here, I discuss the biosynthesis and ecological functions of three classes of phenylpropanoids that are involved in chemical plant defense in poplar: salicinoid phenolic glycosides, hydroxycinnamic acids, and flavonoids including condensed tannins. Other classes of phenylpropanoids not discussed here include the cell wall components lignin and suberin, as well as many other plant defense compounds (Vogt, 2010).

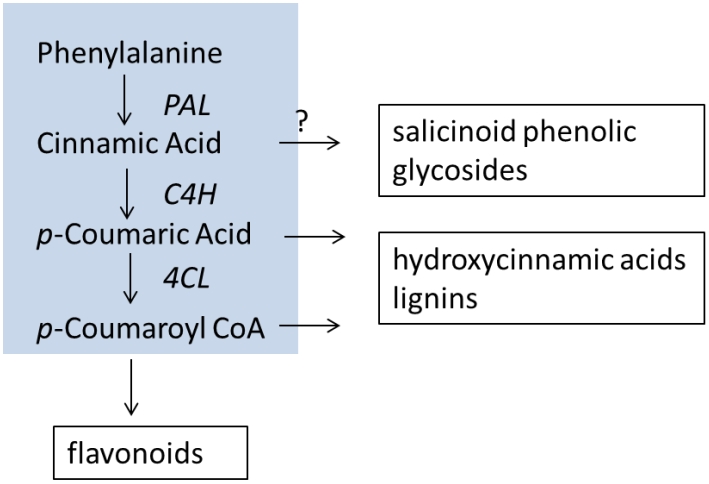

Figure 1-1. Generalized phenylpropanoid pathway
The initial reactions of phenylpropanoid biosynthesis are highlighted in blue. The question mark indicates that this pathway has not been fully characterized. See Figure 1-2 for a detailed representation of the flavonoid pathway. See Figure 1-3 for structures of major phenolic glycosides in poplar. See Figure 1-4 for structures of major hydroxycinnamic acids in poplar. PAL: phenylalanine ammonia lyase. C4H: trans-cinnamate 4-monooxygenase. 4CL: 4-coumaroyl:CoA-ligase.

1.3 Flavonoid and condensed tannin biosynthesis

Flavonoids are a class of compounds that form a branch of the phenylpropanoid pathway. They share a basic structure of C6-C3-C6 phenyl-benzopyran backbone. Flavonoid compounds are derived from chalcone following the condensation of coumaryl-CoA and malonyl-CoA by chalcone synthase (CHS) (Ferrer et al., 1999; Figure 1-2). This reaction is followed by isomerization by chalcone isomerase (CHI) to form flavanones (Jez et al., 2000). Flavanones are precursors of flavones, isoflavones, flavan-4-ols and dihydroflavonols. Flavones, isoflavones and flavan-4-ols represent different end-points of the flavonoid pathway and are therefore not directly involved in the synthesis of condensed tannins. Dihydroflavonols, however, are involved in the synthesis of condensed tannins. Hydroxylation of flavanones by flavanone 3β-hydroxylase (F3H) forms the dihydroflavonols (Lukačin et al., 2000). Dihydroflavonols

can be converted to flavonols by flavonol synthase (FLS) (Holton et al., 1993). Flavonols represent an end point of the flavonoid pathway distinct from condensed tannin synthesis. Alternatively, dihydroflavonols can be converted to leucoanthocyanidins, early precursors of condensed tannins, by the NADPH-dependent dihydroflavonol 4-reductase (DFR)(Fischer et al., 1988).

Condensed tannins, also known as proanthocyanidins, are polymers of flavan-3-ols. Tannin structure varies in composition and length between species (Scioneaux et al., 2011). While branched tannins are common in other plant species such as quebracho, condensed tannins in poplar are linear polymers (Pasch et al., 2001; Scioneaux et al., 2011). The biochemical mechanism that determines polymer length has not yet been characterized. Tannin structure is dependent on subunit composition which is largely characterized by the hydroxylation pattern of the B-ring. Hydroxylation is dependent on the flavonoid hydroxylases, flavonoid 3'-hydroxylase (F3'H) and flavonoid 3'5'-hydroxylase (F3'5'H) (Figure 1-2). The three most common flavan-3-ols in poplar are epicatechin, epigallocatechin and gallocatechin (Schweitzer et al., 2008). Other flavan-3-ols are catechin, afzelechin and epiafzelechin; however, afzelechin and epiafzelechin have not been reported in poplar. The various flavan-3-ols are synthesized by two different enzymes, leucoanthocyanidin reductase (LAR) and anthocyanidin reductase (ANR), which typically can reduce precursors with the various hydroxylation patterns.

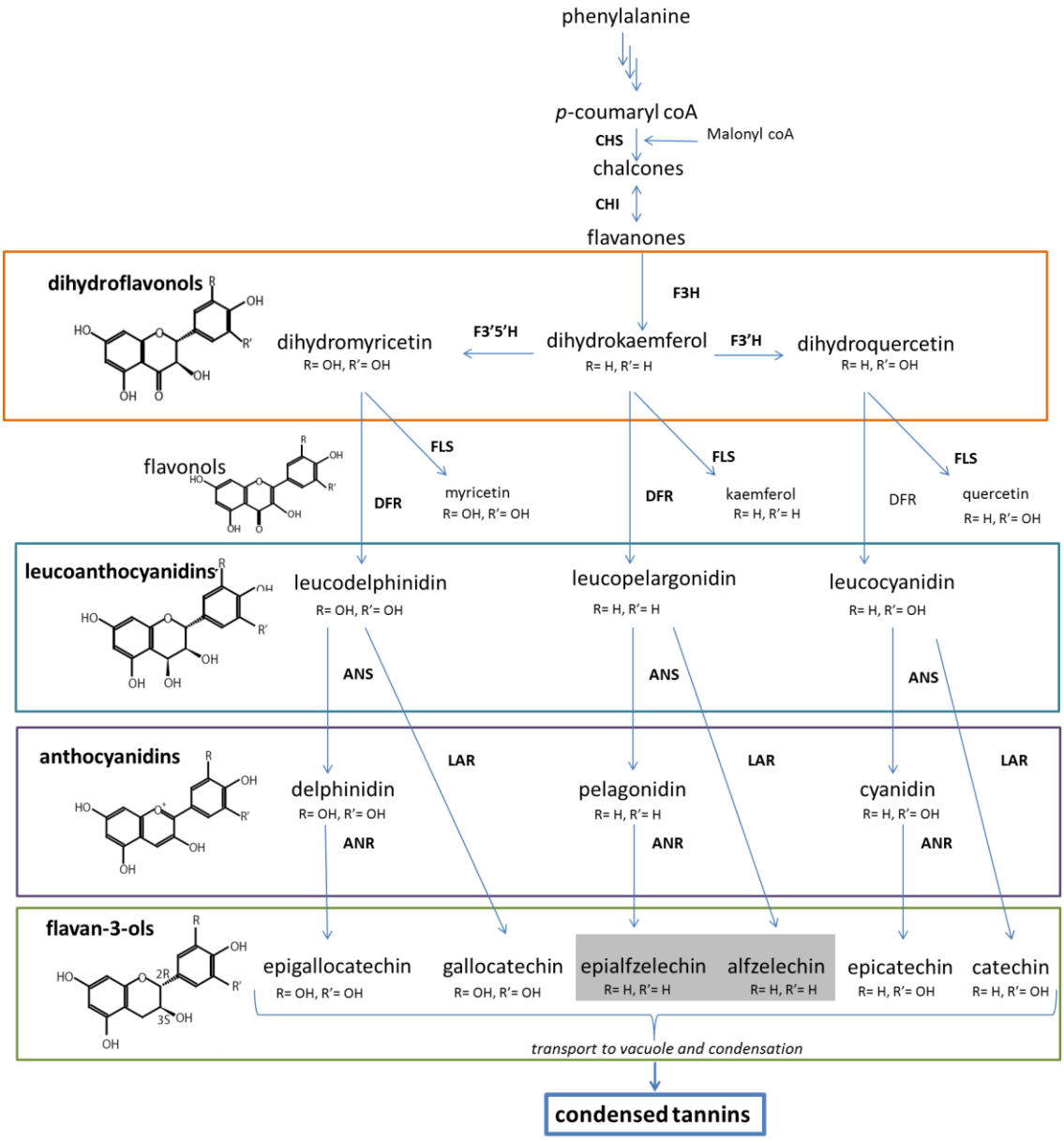

Figure 1-2. A simplified representation of the flavonoid pathway leading the synthesis of condensed tannins.

Only the *trans*- conformation of flavan-3-ols is drawn here. Epialfzelechin and alfzelechin are shaded because they have not been shown to accumulate in poplar. CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3' hydroxylase; F3'5'H, flavonoid 3'5' hydroxylase. FLS, flavonol synthase; DFR, dihydroflavonol reductase; ANS, anthocyanidin synthase; UFGT, UDP-glucose flavonoid glucosyltransferase; ANR, anthocyanidin reductase; LAR, leucoanthocyanidin reductase.

Leucoanthocyanidins are substrates for leucoanthocyanidin reductase (LAR). LAR removes the 4-hydroxyl from leucoanthocyanidin to form 2,3-*trans*-flavan-3-ols (catechin) which are direct precursors of condensed tannins. LAR was first characterized in the legume *Desmodium uncinatum* (Tanner et al., 2003). Recombinant LAR protein incubated with leucocyanidin produced catechin (Tanner et al., 2003). LAR genes have since been characterized in grapevine (Bogs et al., 2005) and *Medicago* (Pang et al., 2007). Arabidopsis lacks an LAR. Three putative genes encoding LAR are found in the poplar genome (Tsai et al., 2006). Two of the three genes, PtLAR1 and PtLAR3, have been shown to function in the synthesis of catechin (Yuan et al., 2012;Wang et al., 2013).

Anthocyanidins are synthesized via reduction of leucoanthocyanidins by anthocyanidin synthase (ANS). Anthocyanins, which represent another end-point of the flavonoid pathway, are formed via glycosylation of anthocyanidins. Flavan-3-ols are synthesized from anthocyanidins. 2,3-*cis*-flavan-3-ols (epicatechin, epigallocatechin and epiafzelechin) are formed from reduction of anthocyanidins by anthocyanidin reductase.

Anthocyanidin reductase (ANR) was first characterized through analysis of mutants that prematurely accumulated anthocyanins in the immature seed coat of Arabidopsis plants (Devic et al., 1999). These mutants had a mutation in the BANYULS gene. The BANYULS gene was originally thought to encode for a leucoanthocyanidin reductase until a study by Xie et al. (2004), demonstrated that recombinant BANYULS protein can catalyze the formation of epicatechin from cyanidin. This shows that BANYULS encodes ANR not LAR. There are two putative genes encoding for ANR in the poplar genome (Tsai et al., 2006). ANR1 was recently characterized in poplar and shown to synthesize epicatechin (Wang et al., 2013).

Most flavonoids localize to the vacuole while immunolocalization studies suggest that some flavonoid biosynthetic enzymes are bound to the endoplasmic reticulum as well as to the tonoplast (Burbulis and Winkel-Shirley, 1999; Saslowsky and Winkel-Shirley, 2001). Flavonoids are likely shuttled into the vacuole by a transporter. The multidrug and toxic compound extrusion family protein, AtTT12, has been characterized

as an H^+-antiporter of glycosylated anthocyanins and glycosylated flavan-3-ols into the large central vacuole (Sharma and Dixon, 2005; Marinova et al., 2007). The ATP dependent proton pump, AtAHA10, is suggested to generate the proton gradient necessary for the activity of TT12 (Baxter et al., 2005). As only glycosylated flavan-3-ols are substrates for AtTT12, a glycosyltransferase is likely involved in flavan-3-ol modification prior to transport (Pang et al., 2013). Further modification may be necessary for transport as a glutathione S-transferase in Arabidopsis, AtTT19, has been characterized to function in modification of flavan-3-ols and anthocyanins prior to transport into the vacuole (Kitamura et al., 2004). The enzymes necessary for transport and condensation of tannins have not been characterized in poplar. However, a putative homolog of AtTT12 is induced by wounding and pathogen stress and shows higher expression in high tannin transgenic plants overexpressing the AtTT2 homolog, PtMYB134 (Mellway et al., 2009).

1.4 Ecological significance of condensed tannins and other flavonoids

1.4.1 Condensed tannins and protection against herbivory

Condensed tannins accumulate in many, but not all, plants. They can accumulate in many plant organs such as leaves, roots, fruits and bark while in some plants tannin accumulation is isolated to specific tissues, such as in Arabidopsis in which tannins only accumulate in the seed coat (Porter, 1992). Condensed tannins accumulate in response to environmental changes. The focus of many studies on condensed tannins is on their potential role in plant-insect interactions. An early study by Feeny (1970) examined the effect of oak tannin levels on feeding by leaf chewing lepidopterous insects. Feeny (1970) found that an increase in insect abundance correlated with a decrease in foliar tannins. Additionally, the same author (Feeny, 1968) found that larvae fed an artificial diet showed decreased growth when fed tannins. These studies suggested that insect fitness and growth are negatively influenced by tannin concentration. Since then, many studies have focussed on tannin-insect interactions (reviewed by Barbehenn and Constabel, 2011). However, other studies have shown that tannin concentration does

not influence feeding on poplar by lepidopterous insects (Hwang and Lindroth, 1997; Osier and Lindroth, 2004). For example, Hwang and Lindroth (1997) found gypsy moth larvae consumption rate correlated with phenolic glycoside concentrations but not condensed tannin concentrations.

In poplar, as in many other plants, condensed tannin structure differs between species (Ayres et al., 1997; Scioneaux et al., 2011). The diversity in structure of condensed tannins could explain the contradictory results of many tannin-herbivore studies. Condensed tannin structure is dependent on chain length, subunit composition and the position of the linkage between subunits. Subunit composition varies depending on the presence of F3'H and F3'5'H which modify the hydroxylation pattern of the B-ring. Ayres et al. (1997) tested the anti-herbivore activity of condensed tannins isolated from 16 woody plants. The authors found that condensed tannin structure greatly influenced anti-herbivore activity. For example, they found that condensed tannins derived from *P. tremuloides*, which has a higher ratio of dihydroxylated subunits to trihydroxylated subunits, had a negative influence on the growth rate of *Chrysomela falsa,* compared to condensed tannins derived from *P. balsamifera* which positively influenced growth rate.

There are a number of hypotheses on how condensed tannins may act to deter pests and include reducing nutrient availability and having direct toxic effects. Tannins are known to form complexes with proteins and could potentially bind digestive enzymes or their substrates in insect guts (Hagerman et al., 1998). Early studies investigated the roles of condensed tannins as anti-nutritive compounds (Feeny, 1969). However, few studies were able to show evidence of any anti-nutritive effect on insect herbivores (Bernays et al., 1981). In mammals, tannins can bind proteins and decrease digestion, although the anti-nutritive effect is mediated by salivary tannin-binding proteins (McArt et al., 2009; Shimada, 2006). This difference in the ability of tannins to bind protein in the guts of mammals but not insects is likely due to differences in chemical conditions, such as pH (Martin et al., 1985).

The negative effect of tannins may be the result of toxicity via oxidation of condensed tannins in the insect gut. Phenolic compounds may be oxidized in the insect gut to form semiquinone radicals and quinones which can damage nutrients or, if absorbed, cause cellular damage (Barbehenn et al., 2005). Oxidation of condensed tannins in insect guts has not been studied widely; however, caterpillars fed condensed tannins show increased concentrations of semiquinone radicals (Barbehenn et al., 2008). The formation of reactive oxygen species from condensed tannins is likely dependent on numerous factors such as chemical structure and gut conditions (eg. pH and presence of oxidants). To determine if condensed tannins function as prooxidants post-ingestion, further studies correlating the production of reactive oxygen species from tannins with insect performance are necessary.

1.4.2 Condensed tannins and other flavonoids protect against microbial pathogen stress

Studies have shown that flavonoid biosynthetic genes are induced following attack by fungal pathogens (Azaiez et al., 2009; Miranda et al., 2007). Flavonoids may function as antimicrobial compounds through prevention of tissue degradation by binding to microbial enzymes or to metal ions necessary for enzyme function. Other studies have shown that plants that are lower in tannins are more susceptible to pathogen stress (Skadhauge et al., 1997; Yuan et al., 2012). In a study by Yuan et al. (2012), overexpression of LAR3 in poplar led to the increased accumulation of condensed tannins. LAR3 overexpressing plants showed increased resistance to the fungal pathogen, *Marssonina brunnea* f.sp. *multigermtubi.* Similarly, mutant barley seeds deficient in condensed tannins were more susceptible to infection by *Fusarium* (Skadhauge et al*.,* 1997). These studies suggest that condensed tannins can function as anti-pathogen compounds.

1.4.3 Flavonoids and light stress

The role of flavonoids in plant defense has been widely studied with regard to light stress. Excess light leads to oxidative stress from free radicals when more light is absorbed than what can be used for photosynthesis. Flavonoids, mainly flavonols, can

protect the plant from light stress by acting as anti-oxidants, or by acting as a physical barrier by absorbing light (reviewed by Harborne and Williams, 2000). Studies of mutant plants deficient in flavonols or plants that accumulate flavonols to high levels have shown that flavonols likely have an important role in UV protection.

A study by Bieza and Lois (2001) showed that Arabidopsis mutants that accumulate flavonoids at enhanced levels were more tolerant to UV stress. These mutants had constitutively higher expression of the early flavonoid biosynthetic gene, chalcone synthase (CHS). Arabidopsis CHS mutants deficient in flavonols were highly sensitive to UV light (Li et al., 1993). Flavonol glycosides have also been shown to be induced by UV stress in poplar (Warren et al., 2003). These studies suggest that both flavonols and flavonol glycosides are key factors in protection against UV light.

Flavonols have been widely studied as UV protectants while few studies have examined the role of condensed tannins (Ryan et al., 2001; Solovchenko and Schmitz-Eiberger, 2003). However, condensed tannins have been implicated as UV protectants as well. Poplar plants exposed to UV light show induction of the tannin specific biosynthetic genes ANR and LAR (Kim et al., 2012; Mellway et al., 2009; Zhang et al., 2013). This suggests that tannins may also have a role in plant response to UV stress.

1.5 Ecological significance of phenolic glycosides

Phenolic glycosides are a class of secondary metabolites that have a core structure of a salicyl alcohol with a glucose moiety (Figure 1-3). The simplest phenolic glycoside is salicin and because of this salicin based phenolic glycosides have been named “salicinoids” (Boeckler et al., 2011). Phenolic glycosides are synthesized at high levels in poplar, and have been reported to accumulate up to 30% leaf dry weight in poplar (Donaldson et al., 2006). Biosynthesis of phenolic glycosides is unknown; however, a recent study has suggested that phenolic glycosides are derived from the phenylpropanoid pathway. By feeding poplar leaves labelled precursors, a study by Babst et al. (2010) suggests that the salicyl moiety of salicin arises from cinnamic acid

and therefore phenolic glycosides synthesis might be a branch of the phenylpropanoid pathway.

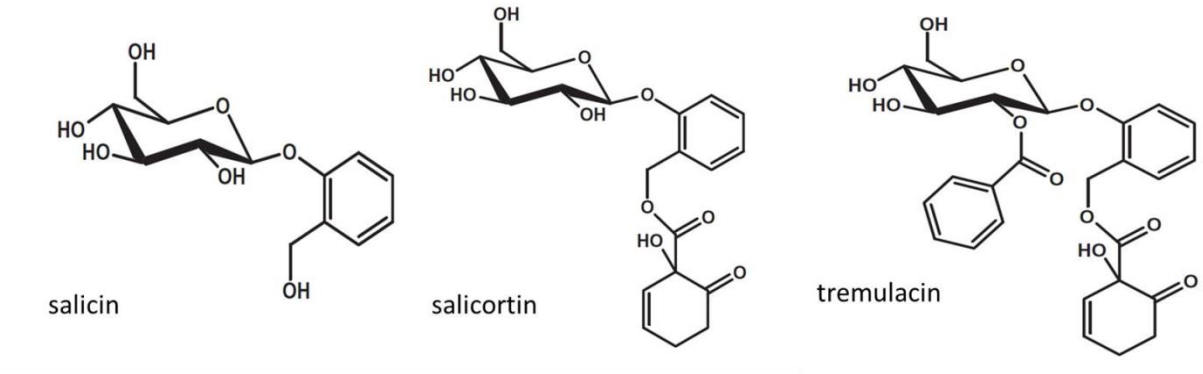

Figure 1-3. Structures of key phenolic glycosides in poplar.

The role of phenolic glycosides as anti-herbivore defense compounds has been widely studied (see Boeckler et al., 2011, for a comprehensive review). There are many examples of phenolic glycosides acting as feeding deterrents against insect herbivores. High levels of phenolic glycosides has been correlated with reduced feeding and reduced fecundity of lepidopteran herbivores (Osier et al., 2000; Young et al., 2010). In insects, some phenolic glycosides are proposed to have toxic effects through the production of 6-hydroxy-2-cyclohexen-on-oyl or o-quinones (Clausen et al., 1990; Haruta et al., 2001). These products of digestion may cause direct damage to the insect gut or bind to digestive enzymes. Contrastingly, phenolic glycosides can stimulate feeding by specialist herbivores that can assimilate phenolic glycosides for their own defense (Prudic et al., 2007).

1.6 Ecological significance of hydroxycinammic acids

Hydroxycinnamic acids (HCA) have a C6-C3 skeleton. Some are precursors to lignin, an integral component of wood. HCAs are often cell wall components where they protect against pathogen attack as well as conferring cell wall integrity. The three major HCAs in poplar are caffeic acid, *p*-coumaric acid and ferulic acid (Figure 1-4).

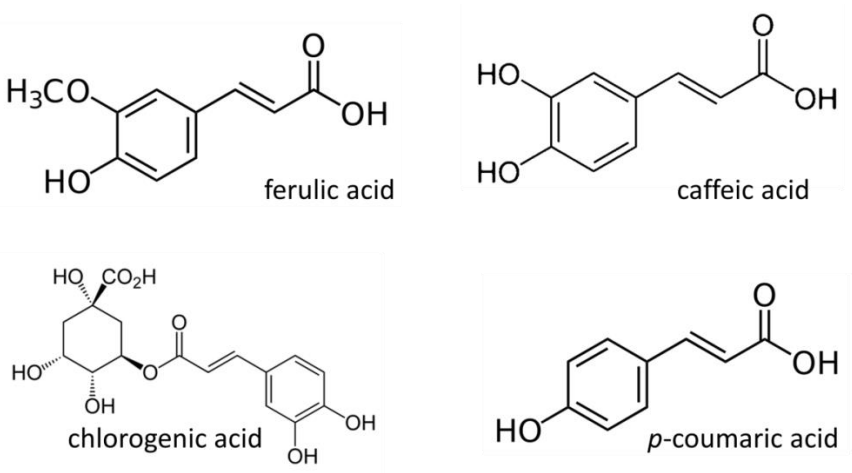

Figure 1-4. Structures of key hydroxycinnamic acids in poplar and chlorogenic acid.

Ferulic acid has been implicated in protection against herbivory. Following aphid infection of barley, ferulic acid increased along with other aromatic compounds (Cabrera et al., 1995). Furthermore, aphids fed an artificial diet with ferulic acid showed a decrease in survival with increasing concentrations of ferulic acid (Cabrera et al. 1995). A study by Abdel-Aal et al. (2001), showed a positive correlation between resistance to midge infection and concentration of ferulic acid in spring wheat. Midge larvae show decreased survival on wheat plants with higher levels of ferulic and *p*-coumaric acids (Ding et al., 2000). Additionally, midge infection induced the synthesis of ferulic and *p*-coumaric acid suggesting that hydroxycinnamic acid biosynthesis is an inducible defense (Ding et al., 2000).

Chlorogenic acids are esters of caffeic acid and quinic acid (Figure 1-4). Chlorogenic acid has been studied as a defensive chemical against insect pests and fungal infection. Studies in transgenic tobacco have shown that chlorogenic acids can protect against infection by the fungal pathogen *Cercospora nicotianae* (Maher et al., 1994; Shadle et al., 2003). Studies in willow have implicated chlorogenic acid as a deterrent and anti-feedant against beetles (Ikonen et al., 2001; Jassbi, 2003). A study by Leiss et al. (2009) demonstrated that chrysanthemums resistant to thrips had higher

concentrations of chlorogenic acid compared to susceptible plants. Furthermore, thrips fed an artificial diet with 5 % chlorogenic acid had reduced growth rate and reduced rates of survival compared to thrips fed an artificial diet lacking chlorogenic acid (Leiss et al., 2009). Together, these studies indicate a role of chlorogenic acid in plant defense against some insect pests and plant pathogens.

1.7 MYB transcription factors: structure and function

Regulation of gene expression occurs at multiple levels and primarily at the transcriptional level. Chromatin structure and transcription factors determine the rate of transcription. Chromatin structure is dynamic and changes during plant development. The structure of chromatin is important for the regulation of transcription as it can alter accessibility to genes by transcription factors. When chromatin is highly condensed transcription cannot occur and must be remodelled into a looser conformation by histone acetyltransferases and ATP-dependent chromatin remodeling enzymes.

Transcription factors regulate transcription by binding to short DNA sequences found in promoter regions and can either repress or induce the expression of the gene. Transcription factors have a modular structure consisting of a DNA-binding domain and an activation or repression domain. In order to regulate the expression of a gene, the transcription factor interacts with *cis* - regulatory DNA motifs and recruits RNA polymerase II. In many cases, interaction with the RNA polymerase II and the basal transcriptional complex requires interaction with additional co-factors.

MYB transcription factors are characterized by a highly conserved DNA binding domain known as the MYB domain consisting of up to four imperfect repeats (R) of about 52 amino acids (Dubos et al., 2010). Each repeat has a helix-helix-loop-helix secondary structure that binds to the major groove of DNA and contains three regularly spaced tryptophan residues that are necessary for binding (Saikumar et al., 1990). The MYB domain also carries the nuclear localization signal (Matus et al., 2008). The C-terminal region of the protein shows sequence divergence and is responsible for activation or repression of transcription. MYB transcription factors are further

characterised by the number of conserved repeats (R). In animals most MYB factors have three MYB domains (3R). In plants, one (R1) or two repeat (R2R3) MYB proteins are most common (Wilkins et al., 2009). The most highly expanded MYB family in plants is the R2R3 MYB transcription factor family. There are 192 putative R2R3 MYB transcription factors in the *P. trichocarpa* genome (Wilkins et al., 2009).

DNA cis-regulatory motifs have been characterized for R2R3 MYB transcription factors in plants and can be found in the plant cis-acting regulatory DNA elements (PLACE) database (http://www.dna.affrc.go.jp/PLACE/). There is significant diversity of cis-elements for MYB binding and a number have been characterized in promoter regions of flavonoid structural genes (Lai et al., 2013). Most commonly, AC-rich motifs are recognized by R2R3 MYB transcription factors (Lai et al., 2013). For example, the 'AC element' found in promoter regions of some flavonoid biosynthetic genes is recognized by regulators of condensed tannin synthesis in both poplar and persimmon (Akagi et al., 2010; Mellway et al., 2009).

A number of MYB factors in plants form a ternary complex with basic helix-loop-helix (bHLH) and WD repeating containing (WDR) type proteins. bHLH cofactors interact directly with DNA while WDR type proteins likely act to stabilize the protein-protein interaction between the MYB factor and the bHLH factor. This complex is plant specific and does not occur in animals. Baudry et al. (2004) showed that the Arabidopsis tannin regulator AtTT2 (AtMYB123) forms a complex with bHLH and WDR proteins to regulate gene expression. This study found that TT2 physically interacts with bHLH (AtTT8) and WDR (AtTTG1) proteins using a yeast three hybrid assay. Furthermore, a yeast one hybrid assay showed that TT2 along with the TT8 could bind to the ANR promoter. When expressed alone, neither AtTT2 nor AtTTG1 could bind to the flavonoid gene promoters. This suggests that the formation of a ternary complex is necessary for activation of the target genes by this type of MYB transcription factor.

1.8 The role of MYB transcription factors in regulating condensed tannin synthesis

The genes encoding regulators of condensed tannin synthesis in Arabidopsis were originally identified through analysis of genetic mutations in Arabidopsis seeds deficient in condensed tannins. These were designated as *transparent testa* (TT) mutations (Shirley et al., 1995). The MYB transcription factor, AtTT2 has been shown to regulate tannin synthesis and forms a regulatory complex with two other factors: AtTT8, a bHLH transcription factor, and AtTTG1, a WDR protein. This regulatory complex regulates the late flavonoid biosynthetic genes: DFR, ANS and ANR (Nesi et al. 2001).

Since the characterization of AtTT2, a number of other MYB regulators of condensed tannin synthesis have been identified in different species. These MYB factors cluster into two clades (Figure 1-5). Grapevine was the first plant in which two MYB regulatory factors of condensed tannin synthesis belonging to both clades were characterized, VvMYBPA1 and VvMYBPA2 (Bogs et al., 2007; Terrier et al., 2009). Because of this, the two condensed tannin clades are named the PA-1 and PA-2 clades. Most MYB regulators of condensed tannin synthesis that have been characterized cluster into the PA-2 clade and include AtTT2 and the poplar regulator, PtMYB134. To date, grape and persimmon are the only plants in which regulators in both clades have been characterized. VvMYBPA1 along with DkMYB4 cluster in the PA-1 clade (Figure 1-5). Since the PA-1 clade lacks an Arabidopsis homolog, this clade could represent a more specialized function in condensed tannin synthesis specific to perennial plants.

VvMYBPA1 is regulated by VvMYBPA2 and activates general flavonoid biosynthetic genes as well as genes specific to condensed tannin synthesis (Bogs et al., 2007; Terrier et al., 2009). Transient promoter activation assays indicate that VvMYBPA1 is able to induce the expression of PA specific branch point enzymes in Arabidopsis and grapevine: anthocyanidin reductase and leucoanthocyanidin reductase (Bogs et al., 2007). Additionally, the authors showed that regulation was specific to the condensed tannin branch of flavonoid synthesis, since the anthocyanin specific gene, UDP-Glc flavonoid glucosyltransferase, was not activated.

DkMYB2 and DkMYB4 have also been shown to specifically regulate condensed tannin synthesis (Akagi et al., 2009, 2010). However, they appear to differ in their function. In transient assays, DkMYB2, which belongs to the PA-2 clade, can activate the promoters of both tannin specific genes, LAR and ANR, while DkMYB4 activates only the ANR promoter (Akagi et al., 2010). Furthermore, DkMYB4 expression responds to changes in seasonal temperature while DkMYB2 does not (Akagi et al., 2011). This supports the idea that clade PA-1 represents an increase in complexity in the regulation of condensed tannin biosynthesis.

A putative regulator of condensed tannin synthesis in poplar, MYB115, also clusters in the PA-1 clade. MYB115 showed a 35.3 fold increase in expression in MYB134 overexpressing plants compared to wild-type plants, suggesting a potential role in condensed tannin synthesis. As MYB115 expression coincides with an increase in transcript abundance of flavonoid biosynthetic genes in MYB134 overexpressing plants, it is predicted to have a similar function to VvMYBPA1 in condensed tannin synthesis.

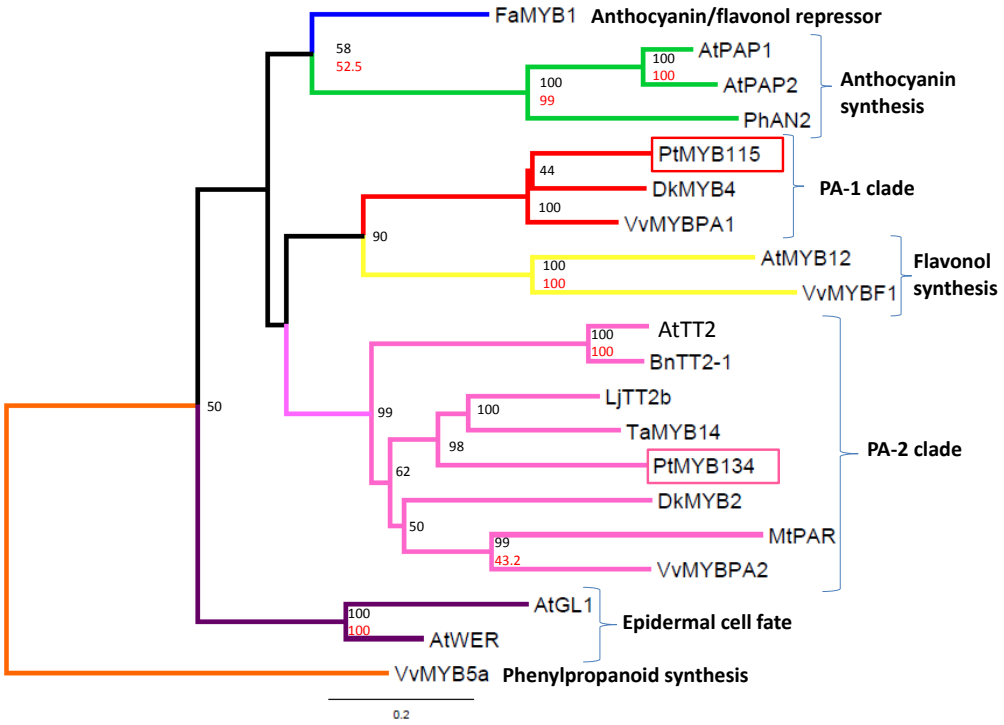

Figure 1-5. Phylogenetic tree representing functionally characterized R2R3 MYB transcription factors involved in flavonoid biosynthesis.

PA-1 and PA-2 clades include regulators of condensed tannin synthesis. Regulators of condensed tannin synthesis in poplar are boxed. MYB transcription factors involved in epidermal cell fate and a general phenylpropanoid pathway regulator were used as out-groups. Alignment was generated using ClustalW alignment software. Phylogeny construction is based on Neighbor-Joining and UPGMA tree drawing from a protein distance matrix using protein distance and neighbour software. Numbers at branch sites represent bootstrap values based on 100 bootstrap replicates using protein distance with neighbour joining (black) and maximum parsimony (red). Accession numbers from GenBank: *Fagaria ananassa* FaMYB1, AAK84064.1; *Arabidopsis thaliana* AtPAP1, AEE33419.1; *Arabidopsis thaliana* AtPAP2, AEE34503.1; *Petunia hybrida* PhAN2, AAF66727.1; *Populus trichocarpa* PtMYB115, EEE81917.1; *Diospyros kaki* DkMYB4, BAI49721.1; *Vitis vinifera* VvMYBPA1, CAJ90831.1; *Arabidopsis thaliana* AtMYB12, NP_182268.1; *Vitis vinifera* VvMYBF1, ACT88298.1; *Arabidopsis thaliana* AtTT2, AED93980.1; *Brassica napus* BnTT2-1, ABI13038.1; *Lotus japonicus* LjTT2b, BAG12894.2; *Trifolium arvense* TaMYB14, AFJ53053.1; *Populus trichocarpa* PtMYB134, EEE92051.1; *Diospyros kaki* DkMYB2, BAI49719.1; *Medicago trunculata* MtPAR, ADU78729.1; *Vitis vinifera* VvMYBPA2, ACK56131.1; *Arabidopsis thaliana* AtGL1, BAA86879.1; *Arabidopsis thaliana* AtWER, NP_196979.1; *Vitis vinifera* VvMYB5a, AAS68190.1.

1.9 Objectives and summary of key findings

The overall objective of this study is to functionally characterize the poplar R2R3 MYB transcription factor, MYB115, and to define its role in the biosynthesis of condensed tannins. MYB115 overexpressing plants were generated and showed an increased accumulation of tannins compared to wild-type plants and induction of flavonoid pathway genes (Chapter Two). Additional changes included decreases in concentrations of phenolic glycosides and changes to the concentrations of hydroxycinnamic acid derivatives. Since tannin synthesis is induced by stress, plants were wounded to test for stress inducibility of MYB115 (Chapter Two). MYB115, along with flavonoid biosynthetic genes and the previously studied regulator of condensed tannin synthesis, MYB134, was induced by wounding. Together, these experiments suggest that MYB115 is a stress-inducible regulator of the flavonoid pathway.

To test for activation of target genes, transient promoter activation experiments were performed (Chapter 3). These experiments demonstrated that MYB115 was able to activate the promoter of the key tannin biosynthetic gene, ANR1 but not LAR3. A second objective of this study was to test for transcriptional regulation of MYB115. Both MYB134 and MYB115 could activate the promoter of MYB115. This suggests that not only is MYB115 regulated by MYB134, but it is also self-regulated.

Together, these experiments demonstrate that MYB115 functions in the regulation of condensed tannin synthesis. MYB115 works in concert with MYB134 to regulate the expression of flavonoid biosynthetic genes required for the accumulation of condensed tannins. MYB115 and MYB134 expression also leads to secondary changes in phenolic metabolism including the reduction in concentrations of phenolic glycosides.

2. Chapter Two: Overexpression of MYB115 leads to induction of flavonoid pathway genes and enhanced accumulation of tannins and other changes to phenolic metabolism

In silico analysis (2.3.1) performed by Jane Guo (PhD candidate, University of Victoria)
All LC/MS analysis performed by Dr. Michael Reichelt (Max Planck Institute for Chemical Ecology in Jena, Germany)

2.1 Introduction

Condensed tannins are environmentally regulated plant compounds that contribute to plant defense. Condensed tannins accumulate at high levels in the leaves of some species of poplar. In trembling aspen, condensed tannins have been shown to accumulate in response to stresses including wounding and exposure to high-intensity light (Mellway et al., 2009; Peters and Constabel, 2002). Understanding the regulation of condensed tannin synthesis in poplar is important for understanding how plants respond to stress and could lead to the development of plants that have increased resistance to biotic and abiotic stresses.

In Arabidopsis, condensed tannins are restricted to the seed. In many other plants, in particular trees, tannins can accumulate in all major plant organs. Other plants also show increased complexity in condensed tannin accumulation due to the development of differential regulation in response to environmental cues. Studies have shown that condensed tannins accumulate in response to seasonal changes (Akagi et al., 2011, 2012; Feeny, 1970), wounding (Akagi et al., 2010; Arnold et al., 2008; Peters and Constabel, 2002), light stress (Mellway et al., 2009), nitrogen stress (Penuelas and Estiarte, 1997), and pathogen stress (Miranda et al., 2007; Wallis and Chen, 2012). Little is known on the signalling processes leading to changes in flavonoid content in response to stress; however, a number of hormones are known to mediate expression of flavonoid pathway genes including jasmonic acid and abscisic acid (Akagi et al., 2012; Peters and Constabel, 2002). Accumulation of condensed has been shown to be transcriptionally regulated and several transcription factors regulating condensed tannin

synthesis in response to environmental stimulus have been identified (Akagi 2010, 2011; Mellway et al., 2009).

Condensed tannin synthesis in persimmon and grape is regulated by at least two positive regulators (Akagi et al., 2011; Terrier et al., 2009). The two regulators fall into two distinct phylogenetic clades of R2R3 MYB factors. Hereafter, these two clades are named the PA1 and PA2 clades based on the grape regulators, VvMYBPA1 and VvMYBPA2 (Figure 1-5). It appears that poplar has a similar regulatory mechanism to persimmon and grape as it possesses MYB transcription factors in both PA-1 and PA-2 clades, called PtMYB115 and PtMYB134 respectively (Figure 1-5). MYB134, the ortholog of AtTT2, was characterized to be a key regulator of tannin synthesis (Mellway et al., 2009). MYB134 overexpressing plants show activation of all the flavonoid biosynthetic genes necessary for the synthesis of tannins. These plants also show a significant increase in the concentration of condensed tannins in leaves and other vegetative organs.

In addition to the induction of flavonoid biosynthetic genes, MYB134 overexpressing plants showed increased expression of many putative regulatory factors. PtMYB115, an R2R3 MYB transcription factor, showed a 35.3 fold change increase in MYB134 overexpressing plants (Mellway, 2009). PtMYB115 shows high sequence similarity to regulators of condensed tannin synthesis in grape, VvMYBPA1, and in persimmon, DkMYB4, and was therefore considered to be a good candidate for further analysis.

In this study, MYB115 was shown to be wound inducible along with MYB134 and flavonoid biosynthetic genes. In addition, transgenic poplar plants overexpressing PtMYB115 under the regulation of the constitutive CAMV 35S promoter were generated. The objective of this study was to use transgenic plants as a tool to determine the function of MYB115 in the regulation of the flavonoid pathway leading to condensed tannin synthesis. Transgenic plants showed strong activation of the flavonoid pathway and enhanced accumulation of condensed tannins. HPLC and LC/MS analysis showed additional changes to phenolic metabolism.

2.2 Methods

2.2.1 Wounding experiments

Populus tremula x *P. tremuloides* clone INRA 353-38 was used for wounding experiments and was originally provided by Steve Strauss (Oregon State University) and Richard Meilan (Purdue University) and is available in the Constabel lab. Plants were grown for three months in the Bev Glover Greenhouse. The leaf margins of leaves from Leaf Plastochron Index (LPI) 10-12 (LPI 1 being the first fully opened leaf with a mid-vein length of approximately 1.5 cm) were wounded by crushing with pliers and re-wounded one hour following initial treatment. Leaves from LPI 10-12 were chosen as they had reached maximum leaf size. Leaves were collected 24 hours after initial wounding. A 24 hour time point was chosen because of previously published time course data showing activation of PtMYB134 (Mellway et al., 2009). Necrotic tissue and mid-veins were removed from harvested leaves and leaf tissue was frozen in liquid nitrogen prior to further analysis. Plants not subjected to wounding were used as controls.

2.2.2 Generation of transgenic plants overexpressing PtMYB115

P. tremula x *P. alba* clone INRA 717-1-B4 was originally provided by David Ellis (CellFor, Vancouver, BC, Canada) and is available in the Constabel lab. Plants were micropropagated *in vitro* on Murashige and Skoog (MS) media (Murashige and Skoog, 1962) supplemented with 0.5 μM final concentration indole-3-butyric acid (IBA). The sequence of PtMYB115 was amplified with the primer set (MYB115-F: GAGTCATACCAGCAGTGACTC and MYB115-R: TCCTGGGAAGGGCTCCTTGTT) and cloned into pBI526 then sub-cloned into the binary plant transformation vector, pRD400, by Lan Tran (Datla et al., 1992; Wang and Constabel, 2004). Sequencing analysis confirmed insertion of the full coding sequence. The pRD400:PtMYB115 plasmid was moved into *Agrobacterium tumefaciens* strain C58 (pMP90) by electroporation.

Agrobacterium cells carrying pRD400:PtMYB115 were grown overnight with shaking at 11 rcf at 28 ^{o}C. Untransformed bacteria were used as a control. Cells were pelleted and resuspended in Induction Medium (MS media with vitamins and 1.28 mM

MES, 10 mM galactose, 50 µM acetosyringone) to an OD_{600} of 0.5. The *Agrobacterium* suspension continued to grow until the suspension reached an OD_{600} of approximately 0.6. Explants were then excised from 717-1-B4 and 353-38 *in vitro* plants (2 to 4 months old) and wounded with multiple cuts with a sterile scalpel across the leaf surface area. Leaves were incubated in the *Agrobacterium* suspension for one hour. Explants were then transferred to Callus Induction Media 1 (CIM1) plates (MS with 5 µM 2-isopentnyl adenine and 10 µM α-naphthalene-acetic acid) and incubated in the dark for two days. Explants were subsequently transferred to CIM2 plates (CIM1 with 250 mg/L cefotaxime, 500 mg/L carbenicillin and 50 mg/L kanamycin) and incubated for three weeks in darkness. The explants were then transferred to shoot induction media (CIM1 with 250 mg/L cefotaxime, 500 mg/L carbenicillin, 50 mg/L kanamycin, 0.2 µM thidiazuron) and grown under light conditions in growth chambers. Once shoots reached 0.5 to 1 cm in height, they were excised and transferred to root induction media (1/2 MS with 1.25 µM IBA). Plants were screened for a high tannin phenotype using a DMACA (*p*-dimethylaminocinnamaldehyde) stain (1% [w/v] in ethanol:6 N HCl, 1:1, [v/v]) (Feucht and Treutter, 1990; Mellway et al., 2009). Positive transformation was confirmed by PCR using primers for the 35S promoter and NOS2 terminator to amplify the transgene.

Positive transformants were micropropagated on solid MS media with 0.5 µM IBA. For HPLC, LC/MS, and gene expression analysis, the plants were grown in the Bev Glover greenhouse for two to four months prior to harvest. Plantlets were acclimated in Sunshine Mix #4 (Sungro, Seba Beach, AB, Canada) in a mist chamber for two weeks before being moved into the greenhouse and grown in Sunshine Mix #4 with fertilizer (21.4 g/gallon soil ACER® 21-7-14 (Plant Products Co. Ltd, Brampton, Ontario), 2.9 g/gallon soil Micromax Micronutrients (Scotts-Sierra, Marysville, OH, USA), 11.4 g/gallon soil dolomite lime (IMASCO, Surrey, BC, Canada), and 1.1 g/gallon soil Superphosphate 0-20-0 (Green Valley, Surrey, BC, Canada)). Plants were randomly placed in the greenhouse to avoid any environmental effects. Supplemental light was provided for a

day length of 16 hour light/8 hour dark and the temperature was maintained between 18 – 28 ^{o}C.

For biochemical and gene expression analysis, leaves from LPI 10-12 were harvested from at least three trees per line. Mid-veins were removed prior to freezing samples in liquid nitrogen. Samples were stored at -80 ^{o}C until analyzed.

2.2.3 Butanol HCL assay for tannin quantification

Freeze-dried leaf tissue was extracted in 100% methanol (see HPLC methods for extraction method). The butanol HCl assay (Porter et al., 1985) was performed using purified poplar tannins from MYB134 overexpressing transgenic plants as a standard. 500 µL of plant extract was added to the reaction mix (2 mL butanol HCL (95:5 v/v) and 66.8 µL iron reagent (2 % w/v $NH_4Fe(SO_4)_2$ in 2 N HCl). The reaction was incubated at 95 ^{o}C for 40 minutes. Following heating the reaction was allowed to cool for 20 minutes. Absorbance at 550 nm was read with a Perkin Elmer Victor™ X5 multilabel plate reader. Samples were corrected for anthocyanins and other pigments by subtracting absorbance readings of unheated controls.

All statistical tests were done in R (www.r-project.org). To perform analysis of variance, the aov function from the *stats* version 2.15.0 package was used. For correlation analysis, the *lm* function from the *stats* version 2.15.0 package was used.

2.2.4 High-performance liquid chromatography methods

25 mg of finely ground, freeze-dried tissue was extracted in 4.5 mL of 100 % HPLC grade methanol by sonication. Extracts were centrifuged for 10 minutes at 15 871 rcf to remove solid debris. Four mL of extracts was used for further analysis (0.5 mL was used for the butanol-HCl assay). Extracts were dried in a SC110A SpeedVac® Plus concentrator. Dried extracts were resuspended in 300 µL 100 % methanol and chlorophyll was removed using a Strata-X 33-µm solid-phase extraction columns (Phenomenex, Torrance, CA, USA). The Strata-X column was rinsed with 100% methanol followed by dH_20 before eluting sample. The sample was eluted in approximately 9 mL of 100% methanol into glass tubes and dried in a SC110A SpeedVac® Plus concentrator.

The dried extracts were resuspended in 100 % methanol to a final concentration of 10 mg/mL. 20 µL of sample was injected onto an HPLC system (Beckman Coulter System Gold 126 solvent module with a System Gold 168 diode array detector with a Phenomenex Kinetex C18 column [150 x 4.6, 2.6µm; 100 Å]). Separation was performed with an elution gradient with solvent A (dH_20 with 0.4 % formic acid) to solvent B (acetonitrile with 0.4 % formic acid) over 55 minutes at a flow rate of 1 mL min^{-1}. The gradient profile was 5 % B (0-5 min), 14 % B (6-11 min), 38 % B (12-40 min), 100 % B (41-47 min), 5 % B (48-49 min), and 0 % B (50-55 min).

Analysis was performed with 35 Karat Software Verson 5.0 (Beckman Coulter, Inc, Pasadena, USA). The baseline was manually added for integration of peak area. Compounds were identified by retention time and by comparing to UV spectra of representative standards. The identity of the compounds was further verified based on the fragmentation pattern from liquid chromatography/mass spectrometry (LC/MS) analysis (M. Reichelt and C.P. Constabel, unpublished data). Compounds were quantified using representative standards: phenolic glycosides were quantified as salicin equivalents, phenolic acids were quantified as chlorogenic acid equivalents, flavan-3-ols were quantified as catechin equivalents, and flavonoid glycosides were quantified as rutin equivalents. Compounds were quantified at 280 nm.

2.2.5 Liquid Chromatography/Mass Spectrometry methods

The LC/MS analysis of leaf extracts was performed by Dr. Michael Reichelt, Max Planck Institute for Chemical Ecology in Jena, Germany, using the following procedures. Chromatography was performed on an Agilent 1200 high performance liquid chromatography (HPLC) system (Agilent Technologies, Boeblingen, Germany). Separation was achieved on a Zorbax Eclipse XDB-C18 column (50 x 4.6 mm, 1.8 µm, Agilent, Waldbronn, Germany). 5 µL of extract reconstituted to 10 mg dry weight (DW) extract/mL methanol were injected for each sample. Separation was performed using the same gradient as described above for HPLC anlaysis. The column temperature was maintained at 25 °C. An API 3200 tandem mass spectrometer (Applied Biosystems,

Darmstadt, Germany) equipped with a Turbospray ion source was operated in negative ionization mode. The instrument parameters were optimized by infusion experiments with pure standards, where available. The ion-spray voltage was maintained at -4500 eV. The turbo gas temperature was set at 700 °C. Nebulizing gas was set at 60psi, curtain gas at 25 psi, heating gas at 60 psi and collision gas at 7 psi. Multiple reaction monitoring (MRM) was used to monitor product ions from precursor ions. Both Q1 and Q3 quadrupoles were maintained at unit resolution. Analyst 1.5 software (Applied Biosystems, Darmstadt, Germany) was used for data acquisition and processing. Quantification of compounds that were not detected by HPLC analysis were quantified as peak areas from LC/MS analysis.

2.2.6 RNA extraction and semi-quantitative RT-PCR and qPCR

RNA was extracted using a method as described by Muoki et al. (2012). RNA was extracted from approximately 50 mg of frozen and ground leaf tissue. Tissue was incubated for 15 minutes at 65 ^{o}C in pre-heated Buffer I (2 % cetyltrimethylammonium bromide (CTAB) (w/v), 2 % polyvinylpolypyrrolidone (PVPP) (w/v), 100 mM (hydroxymethyl)aminomethane [Tris–HCl (pH 8.0)], 125 mM ethylenediaminetetra acetic acid [EDTA(pH, 8.0)], 2 M sodium chloride, and 2 % β-mercaptoethanol). The lysate was mixed with chloroform:isoamyl alcohol (CIA) [24:1 (v:v)] and centrifuged at 15 871 rcf for 10 minutes. The supernatant was removed and mixed with CIA and centrifuged at 15 871 rcf for 10 minutes. The supernatant was removed and mixed with Buffer II (phenol saturated with Tris buffer to a pH of 8, sodium dodecyl sulfate (SDS) [0.1 % (w/v)], sodium acetate (NaOAc) [0.32 M (w/v)], and EDTA (0.01 M) (pH 8.0). Following the addition of chloroform, the samples were centrifuged at 15 871 rcf for 10 minutes at 4 ^{o}C. Isopropanol was added to the supernatant and the sample was incubated for 10 minutes at room temperature. The sample was centrifuged at 15 871 rcf for 10 minutes at 4 ^{o}C and supernatant was removed. The pellet was rinsed with 70 % ethanol. The pellet was resuspended in diethylpyrocarbonate (DEPC) treated dH_20.

Quality of the extracted RNA was assessed by denaturing gel electrophoresis. RNA was stored at -20 ^{o}C for short periods and -80 ^{o}C for longer periods of time.

For qPCR analysis, total RNA was treated with Amplification Grade DNase I (Invitrogen, Carlsbad, CA, USA) to remove genomic DNA contamination. cDNA was generated using SuperScript II (Invitrogen, Carlsbad, CA, USA) reverse transcriptase. Expression was normalized against expression of the poplar housekeeping gene, elongation factor 1-β (accession number: XM_002299613; Miranda et al., 2007) using the $2^{(-\Delta\Delta Ct)}$ method (Livak and Schmittgen, 2001). QuantiTect SYBR Green RT-PCR kit (Qiagen, Mississauga, Canada) was used for quantitative polymerase chain reaction (qPCR). Each reaction contained 2 µL 1:20 diluted cDNA template (5 ng), 1 µL of 10 µM forward and reverse primers (Table 2-1), 7.5 µL of 2X QuantiTect master mix (HotStarTaq DNA polymerase, deoxyribonucleotide triphosphate mix, SYBR Green I dye, ROX reference dye, and PCR buffer), and 4.5 µL of water for a final reaction volume of 15 µL. For each reaction, two identical technical replicates were analyzed. No-reverse transcriptase controls were included for each sample. qPCR was performed on an Mx3005p QPCR System (Stratagene, Stratagene, La Jolla, CA, USA). The PCR conditions were 10 minutes at 95 ^{o}C followed by 40 cycles of 95 ^{o}C for 30 seconds, 56 to 60 ^{o}C for 30 seconds (see Table 2-1 for annealing temperatures), 72 ^{o}C for 30 seconds followed by one cycle of 95 ^{o}C for one minute, 56 to 60 ^{o}C for 30 seconds and 95 ^{o}C for 30 seconds. For each sample run, a melt curve analysis was performed using the Mx3005p default parameters (60 seconds at 95 °C, 30 seconds at 55-95 °C in one degree increments, 30 seconds at 95 °C), which yielded one peak after normalization with the ROX signal for each set of primers. Annealing temperatures were optimized for a high primer efficiency. Primer efficiency was estimated using by calculating the slope of a dilution series of template concentration. Efficiency was calculated using the slope in the following equation: primer efficiency % = $((10^{-(1/slope)})- 1)100$.

Table 2-1. List of primers used for qPCR analysis.
Expression of all genes was normalized against the housekeeping gene, elongation factor 1-β (EF 1β).

Gene Name	Forward primer	Reverse primer	Amplicon length (bp)	Annealing temperature (°C)
EF 1β	AAGAGGACAAGAAGGCAGCA	CTAACCGCCTTCTCCAACAC	145	58
MYB115	GGATTGTGATAATGGGGTTGCC	GTGACTCGGCGAAGGAGTTT	185	58
MYB134	GGACACTGGAATGAGTTTCAA	ATGTGCCAAAGATTCCAAGTC	184	60
MYB182	GAATCTTTGGTGACACAGCAAGC	GAAGCAGAGTTGGCAATGATGA	181	58
F3'5'H1	GCAACGGCTCATGAACGCAAGG	ATGCTCGAGGAAGTGTCAGTGC	152	60
DFR2	CTTATAACTGCCCTTTCTCTGA	AGATCATGAATGGTGGCTT	173	58
ANR1	CCATCACTTCAGAGAAGCTCAT	ACACCAGATACAGCCAAGCTAG	191	56

2.2.7 Microarray analysis

RNA extraction was performed as described (2.2.6). RNA was purified with NucleoSpin RNA II clean-up kit (Clontech, Mountain View, CA, USA). Affymetrix GeneChip® Poplar Genome Array microarray hybridizations were conducted at the Genome Quebec Innovation Centre at McGill University (Montreal, QC, Canada). Data was normalized with FlexArray (genomequebec.mcgill.ca/FlexArray) using a Robust Multi-array Average (RMA) algorithm. To identify differentially expressed genes, Empirical Bayes (Wright and Simon) algorithm was performed in FlexArray. Only genes with a *p*- value ≤ 0.05 and a fold change > 2 or < 0.5 were defined to be up- or down-regulated.

Annotations were obtained from annotation files provided by Affymetrix. Gene model IDs were obtained from POParray (Tsai et al., 2011) and further annotations were obtained from Blast2Go analysis (Conesa et al., 2005). Annotations were further verified by BLASTn searches in the NCBI Transcript Reference Sequences database (Benson et al., 2005) and keyword searches in Phytozome v9.1 *Populus trichocarpa* JGI assembly release v3.0, annotation v3.0 (Goodstein et al., 2012) databases. If there was more than one probe for a given gene, the probeset that showed a smaller *p*-value was presented.

2.3 Results

2.3.1 In silico *co-expression analysis of MYB115*

A co-expression analysis of MYB115 against a developmental series of 365 public, poplar Affymetrix microarray datasets was performed by Jane Guo (PhD candidate, University of Victoria). MYB115 expression correlated strongly with the expression of flavonoid biosynthetic enzymes including condensed tannin specific enzymes, ANR and LAR (Table 2-2). This is consistent with the hypothesis that MYB115 is a regulator of condensed tannin synthesis. Expression also correlated with two MATE efflux family proteins which potentially function in transport of tannin subunits into the vacuole (Marinova et al., 2007). Expression also correlated with regulatory proteins including one MYB family protein (not yet characterized) and two WD repeat type proteins that show close homology to the Arabidopsis light responsive regulators, AtLWD1 (Wu et al., 2008). Both AtLWD1 homologs are induced in MYB134 and MYB115 overexpressing plants (Mellway, 2009; Table 2-3). These regulatory factors could function along with MYB115 to regulate expression of flavonoid pathway genes. A glutathione transferase like protein is also co-expressed with MYB115 that could potentially be involved in modification of flavonoids for transport into the vacuole (Mueller et al., 2000).

Table 2-2. Co-expression analysis of MYB115.
Expression analysis was completed using a developmental series dataset compiled by Jane Guo a PhD candidate in Dr. Juergen Ehlting's lab. The dataset represents data from 365 Poplar Affymetrix arrays. Data was normalized with R bioconductor using Affymetrix's MAS 5.0 expression measure. Colours indicate genes sharing a common function. Yellow, flavonoid biosynthesis; red, putative function in phenylpropanoid biosynthesis; blue, putative flavonoid transporter; green, regulatory proteins.

r-value to bait	GeneChip-ID	Annotation
1	BAIT PtpAffx.30659.1.A1_at	MYB115
0.83	Ptp.8030.1.S1_at	cytochrome b5
0.801	Ptp.8030.1.S1_s_at	cytochrome b5
0.772	PtpAffx.7896.3.S1_a_at	chalcone synthase (CHS3)
0.766	Ptp.323.1.S1_s_at	flavanone 3-hydroxylase (F3H6)
0.762	PtpAffx.25553.1.A1_at	dihydroflavonol 4-reductase(DFR2)
0.754	Ptp.6057.1.S1_at	anthocyanidin synthase (ANS2)
0.752	PtpAffx.5092.2.S1_a_at	anthocyanidin reductase (ANR1)
0.738	PtpAffx.5092.1.A1_at	anthocyanidin reductase (ANR1)
0.733	PtpAffx.7896.4.A1_a_at	chalcone synthase (CHS6)
0.727	PtpAffx.224485.1.S1_s_at	(MATE) family transporter-related (AtTT12 like)
0.718	PtpAffx.161181.1.S1_at	cinnamoyl-CoA reductase-like
0.711	Ptp.6753.1.S1_s_at	putative protein -dihydrofolate reductase
0.708	PtpAffx.204062.1.S1_at	cinnamoyl CoA reductase-like
0.703	Ptp.1512.1.S1_s_at	chalcone isomerase-like (CHIL2)
0.694	PtpAffx.18705.2.A1_a_at	leucoanthocyanidin reductase (LAR3)
0.689	PtpAffx.94822.1.A1_at	(MATE) family transporter-related (AtTT12 like)
0.678	PtpAffx.212699.1.S1_at	MYB203
0.675	PtpAffx.213439.1.S1_at	WD40 repeat-containing protein (AtLWD1 like)
0.673	Ptp.4458.1.S1_s_at	glutathione transferase-like
0.665	PtpAffx.6065.2.S1_at	leucoanthocyanidin reductase (LAR1)
0.664	Ptp.5716.1.S1_at	monosaccharide transporter-like
0.651	PtpAffx.37082.1.A1_at	dihydroflavonol 4-reductase (DFR1)
0.628	PtpAffx.127289.1.A1_at	WD40 repeat-containing protein (AtLWD1 like)
0.628	PtpAffx.142603.1.A1_s_at	flavonoid 3'-hydroxylase (F3'H1)

2.3.2 MYB115 expression is induced by wounding and coincides with induction of flavonoid biosynthetic genes

PtMYB134 has been previously shown to be a stress-responsive regulator of condensed tannin biosynthesis. PtMYB134 is induced by high-intensity light, UV-B light and wounding (Mellway et al., 2009). To determine if MYB115 is stress inducible, wounded 353-38 plants were analyzed for induced expression of MYB115 along with MYB134 and flavonoid biosynthetic genes. Leaf margins of 353-38 plants were crushed with pliers 24 hours prior to harvest. MYB115 was expressed on average 4.5 fold higher in wounded plants compared to unwounded plants (Figure 2-1). The experiment was repeated once and MYB115 again showed higher expression (4.1 fold) in the wounded plant compared to the unwounded plant (n = 1; Appendix A, Figure A-1). MYB134 showed an average 3.3 fold change compared to control plants.

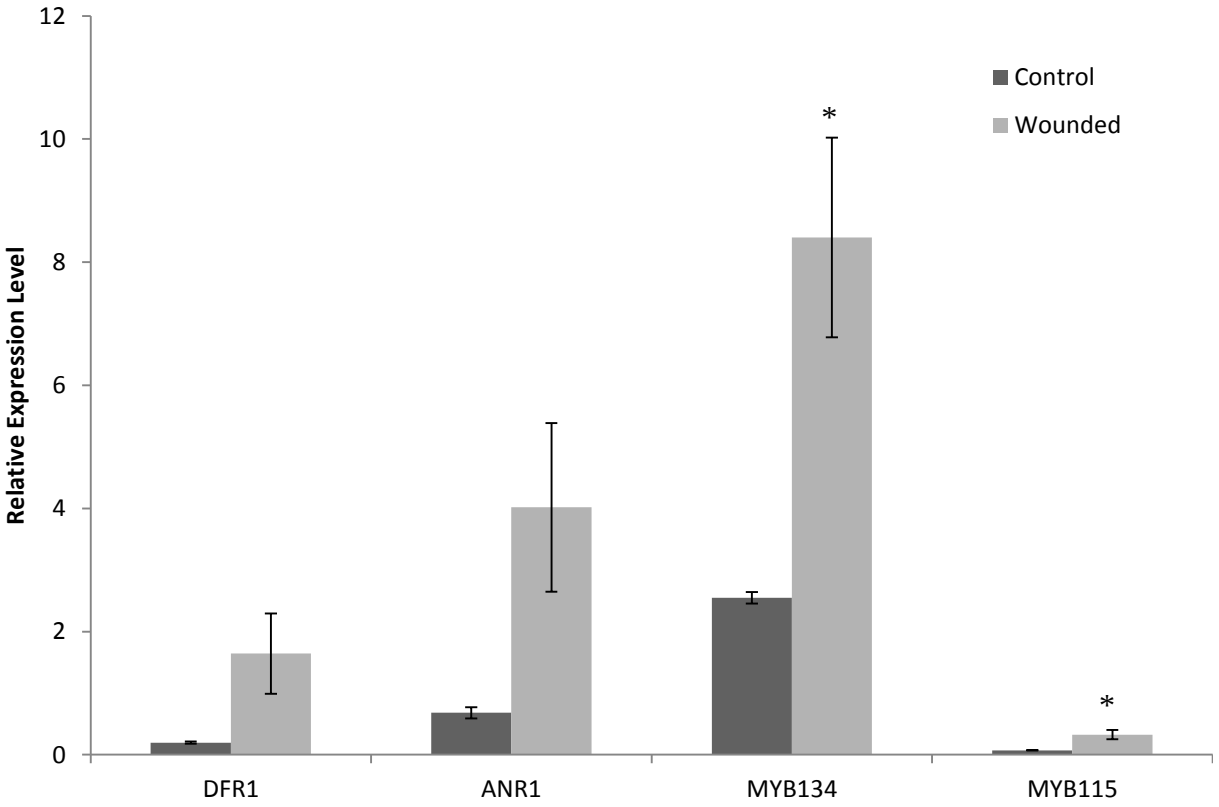

Figure 2-1. Expression analysis of 353-38 hybrid aspen wounded and control plants.
Leaf margins were wounded by crushing with pliers 24 hours prior to harvest. LPI 10 - 12 were collected from each plant and pooled for analysis. Error bars indicate standard error (n = 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *).

2.3.3 Generation of transgenic plants overexpressing MYB115

Both INRA clones 353-38 (*P. tremula* x *P. tremuloides*) and 717-1-B4 (*P. tremula* x *P. alba*) were transformed with an overexpression construct carrying the gene coding for MYB115 under the control of the double CaMV 35S promoter with the alfalfa mosaic virus RNA4 transcriptional enhancer sequence (Datla et al., 1992). For each poplar hybrid, three sets of *Agrobacterium* infections were performed. For each infection, approximately 45 leaf explants were infected with *Agrobacterium* carrying the MYB115 overexpression construct for a total of approximately 135 infected explants for each poplar hybrid. After explants were transferred to shoot induction media, only one set of infections per poplar hybrid resulted in successful shoot development and transformed plants.

Explants were placed on kanamycin media for selection of positive transformants. 30-35 shoots per poplar hybrid that successfully developed on kanamycin-containing media were transferred to rooting media. It is important to note that the number of shoots does not reflect the number of explants infected since explants were often cut into more than one section following infection. For example, if calli developed on two sides of the explant the explant would be cut in half in order to separate the two calli. To select for a high tannin phenotype, the shoots were stained with DMACA. Five 353-38 and three 717-1-B4 lines stained dark blue indicating accumulation of condensed tannins. Each plant represented a single transformation event. Successful transformation was confirmed by amplifying the transgene from genomic DNA.

These individual plantlets were then micropropagated on MS media to generate enough plants (> 12) to transfer to the greenhouse and to maintain a stock in tissue culture. With the goal of having three to four replicates for each line, eight plants were transferred to the greenhouse since the survival rate was unpredictable. Plants were grown for two to four months prior to harvest for gene expression and chemical analyses. Expression of MYB115 was measured in all lines by qPCR. The overexpressed transgene, MYB115, showed significantly higher expression in all transgenic lines

compared with WT plants with the exception of one 717-1-B4 transgenic line (Line 2) (Figure 2-2). These results further verified that the plants were successfully transformed and are expressing the transgene at high levels.

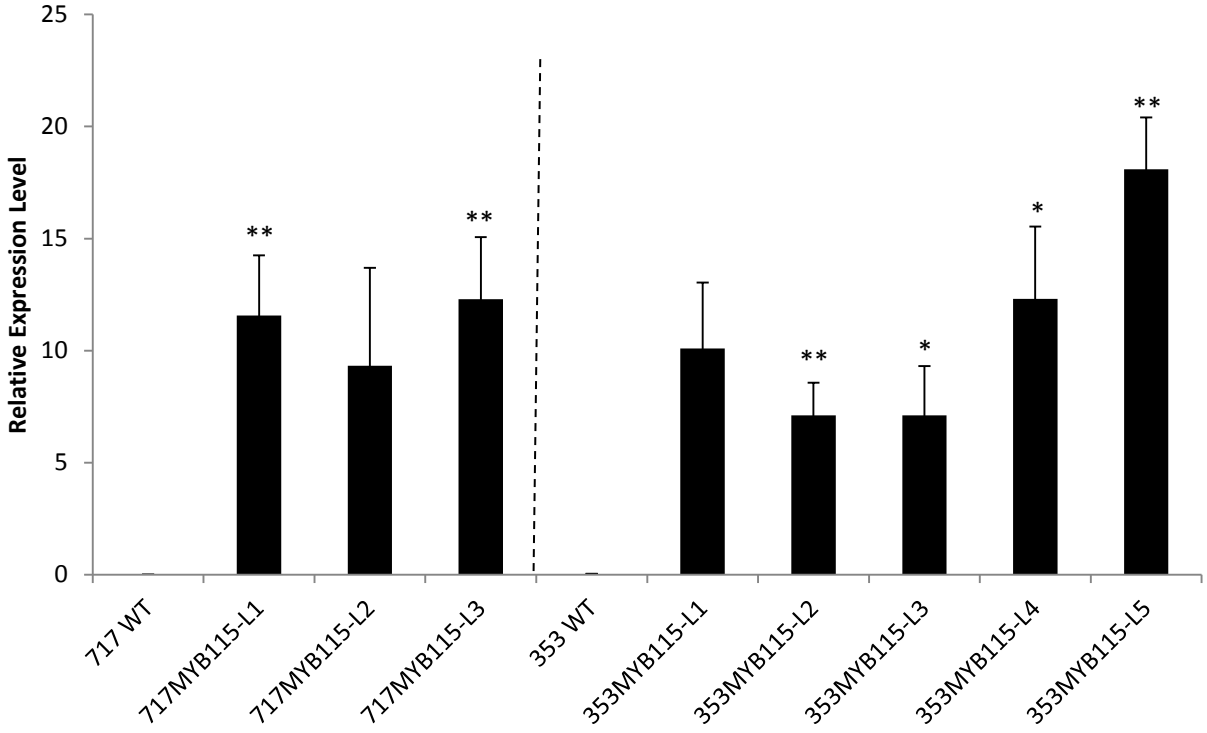

Figure 2-2. Relative expression of MYB115 in MYB115 overexpressing poplar as analyzed by qPCR.

Data shown include all positive transformants for both the 717 (*P. tremula* x *P. alba* [INRA 717-1-B4]) and 353 (*P. tremula* x *P. tremuloides* [INRA 353-38]) backgrounds. Lines with different genetic backgrounds are separated with a dashed line. Expression is normalized against elongation factor 1-β, a housekeeping gene. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants; 353MYB115-L1: n = 2 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***).

2.3.4 Transgenic plants overexpressing MYB115 show enhanced accumulation of tannins and the tannin pre-cursor, catechin

To test if MYB115 overexpression results in enhanced accumulation of condensed tannin synthesis as predicted, condensed tannin concentrations were measured using the butanol HCl assay (Porter et al., 1985). All 717-1-B4 MYB115 overexpressing transgenic lines showed significant increases (ANOVA, $p \leq 0.05$) in tannin levels compared to wild-type (WT) plants (Figure 2-3). Two of the five 353-38 MYB115 overexpressing lines (Lines 4 and 5) showed very strong and significant increases in condensed tannin compared to wild-type (WT) plants (Figure 2-3).

Tannins are polymers of flavan-3-ol subunits. The most common subunits in poplar are catechin, epicatechin, epigallocatechin and gallocatechin (Schweitzer et al., 2008). Flavan-3-ol monomers differ based on the hydroxylation pattern of the B-ring and their *cis*- or *trans*- configuration. Catechin was quantified by HPLC and was identified based on its UV spectrum and retention time compared to a catechin standard. An example HPLC trace is shown in Appendix A, Figure A-2. Catechin showed enhanced accumulation in all 717-1-B4 MYB115 overexpressing lines and the two high tannin 353-38 MYB115 overexpressing lines (Lines 4 and 5) (Figure 2-4). Catechin was not detectable in WT plants. In 353-38 and 717-1-B4 plants, as expected there was a strong correlation between catechin and condensed tannins (Appenix A, Table A-3). Epicatechin, a stereoisomer of catechin, was below the detection limit of HPLC in both MYB115 overexpressing plants and WT plants. As both condensed tannins and the tannin precursor, catechin, showed enhanced accumulation in MYB115 overexpressing plants compared to WT plants, MYB115 is likely a positive regulator of condensed tannin synthesis.

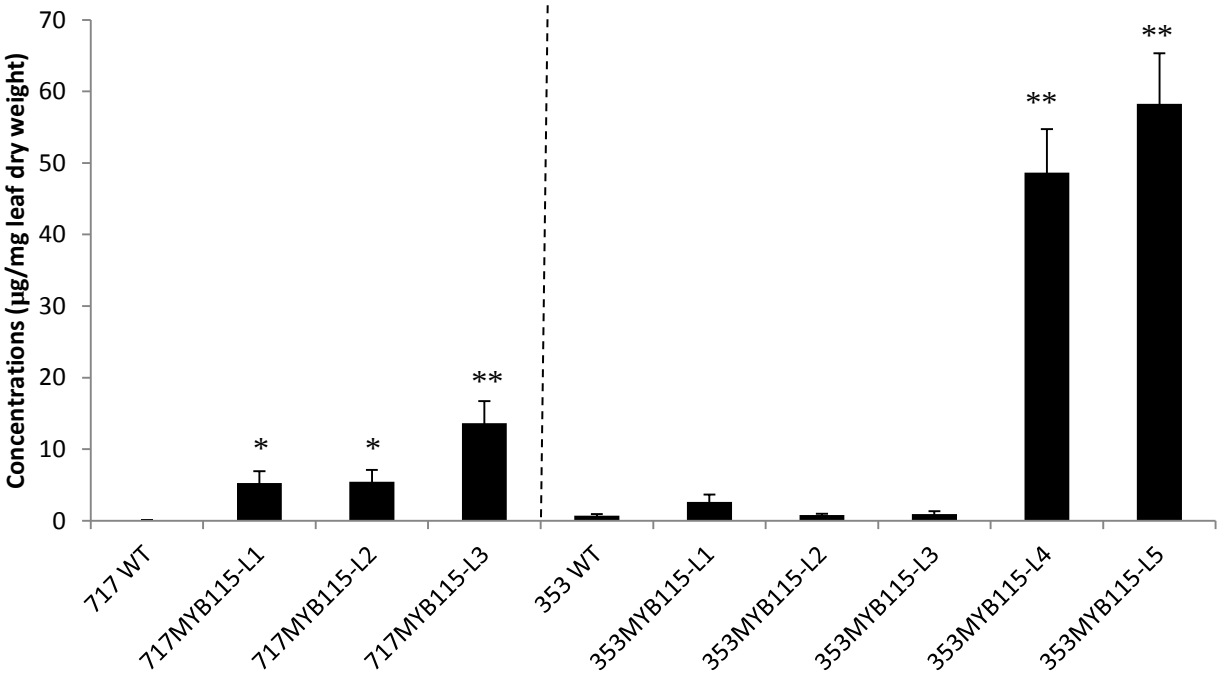

Figure 2-3. Concentration of condensed tannins in wild-type (WT) and MYB115 overexpressing poplar as analyzed by the butanol HCl assay.
Data shown include all positive transformants for both the 717 (*P. tremula* x *P. alba* [INRA 717-1-B4]) and 353 (*P. tremula* x *P. tremuloides* [INRA 353-38]) backgrounds. Lines with different genetic backgrounds are separated with a dashed line. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **).

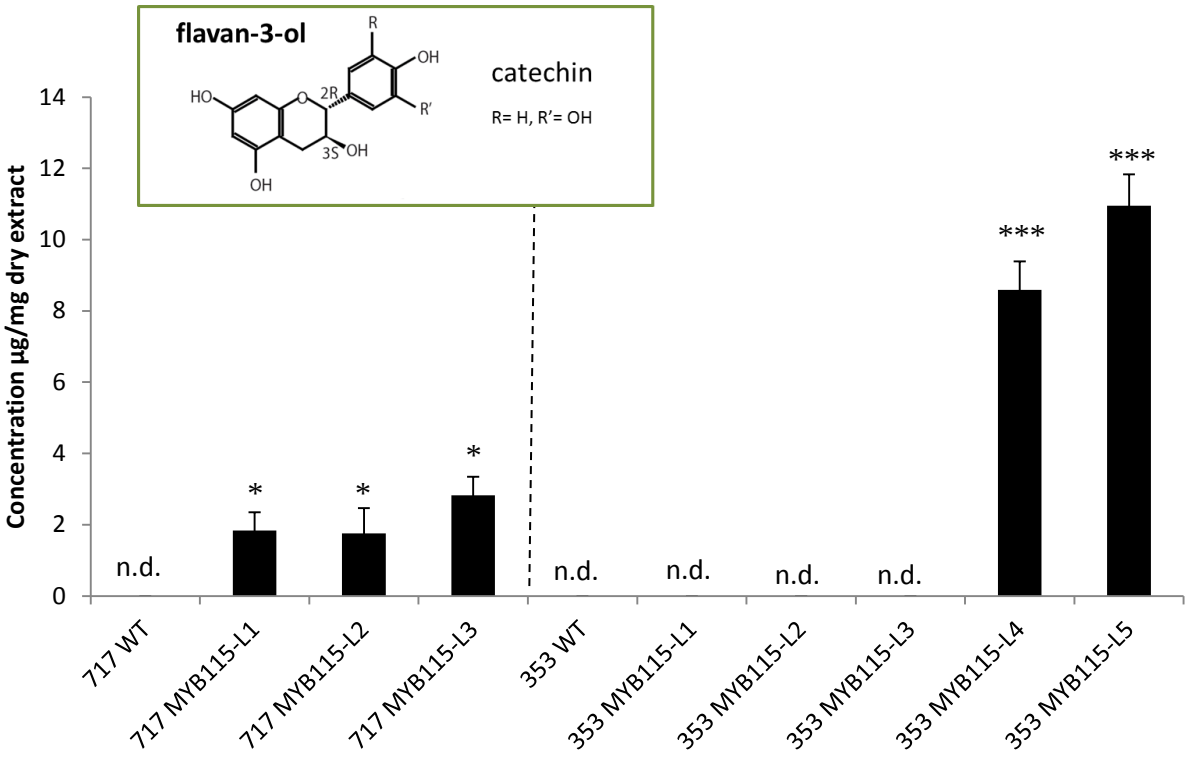

Figure 2-4. Concentration of catechin in wild-type (WT) and MYB115 overexpressing poplar as quantified by HPLC in both 353-38 and 717-1-B4 backgrounds.
Purified catechin was used as a standard. Each line (L1, L2 etc.) represents an independent transformation event. Lines with different genetic backgrounds are separated with a dashed line. Samples labelled n.d. indicate that the concentration of catechin was below the detection limit of the HPLC. Error bars indicate standard error (n ≥ 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). Also shown is the location of catechin in the flavonoid pathway (refer to Figure 1-2).

2.3.5 MYB115 overexpressing plants show enhanced accumulation of dihydroflavonols

HPLC analysis showed an unknown peak in the two high tannin 353-38 MYB115 overexpressing lines (Lines 4 and 5) that was absent in the low tannin transgenic lines and WT plants (Figure 2-5; Appendix A, Figure A-2). This peak was also not observed in any of the 717-1-B4 plants. This peak was tentatively characterized to be dihydromyricetin based on the fragmentation pattern observed in LC/MS analysis (M. Reichelt and C.P. Constabel, unpublished data). Dihydromyricetin, a dihydroflavanol, is synthesized by hydroxylation of dihydrokaemferol by a F3'5'H (Figure 1-2). Dihydromyricetin is a precursor of the trihydroxylated flavan-3-ols epigallocatechin and gallocatechin, which are both direct precursors of condensed tannins. There was a strong correlation between the concentration of dihydromyricetin and condensed tannins (Appendix A, Table A-3). Another dihydroflavanol, tentatively identified as dihydroquercetin by LC/MS, showed significantly enhanced accumulation (ANOVA, $p \leq 0.05$) in both 717-1-B4 and 353-38 plants as analyzed by LC/MS (Figure 2-6). This compound was not detected by HPLC analysis and is therefore quantified only as peak areas. Dihydroquercetin is a precursor of the dihydroxylated flavan-3-ols, catechin and epicatechin. The hydroxylation pattern of dihydroflavonols influences the subunit composition of condensed tannins. These data suggest that MYB115 is regulating the synthesis of dihydroflavonols and therefore it is likely that MYB115 is also influencing the subunit composition of condensed tannin as well. However, this was not tested here.

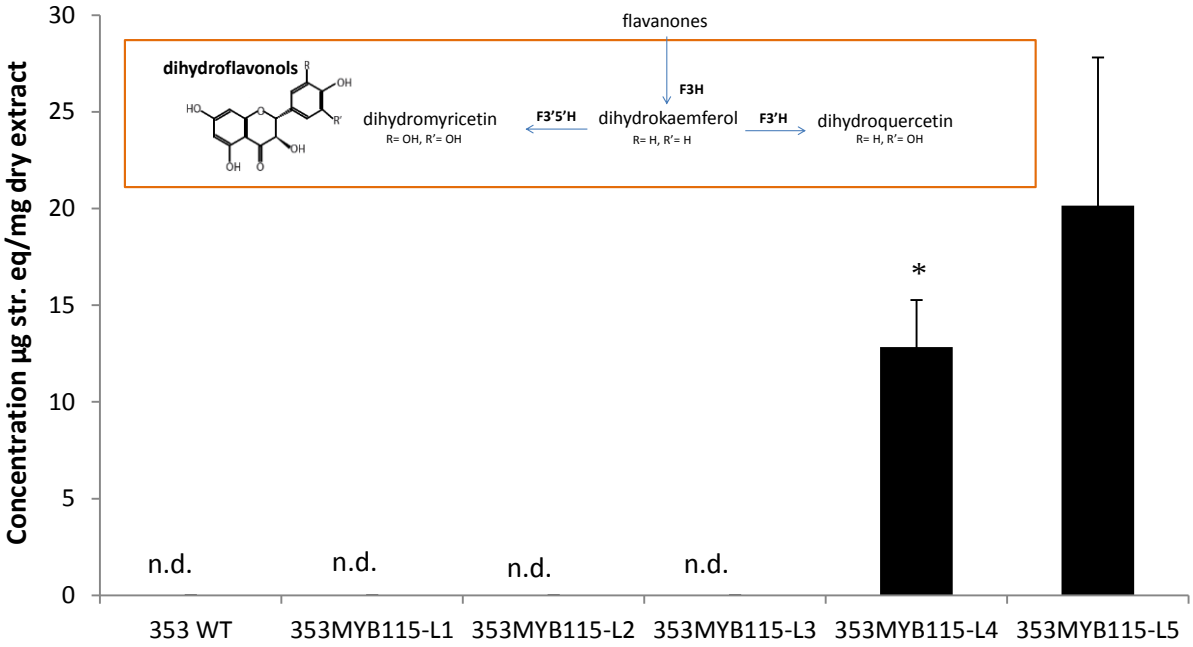

Figure 2-5. Concentration of dihydromyricetin in 353-38 wild-type (WT) and 353-38 MYB115 overexpressing poplar as analyzed by HPLC.

Concentration was calculated as a rutin standard equivalent. Each line (L1, L2 etc.) represents an independent transformation event. Samples labelled n.d. indicate that the concentration of dihydromyricetin was below the detection limit of the HPLC. Error bars indicate standard error (n = 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). Also shown is the location of dihydromyricetin in the flavonoid pathway (refer to Figure 1-2).

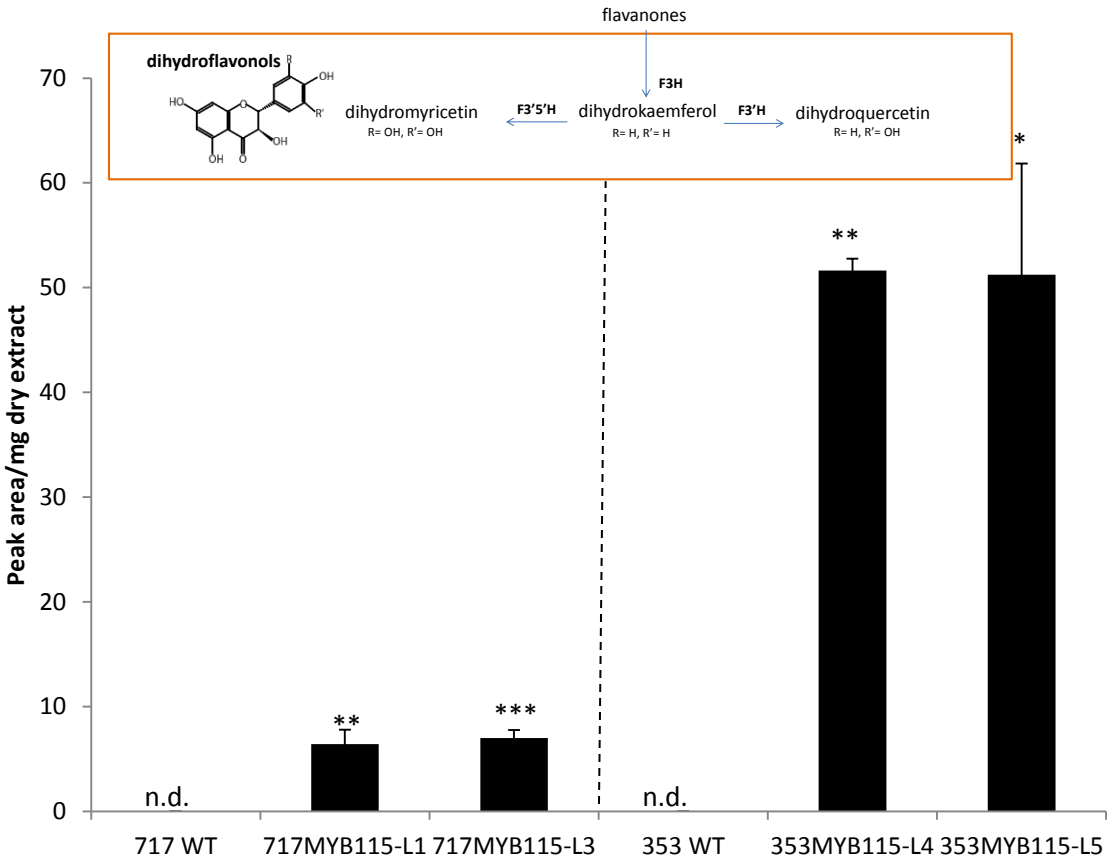

Figure 2-6. Concentration of dihydroquercetin in wild-type (WT) and MYB115 overexpressing poplar as analyzed by LC/MS in both 717-1-B4 and 353-38 backgrounds.
Each line (L1, L2 etc.) represents an independent transformation event. Lines with different genetic backgrounds are separated with a dashed line. Samples labelled n.d. indicate that the concentration of dihydroquercetin was below the detection limit of the LC/MS. Error bars indicate standard error (n = 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). Also shown is the location of dihydroquercetin in the flavonoid pathway (refer to Figure 1-2). This compound could only be detected by LC/MS and therefore only relative peak area quantification was possible.

2.3.6 LC/MS identification of flavan-3-ols and dimeric proanthocyanidins in MYB115 overexpressing plants

LC/MS analysis of leaf tissue from 717-1-B4 and 353-38 MYB115 overexpressing plants verified the enhanced accumulation of the flavan-3-ol, catechin (Figure 2-7, 2-8). This LC/MS analysis did not distinguish between *cis*- and *trans*- isomers; however, it likely is measuring catechin levels as epicatechin was below the detection limit of HPLC. Gallocatechin was not detected by HPLC but was tentatively identified by LC/MS and shown to be present at higher levels in MYB115 overexpressing plants compared to WT plants in both 717-1-B4 and 353-38 lines (Figure 2-7, 2-8).

Compounds that were tentatively identified as dimeric proanthocyanidins based on their fragmentation pattern could be detected by LC/MS (M. Reichelt and C.P. Constabel, unpublished data). Dimeric proanthocyanidins showed significantly enhanced accumulation in MYB115 overexpressing plants compared to WT plants. Catechin-catechin dimers were significantly enhanced (ANOVA, $p \leq 0.05$) in both 717-1-B4 and 353-38 MYB115 overexpressing lines as analyzed by LC/MS (Figure 2-7, 2-8). Gallocatechin-catechin dimers were only detected in 353-38 transgenic plants, similarly to dihydromyricetin, and were significantly higher (ANOVA, $p \leq 0.05$) in the MYB115 overexpressing plants compared to WT plants (Figure 2-8). The strong increase in gallocatechin and gallocatechin-catechin dimers in 353-38 transgenic plants is likely a result of the increase in dihydromyricetin which is a precursor to gallocatechin.

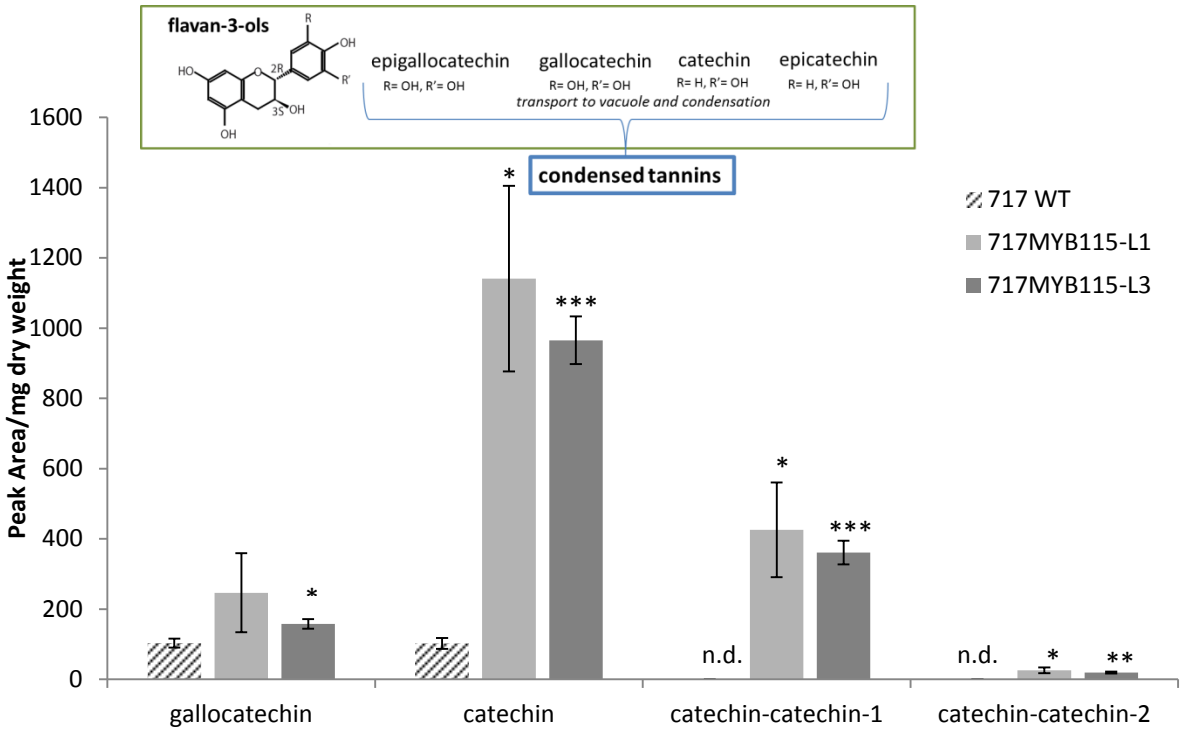

Figure 2-7. Levels of flavan-3-ols and flavan-3-ol dimers in 717-1-B4 wild-type (WT) and 717-1-B4 MYB115 overexpressing poplar as analyzed by mass spectrometry.
Each line (L1 and L2) represents an independent transformation event. Samples labelled n.d. indicate that the concentration was below the detection limit of LC/MS. Error bars indicate standard error (n ≥ 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). This analysis did not differentiate between *cis*- and *trans*- isomers. This compound could only be detected by LC/MS and therefore only relative peak area quantification was possible.

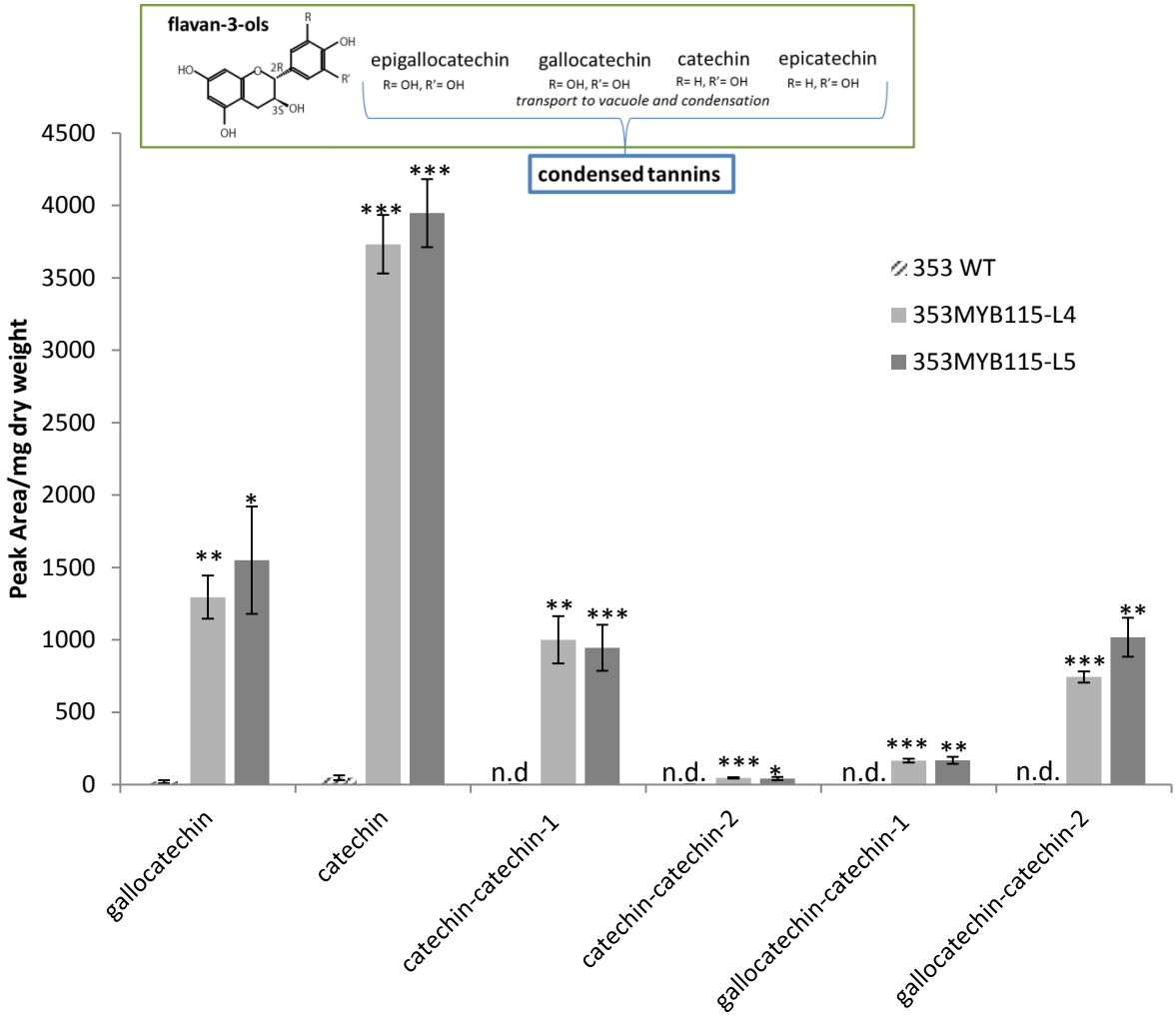

Figure 2-8. Levels of flavan-3-ols and flavan-3-ol dimers in 353-38 wild-type (WT) and 353-38 MYB115 overexpressing poplar as analyzed by mass spectrometry.
Each line (L1 and L2 represents an independent transformation event. Samples labelled n.d. indicate that the concentration was below the detection limit of LC/MS. Error bars indicate standard error (n = 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). This analysis did not differentiate between *cis*- and *trans*- isomers. This compound could only be detected by LC/MS and therefore only relative peak area quantification was possible.

2.3.7 Expression analysis of transgenic plants overexpressing MYB115 compared with wild-type plants

Overall changes in gene expression as a result of MYB115 overexpression were analyzed in 353-38 MYB115 overexpressing transgenic plants (Line 5) compared to 353-38 WT plants using the Affymetrix GeneChip® Poplar Genome Array. RNA was extracted from leaves LPI 10-12 of three transgenic plants and three WT plants for analysis. The same tissue was used for the biochemical analysis. Data was normalized and annotated as described in the methods section. After analysis, a total of 208 probes showed increased transcript abundance with greater than two fold change in the MYB115 overexpressing plants compared to WT plants. These 208 probes represent 139 unique gene models. Seventeen probes had no gene model associated with them. A total of 284 probes showed < 0.5 fold change. These 284 probes represent 213 unique gene models. Thirteen probes had no gene model ID associated with them.

Deregulated genes were annotated and categorized based on function. A subset of deregulated genes, were chosen based on their putative role in phenylpropanoid biosynthesis and are presented in Table 2-3. These genes were annotated and verified using NCBI BLAST and Phytozome Keyword Search as well as from Gene Ontology IDs (see section 2.2.6 for details). A complete list of all deregulated genes is included in the appendix (Appendix A, Tables A-1, A-2). A number of transcriptional regulators were deregulated in MYB115 overexpressing plants, including the previously characterized regulator of condensed tannin synthesis, PtMYB134. Putative negative regulators of condensed tannin synthesis, PtMYB182, PtMYB165 and PtMYB194, were also upregulated by MYB115 overexpression. Putative WDR and bHLH co-factors of MYB transcription factors were also upregulated in MYB115 overexpressing plants.

The most highly upregulated gene in MYB115 overexpressing plants was a flavonoid 3'5'-hydroxylase (F3’5’H1). This enzyme is required for trihydroxylation of the B-ring of flavonoids and may function in the biosynthesis of dihydromyricetin. In addition to F3’5’H1, all genes encoding for the flavonoid biosynthetic enzymes

necessary for condensed tannin synthesis were upregulated in MYB115 overexpressing plants. A MATE transporter protein was upregulated that showed high sequence homology to the characterized Arabidopsis transporter involved in condensed tannin biosynthesis. Five genes belonging to the cytochrome p450 family CYP92A that were annotated as flavonoid hydroxylase-like were downregulated in MYB115 overexpressing plants.

Genes involved in the general phenylpropanoid pathway showed higher expression in MYB115 overexpressing plants including 4CL and C4H. Three transcripts annotated as cinnamoyl-CoA reductase like proteins were activated that are also potentially involved in the general phenylpropanoid pathway. Three transcripts that may function in lignin biosynthesis were downregulated in MYB115 overexpressing plants, two peroxidases and a cinnamyl alcohol dehydrogenase-like protein. One transcript encoding for an aldehyde dehydrogenase like protein possibly involved in HCA biosynthesis was downregulated. Two transcripts encoding for BAHD acyltransferase like proteins which might be involved in phenolic glycoside biosynthesis were also downregulated.

The results of this microarray analysis demonstrate that MYB115 likely has a key role in the regulation of the flavonoid pathway leading to tannin synthesis. All genes necessary for condensed tannin synthesis are represented in this analysis, including the tannin specific genes ANR1 and LAR1/2/3. Furthermore, putative genes involved in flavonoid modification and transport that may also be involved in condensed tannin synthesis are upregulated.

Table 2-3. Deregulated genes related to phenylpropanoid and flavonoid biosynthesis in MYB115 overexpressing transgenic poplar (353-38 MYB115 Line 5, n = 3) compared to wild-type plants (n = 3 plants) as analyzed by Affymetrix GeneChip® Poplar Genome Array.
Table is ranked by decreasing fold change. Green indicates an increase in gene expression greater than two fold. Red indicates a decrease in gene expression less than 0.5 fold. Annotations are from manual BLASTn searches in NCBI GenBank unless otherwise indicated.

Affymetrix probe set ID	Gene ID	Annotation	Fold Change	*p*-value
MYB transcription factors (positive regulators)				
PtpAffx.224252.1.S1_at	Potri.002G173900	MYB115	121.7093	5.27E-10
PtpAffx.8131.4.A1_a_at	Potri.006G221800	MYB134	22.59711	1.93E-10
PtpAffx.224943.1.S1_s_at	Potri.019G036400	MYB061	3.897258	2.63E-07
PtpAffx.224878.1.S1_at	Potri.014G100800	MYB201	3.168808	1.03E-06
Ptp.4185.1.S1_at	Potri.006G275900	MYB097	0.2146141	2.93E-05
MYB transcription factors (negative regulators)				
PtpAffx.137746.1.S1_at	Potri.004G088100	MYB182	3.744134	6.71E-05
PtpAffx.224650.1.S1_s_at	Potri.010G114000	MYB165	3.237289	0.000735
PtpAffx.224602.1.S1_at	Potri.008G128500	MYB194	2.503994	0.000485
PtpAffx.2935.2.A1_a_at	Potri.004G021300	MYB166	0.45789	0.0296401
PtpAffx.104683.1.S1_at	Potri.013G149200	MYB049	0.343932	1.43E-06
Other transcription factors				
PtpAffx.46526.1	Potri.004G181900	NAC domain trancription factor (NAC074)	0.496776	0.003228
PtpAffx.102143.1	Potri.001G079800	AP2 ERF domain-containing transcription factor (EFR6)	0.485004	0.01708
PtpAffx.163890.1	Potri.003G151000	AP2 ERF domain-containing transcription factor (ERF2)	0.479877	0.025837
PtpAffx.141636.1	Potri.009G019200	NAC domain trancription factor (NAC113)	0.433646	0.001565
PtpAffx.572.2	Potri.010G129200	AP2 ERF and B3 domain-containing (RAV1)	0.413082	0.006575
Ptp.3323.1	Potri.003G150800	AP2 ERF domain-containing transcription factor (ERF5)	0.372843	0.001773
PtpAffx.2179.1	Potri.007G099400	NAC domain trancription factor (NAC045)	0.371855	0.0112
PtpAffx.2179.2	Potri.005G069500	NAC domain trancription factor (NAC043)	0.357166	0.01988
PtpAffx.202716.1	Potri.002G255700	MADS box protein * (PANTHER:11945)	0.326733	2.75E-05
PtpAffx.203341.1	Potri.003G150700	AP2 ERF domain-containing transcription factor (ERF13)	0.265958	0.002664
Putative MYB co-factors				
PtpAffx.127289.1.A1_at	Potri.006G209000	WD40 repeat-containing protein [i] (AtLWD1 like [72.8 % identity])	16.70971	2.10E-07
PtpAffx.256.2.S1_at	Potri.005G208600	bHLH type protein [i] (TT8 like [50.5 % identity])	3.407876	8.15E-07
PtpAffx.213439.1.S1_at	Potri.016G075800	WD40 repeat-containing protein [i] (AtLWD1 like [67.5 % identity])	2.690502	0.000231

* indicates functional annotation from phytozome or Blast2Go

[i] indicates annotation is based on top Arabidopsis homolog through protein homology search in phytozome

Table 2-4. continued.

Affymetrix probe set ID		Annotation	Fold Change	*p*-value
Early Flavonoid Biosynthetic Genes				
PtpAffx.83404.1.A1_at	Potri.009G069100	flavonoid 3',5'-hydroxylase (CYP75A12|F3'5'H1)	138.3866	1.45E-07
PtpAffx.7896.2.S1_at	Potri.003G176700	chalcone synthase (CHS4)	64.90673	9.49E-11
PtpAffx.7896.4.A1_a_at	Potri.003G176900	chalcone synthase (CHS6)	50.9888	2.07E-11
PtpAffx.7896.3.S1_a_at	Potri.001G051600	chalcone synthase (CHS3)	39.651	1.80E-11
Ptp.6711.1.S1_s_at	Potri.014G145100	chalcone synthase (CHS1)	33.89923	3.20E-09
PtpAffx.142603.1.A1_s_at	Potri.013G073300	flavonoid 3'-hydroxylase (F3'H1)	23.07233	1.71E-10
Ptp.323.1.S1_s_at	Potri.005G113900	flavanone 3-hydroxylase (F3H6)	19.61242	1.62E-10
Ptp.1512.1.S1_s_at	Potri.019G057800	chalcone isomerase-like protein (CHIL2)	7.201737	5.13E-08
PtpAffx.4850.1.A1_s_at	Potri.010G213000	chalcone isomerase (CHI1)	4.163411	2.73E-08
PtpAffx.224381.1.S1_s_at	Potri.004G071000	quercetin O-glucosyltransferase-like protein * (GO:0080045)	2.466033	0.008391
PtpAffx.25980.1.S1_at	Potri.012G032700	flavonol 3-sulfotransferase-like protein * (GO:0047364)	2.191155	0.012064
PtpAffx.224381.1.S1_at	Potri.004G069600	quercetin O-glucosyltransferase-like protein * (GO:0080045)	2.177728	0.034052
PtpAffx.224177.1.S1_at	Potri.001G274600	flavonoid 3',5'-hydroxylase (CYP75A13 |F3'5'H2)	2.031447	0.001723
Ptp.323.1.S1_at	Potri.005G113700	flavonone 3-hydroxylase (F3H3)	2.000638	0.030033
PtpAffx.160390.2.A1_at	Potri.015G002800	flavanone 3-hydroxylase-like protein * (GO:0001666)	0.3542198	0.000641
Ptp.6598.1.S1_s_at	Potri.018G051300	flavonoid 3'5' hydroxylase -like (CYP92A25) * (GO:0033772)	0.2289472	2.82E-05
Ptp.6598.1.S1_at	Potri.006G141400	flavonoid 3'5'-hydroxylase-like (CYP92A24) * (GO:0033772)	0.2257996	5.85E-05
PtpAffx.3847.3.A1_at	Potri.001G167900	flavonoid 3'5'-hydroxylase-like (CYP92A17v1) * (GO:0033772)	0.1801926	5.85E-05
PtpAffx.103190.1.S1_at	Potri.003G066400	flavonoid 3'5'-hydroxylase-like (CYP92A20v2) * (GO:0033772)	0.1649159	0.000284
Ptp.4381.2.A1_a_at	Potri.003G066800	flavonoid 3'5'-hydroxylase-like (CYP92A20v2) * (GO:0033772)	0.1357423	3.28E-05
Late Flavonoid Biosynthetic Genes				
Ptp.6057.1.S1_at	Potri.001G113100	anthocyanidin synthase (ANS2)	41.48594	2.15E-09
PtpAffx.37082.1.A1_at	Potri.002G033600	dihydroflavonol 4-reductase (DFR1)	38.50957	1.21E-09
PtpAffx.6065.2.S1_at	Potri.008G116500	leucoanthocyanidin reductase (LAR1)	33.23252	1.56E-09
Ptp.3138.2.A1_a_at	Potri.009G133300	UDP-glucose:flavonoid glucosyltransferase (PtUGT78L1)	29.8633	1.67E-06
PtpAffx.18705.2.A1_a_at	Potri.015G050200	leucoanthocyanidin reductase (LAR3)	21.52007	2.85E-07
PtpAffx.5092.2.S1_a_at	Potri.004G030700	anthocyanidin reductase (ANR1)	19.81152	2.29E-10
PtpAffx.25553.1.A1_at	Potri.005G229500	dihydroflavonol 4-reductase (DFR2)	15.60789	2.33E-08
Ptp.1080.1.S1_at	Potri.010G129800	leucoanthocyanidin reductase (LAR2)	7.835983	5.49E-08
PtpAffx.159543.1.A1_at	Potri.006G101100	similar to putative ANS in *Ricinus communis* [i] (86.7% identity)	4.057872	0.000112
PtpAffx.220100.1.S1_at	Potri.001G133000	anthocyanidin 3-O-glucosyltransferase-like protein *(GO:0047213)	2.440327	0.001164

* indicates functional annotation from phytozome or Blast2Go

[i] indicates annotation is based on top Arabidopsis homolog through protein homology search in phytzome

Table 2-4. continued.

Affymetrix probe set ID		Annotation	Fold Change	*p*-value
Putative flavonoid transporters				
PtpAffx.224485.1.S1_s_at	Potri.005G207600	(MATE) family transporter-related [i] (AtTT12 like [82.6 % identity])	48.93626	4.81E-10
PtpAffx.94822.1.A1_at	Potri.002G055100	(MATE) family transporter-related (AtTT12 like [81.5 % identity])	7.284551	1.57E-07
PtpAffx.211115.1.S1_at	Potri.013G069200	(MATE) family transporter-related * (PANTHER:11206)	3.945352	6.92E-07
PtpAffx.119354.1.A1_at	Potri.012G144900	(MATE) family transporter-related (AtFFT like [71.8 % idenitity])	0.3980224	8.72E-05
PtpAffx.119354.2.A1_s_at	Potri.015G147600	(MATE) family transporter-related (AtFFT like [68.0 % identity])	0.3806387	6.65E-06
PtpAffx.224677.1.S1_at	Potri.011G002400	(MATE) family transporter-related * PANTHER:11206	0.3143611	3.59E-05
Ptp.3040.1.A1_at	Potri.011G002500	(MATE) family transporter-related * PANTHER:11206	0.2692467	0.000129
Phenylpropanoid pathway				
PtpAffx.161181.1.S1_at	Potri.006G178700	cinnamoyl-CoA reductase-like protein * (KOG:1502)	35.7962	3.53E-08
PtpAffx.7740.2.A1_a_at	Potri.018G100500	cinnamoyl-CoA reductase-like protein * (KOG:1502)	11.3066	1.27E-08
Ptp.6632.1.S1_at	Potri.019G130700	cinnamate 4-hydroxylase (C4H1|CYP73A43)	8.199498	3.11E-07
PtpAffx.150025.1.S1_s_at	Potri.013G157900	cinnamate 4-hydroxylase (C4H2|CYP73A42)	7.907291	8.37E-08
PtpAffx.30128.1.S1_at	Potri.001G140700	cinnamoyl-CoA reductase-like protein * (KOG:1502)	6.969607	1.60E-05
PtpAffx.12353.1.A1_at	Potri.018G146100	trans-cinnamate 4-monooxygenase (C4H3)	5.992858	5.54E-05
PtpAffx.12056.3.S1_a_at	Potri.T071600	4-Coumarate:CoA ligase (4CL4)	5.802742	2.97E-07
PtpAffx.34213.1.S1_at	Potri.013G083600	peroxidase (lignin biosynthesis) * (GO:0004601)	4.7162	3.79E-06
PtpAffx.225544.1.S1_s_at	Potri.016G031000	p-coumaroyl shikimate 3'- hydroxylase (C3H3)	2.560835	8.08E-06
PtpAffx.55376.1.S1_at	Potri.014G143200	peroxidase (lignin biosynthesis) * (GO:0004601)	2.409356	0.004393
PtpAffx.225544.1.S1_x_at	Potri.016G031100	p-coumaroyl shikimate 3'- hydroxylase (CYP98A23)	2.381891	1.49E-05
PtpAffx.20905.1.S1_at	Potri.018G075000	aldehyde dehydrogenase-like protein * (Panther:11699)	0.442972	0.000313
Ptp.4093.1.S1_at	Potri.019G043600	PtACT49[†]	0.3321203	0.000169779
PtpAffx.6696.2.S1_at	Potri.013G074500	PtACT47[†]	0.2396261	0.000196127
PtpAffx.3594.1.S1_at	Potri.002G018300	cinnamyl alcohol dehydrogenase-like protein (CADL2)	0.1573671	8.41E-05

* indicates functional annotation from phytozome or Blast2Go.

[i] indicates annotation is based on top Arabidopsis homolog through protein homology search in phytzome.

[†] indicates genes are named based on naming system developed by (Yu et al., 2009).

2.3.8 Validation of microarray analysis by qPCR

In order to validate the results of microarray analysis of changes in gene expression in MYB115 overexpressing plants compared to WT plants, six genes of biological interest that showed higher expression in 353-38 MYB115 Line 5 were further tested by qPCR. Three deregulated MYB transcription factors were tested, the transgene MYB115, the characterized regulator of condensed tannin synthesis, MYB134, and the putative repressor of condensed tannin synthesis, MYB182. Three biosynthetic genes were also tested, the late flavonoid biosynthetic gene, DFR2, the tannin specific gene, ANR1, and F3'5'H1 which has a putative function in the synthesis of dihydromyricetin. The log fold change of gene expression from qPCR analysis was calculated for 353-38 MYB115 Line 5 compared to WT plants and correlated with log fold change from microarray analysis. There was a strong linear correlation between qPCR and microarray results ($R^2 = 0.85$; Figure 2-9). This confirms that the microarray results accurately reflect the actual changes in gene expression.

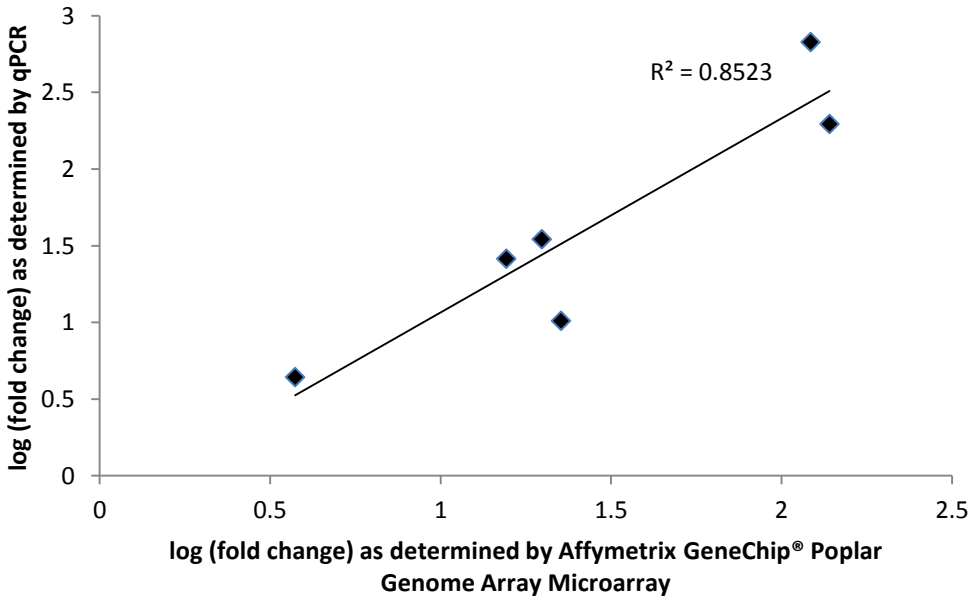

Figure 2-9. Validation of microarray results using qPCR.
The log fold change of Affymetrix microarray results and qPCR analysis of six genes of biological interest are plotted. Data represents three plants for 353 MYB115 Line 5 compared with three 353 WT plants.

2.3.9 qPCR analysis of key flavonoid biosynthetic genes and regulatory factors in MYB115 overexpressing plants

Expression of the six genes used for validation of the microarray data was also tested in the remaining 353-38 transgenic lines and all 717-1-B4 transgenic lines. MYB134 is the known regulator of condensed tannin synthesis in poplar (Mellway et al., 2009). Microarray analysis showed a 22.6 fold induction of MYB134 in MYB115 overexpressing plants (Table 2-4). qPCR analysis showed an apparent 10.2 fold change in expression in 353-38 Line 5 plants compared to wild-type plants but the effect was not significant (one-way ANOVA, $p > 0.05$) (Figure 2-10). 353-38 Line 4 showed an apparent fold change of 2.8 but the effect was not significant. None of the other lines showed a change in expression greater than two fold. Expression of MYB134 did not correlate with expression of MYB115 in either 717-1-B4 or 353-38 plants (Table 2-4). This indicates that MYB115 is likely not a regulator MYB134 expression.

MYB182 is a putative repressor of condensed tannin synthesis (K. Yoshida and C.P. Constabel, unpublished). MYB182 showed a significant induction (3.7 fold) in MYB115 overexpressing plants as assayed by microarray analysis (Table 2-3). qPCR analysis resulted in an apparent comparable fold change induction in 353-38 Line 5 (4.4 fold), but the effect was not significant (one-way ANOVA, $p > 0.05$) (Figure 2-11). No other transgenic lines showed an induction greater than two fold as tested by qPCR. MYB182 expression correlated strongly with the expression of MYB134, but not MYB115, in 353-38 plants (Table 2-4).

F3’5’H1 showed the highest fold change (138.4 fold) in MYB115 overexpressing plants compared to WT plants as analyzed by microarray (Table 2-3). This fold change was greater than twice as high as the next highest upregulated flavonoid biosynthetic gene, CHS4 (64.9 fold). qPCR analysis of 353-38 Line 5 validated the dramatic increase in transcript abundance showing a significant increase in expression levels of F3’5’H1 (196.3 fold change) compared to WT plants (one-way ANOVA, $p \leq 0.05$) (Figure 2-12). 353-38 Line 4 showed an apparent 95.4 fold change in expression of F3’5’H1 but the

effect was not significant (one-way ANOVA, $p > 0.05$). An increase F3'5'H1 expression was only seen in high tannin MYB115 transgenic plants (e.g. 353-38 Lines 4 and 5). 353-38 Lines 1, 2 and 3 showed an apparent slight but not significant decrease in expression (one-way ANOVA, $p > 0.05$). All 717-1-B4 lines showed an average fold change greater than two fold above WT plants (Figure 2-12). However, only Line 1 showed a significant effect of MYB115 overexpression on expression of F3'5'H1 (one-way ANOVA, $p \leq 0.05$).

DFR2 showed significantly higher expression (15.6 fold) in MYB115 transgenic lines compared to WT as analyzed by microarray (Table 2-3). DFR2 converts dihydroflavonols into leucoanthocyanidins and functions in the synthesis of condensed tannins (Huang et al., 2012). qPCR showed a significant fold change (26.0 fold) in expression of DFR2 in 353-38 Line 5 compared to wild-type plants (one-way ANOVA, $p \leq 0.001$) (Figure 2-13). All other transgenic lines showed a fold change of DFR2 expression greater than two fold above WT plants with the exception of 717-1-B4 Line 2 (Figure 2-14). However, the effect of MYB115 overexpression on DFR2 expression was only significant for 353-38 Line 5 and 717-1-B4 Line 3 (one-way ANOVA, $p \leq 0.05$). DFR2 expression correlated with tannin levels in both 717-1-B4 and 353-38 plants and weakly correlated with expression of both MYB115 and MYB134 in 353-38 plants and (Table 2-4; Appendix A, Table A-3).

ANR1 also showed significantly higher expression (19.8 fold) in MYB115 overexpressing plants compared to WT as analyzed by microarray (Table 2-3). ANR1 is involved in the synthesis of flavan-3-ols, direct pre-cursors of condensed tannins. qPCR analysis showed all transgenic lines analyzed had a greater than two fold change in expression of ANR1 compared to WT plants (Figure 2-14). Expression of ANR1 shows correlation with MYB115 expression in 353-38 plants but not tannin levels (Table 2-4). Of all genes tested by qPCR, ANR1 was the only gene that correlated with MYB115 expression in 717-1-B4 plants (Table 2-4).

Overall, these results show that MYB115 overexpression led to an increase in expression to flavonoid pathway genes leading to condensed tannin synthesis. While MYB115 expression is high in all transgenic lines tested, both DFR2 and F3'5'H1 are only

upregulated in plants that show a high tannin phenotype. By contrast, ANR1 expression correlates with MYB115 expression and not tannin levels. MYB115 expression does not correlate with expression of the positive tannin regulator, MYB134, nor the putative negative regulator, MYB182. This suggests that MYB115 expression does not directly influence changes in expression of these other condensed tannin regulators, and that other regulators and feedback mechanisms are likely involved (Mellway, 2009; Chapter 3).

Table 2-4. Linear regression analysis between relative gene expression of the six genes analyzed by qPCR and expression of MYB115 and MYB134 in WT and MYB115 overexpressing lines.

Data shown include all positive transformants for both the 717 (*P. tremula* x *P. alba* [INRA 717-1-B4]) and 353 (*P. tremula* x *P. tremuloides* [INRA 353-38]) backgrounds. Relative gene expression was measured by qPCR and normalized against elongation factor 1-β. R^2 values of > 0.5 are bolded. Asterisks indicate the level of significance of the R^2 value ($p \leq 0.05$, *; $p \leq 0.01$, **, $p \leq 0.001$, ***).

	353 (R^2)		717 (R^2)	
	MYB115	MYB134	MYB115	MYB134
MYB134	0.2581*		0.1647	
MYB182	0.1604	**0.8825*** **	0.3525*	0.07455
F3'5'H1	**0.5488*** **	0.4846**	0.02028	0.01741
DFR2	**0.6168*** **	**0.5358*** **	0.3444*	0.1603
ANR1	**0.7362*** **	0.3633*	**0.5985****	0.2683

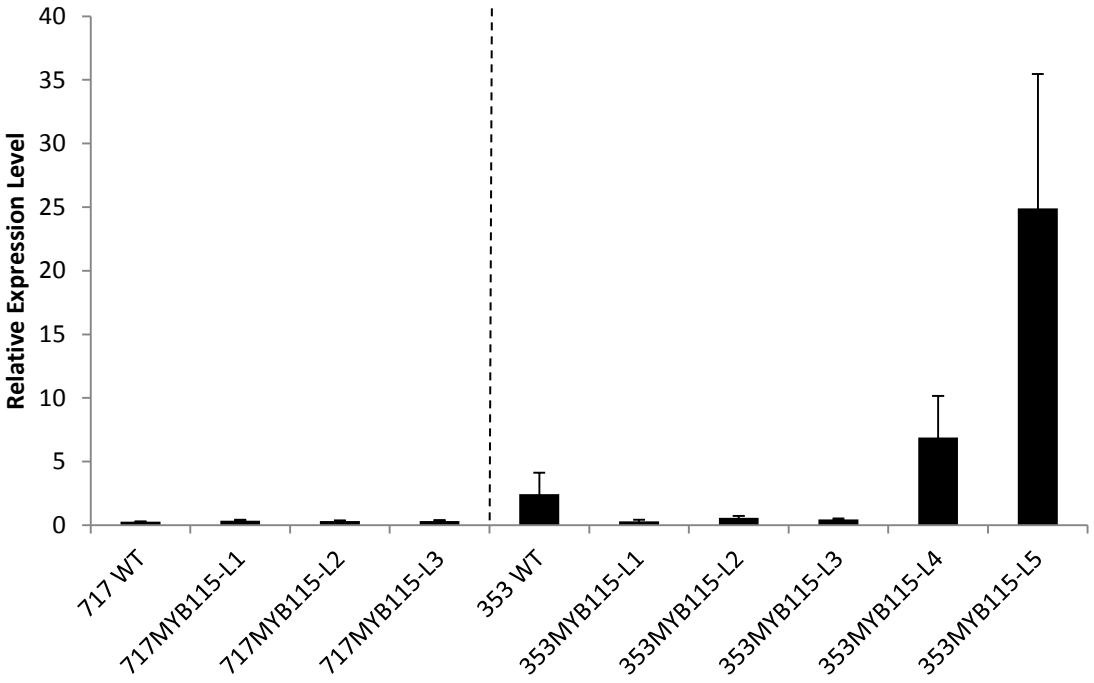

Figure 2-10. Relative expression of MYB134 in MYB115 overexpressing poplar as analyzed by qPCR.

Data shown include all positive transformants for both the 717 (*P. tremula* x *P. alba* [INRA 717-1-B4]) and 353 (*P. tremula* x *P. tremuloides* [INRA 353-38]) backgrounds. Lines with different genetic backgrounds are separated with a dashed line. Expression is normalized against elongation factor 1-β, a housekeeping gene. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants; 353MYB115-L1: n = 2 plants).

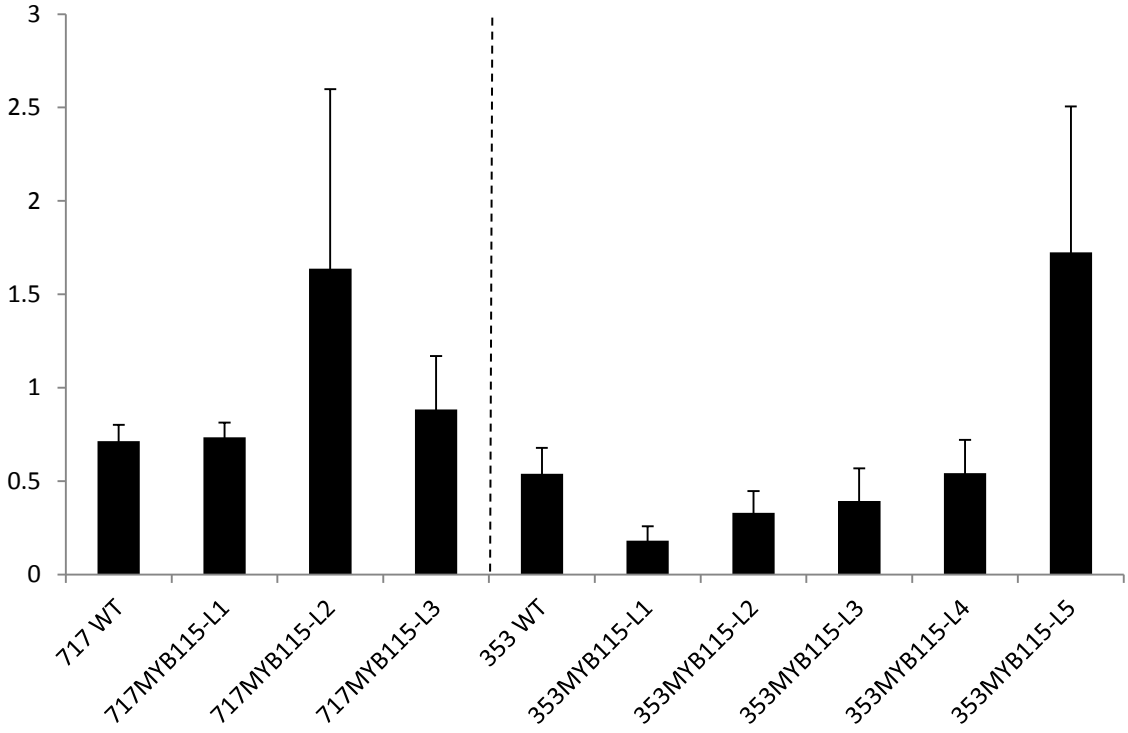

Figure 2-11. Relative expression of MYB182 in MYB115 overexpressing poplar as analyzed by qPCR.

Expression is normalized against elongation factor 1-β, a housekeeping gene. Each line (L1, L2 etc.) represents an independent transformation event. Lines with different genetic backgrounds are separated with a dashed line. Error bars indicate standard error (n ≥ 3 plants; 353MYB115-L1: n = 2 plants).

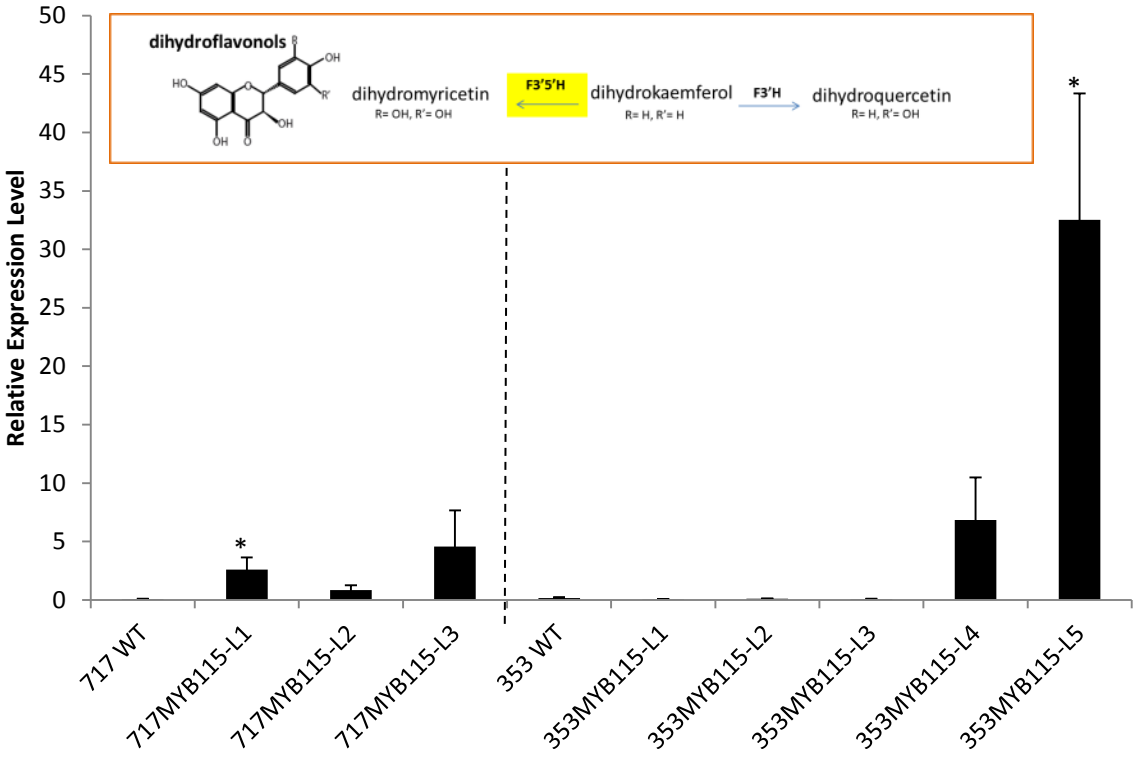

Figure 2-12. Relative expression of F3'5'H1 in MYB115 overexpressing poplar as analyzed by qPCR.

Expression is normalized against elongation factor 1-β, a housekeeping gene. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants; 353MYB115-L1: n = 2 plants). Lines with different genetic backgrounds are separated with a dashed line. Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). Also shown is the functional role of F3'5'H in the flavonoid pathway (refer to Figure 1-2).

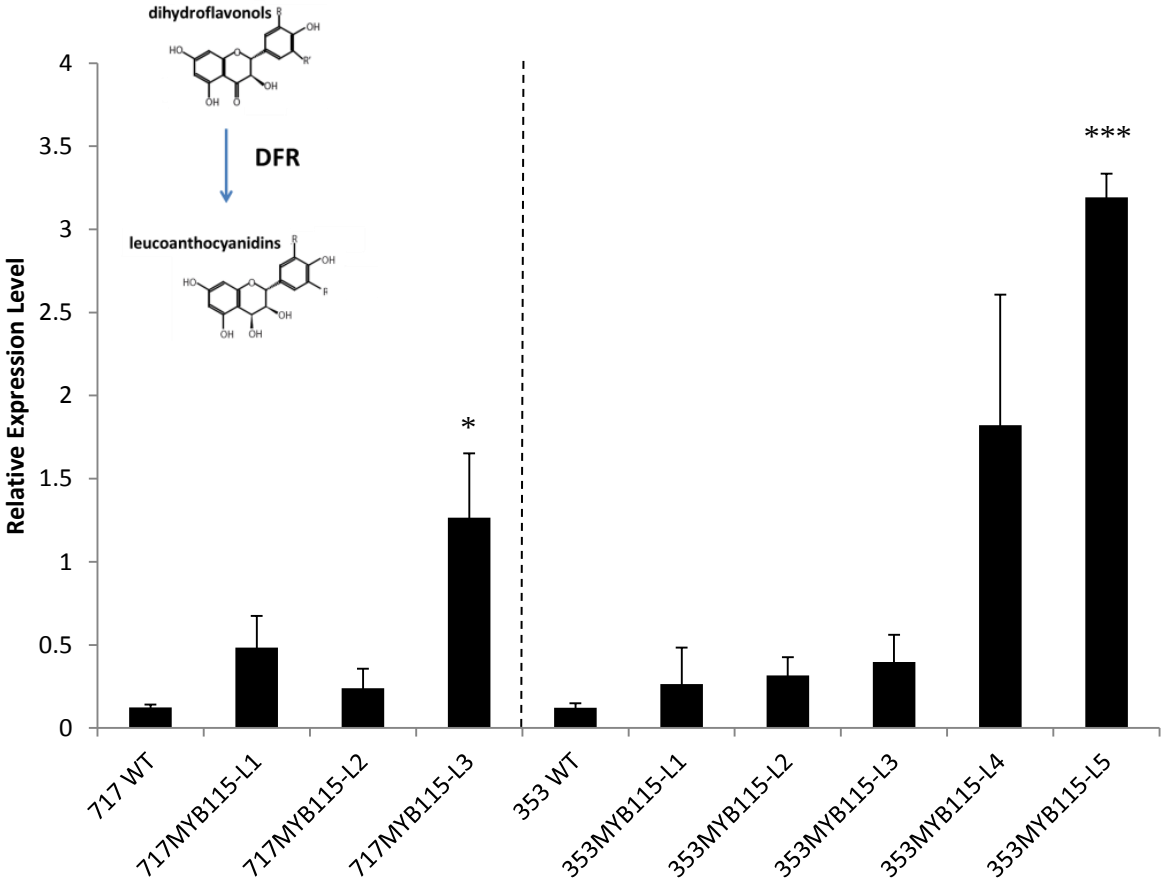

Figure 2-13. Expression of PtDFR2 in MYB115 overexpressing poplar.
Expression is normalized against elongation factor 1-β, a housekeeping gene. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants; 353MYB115-L1: n = 2 plants). Lines with different genetic backgrounds are separated with a dashed line. Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). Also shown is the functional role of DFR in the flavonoid pathway.

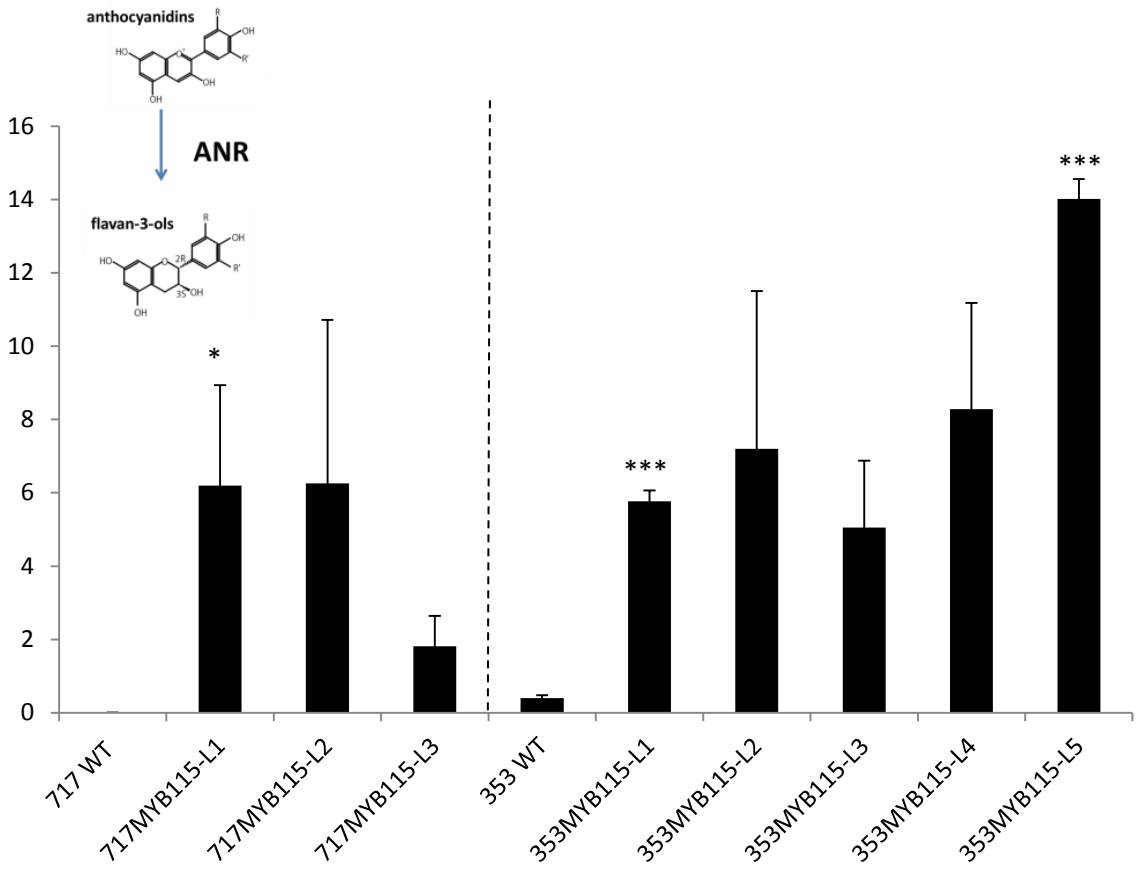

Figure 2-14. Expression of PtANR1 in MYB115 overexpressing poplar.
Expression is normalized against elongation factor 1-β, a housekeeping gene. Each line (L1, L2 etc.) represents an independent transformation event. Lines with different genetic backgrounds are separated with a dashed line. Error bars indicate standard error (n ≥ 3 plants; 353MYB115-L1: n = 2 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). Also shown is the functional role of ANR in the flavonoid pathway.

2.3.4 Transgenic plants overexpressing MYB115 show reduced accumulation of phenolic glycosides compared with wild-type plants

Previous work found that in MYB134 overexpressing high tannin plants levels of phenolic glycosides were unexpectedly reduced relative to control plants (Mellway et al., 2009). In the current project, phenolic glycosides were quantified by HPLC using purified salicin as a standard. The major phenolic glycosides salicin, salicortin and tremulacin were identified based on their UV spectra, their retention time, and fragmentation pattern from LC/MS analysis. Salicortin and tremulacin were significantly reduced in only the high tannin 353-38 MYB115 transgenic plants (Lines 4 and 5) (Figure 2-16). Therefore, there was a strong inverse correlation in 353-38 plants between tannin concentration and salicortin or tremulacin (Appendix A, Table A-3). None of the 353-38 MYB115 overexpressing lines showed a significant reduction in salicin (Figure 2-16).

717-1-B4 MYB115 overexpressing plants also showed a significant reduction in two salicinoid phenolic glycosides, tremulacin and salicin (Figure 2-15). Two of the three 717-1-B4 transgenic lines showed a significant reduction in the salicinoid phenolic glycoside, salicortin. In 717-1-B4 plants, there was a weak inverse correlation between the concentrations of salicortin and tremulacin and the concentration of condensed tannins (Appendix A, Table A-3).

Grandidentatin is a *p*-coumaroyl-glucopyranoside which has been detected in *Populus* and *Salix* (Pearl and Darling, 1962, 1970). The identities of two isomers of grandidentatin were tentatively identified by LC/MS (M. Reichelt and C.P. Constabel, unpublished data). They were then quantified by HPLC using chlorogenic acid as a standard equivalent. Two isomers of grandidentatin were significantly reduced in high tannin 353-38 MYB overexpressing plants (Lines 4 and 5) (Figure 2-18). Therefore, there was a strong inverse correlation in 353-38 plants between tannin concentration and concentrations of grandidentatin derivatives (Table 2-3). There was a slight apparent reduction of grandidentatin derivative-2 in 717-1-B4 plants but the effect is not significant (Figure 2-17).

The overall trend observed from these data is that concentrations of phenolic glycosides, both salicin based phenolic glycosides and grandidentatin derivatives, are decreased in MYB115 overexpressing plants and are therefore inversely correlated with condensed tannin levels. This suggests that MYB115 is negatively regulating the synthesis of phenolic glycosides or that they are reduced by an indirect mechanism.

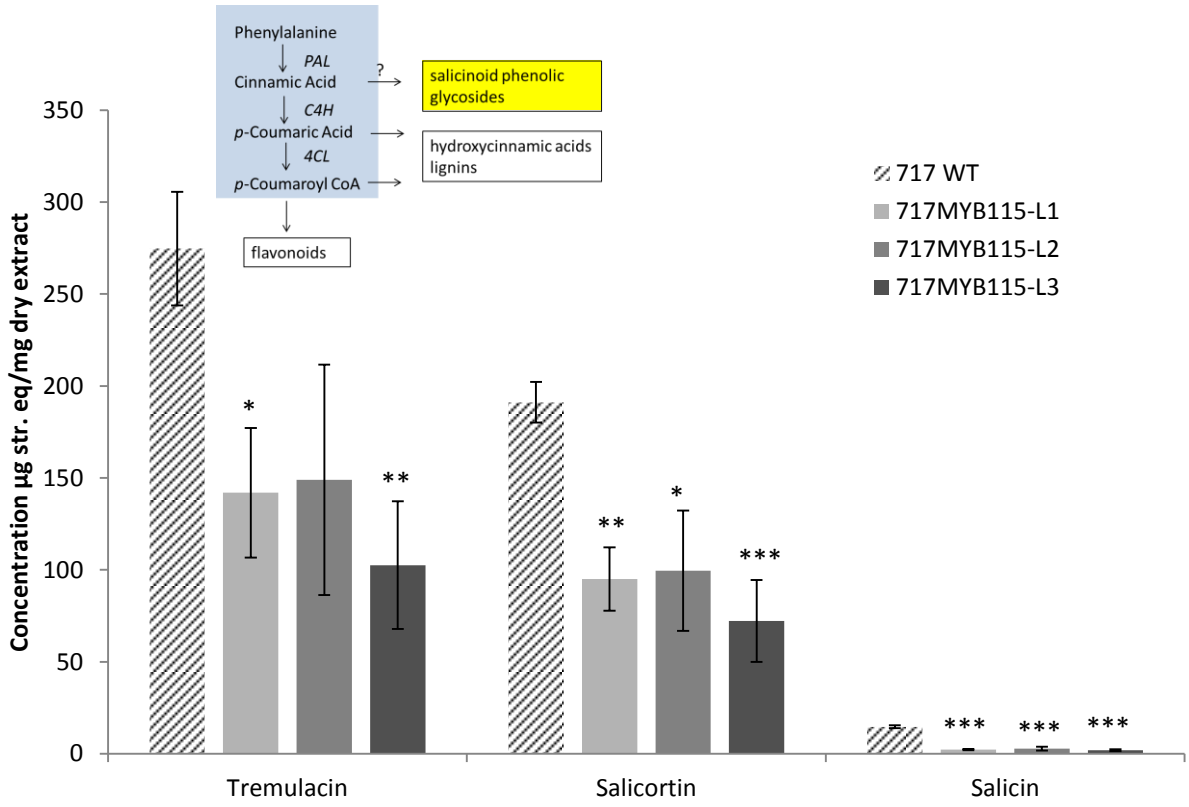

Figure 2-15. Concentration of salicinoid phenolic glycosides in 717-1-B4 wild-type (WT) and 717-1-B4 MYB115 overexpressing poplar as anlayzed by HPLC.
Concentrations are calculated as a salicin standard equivalent. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). The location of phenolic glycosides in the general phenylpropanoid pathway is also shown (refer to Figure 1-1).

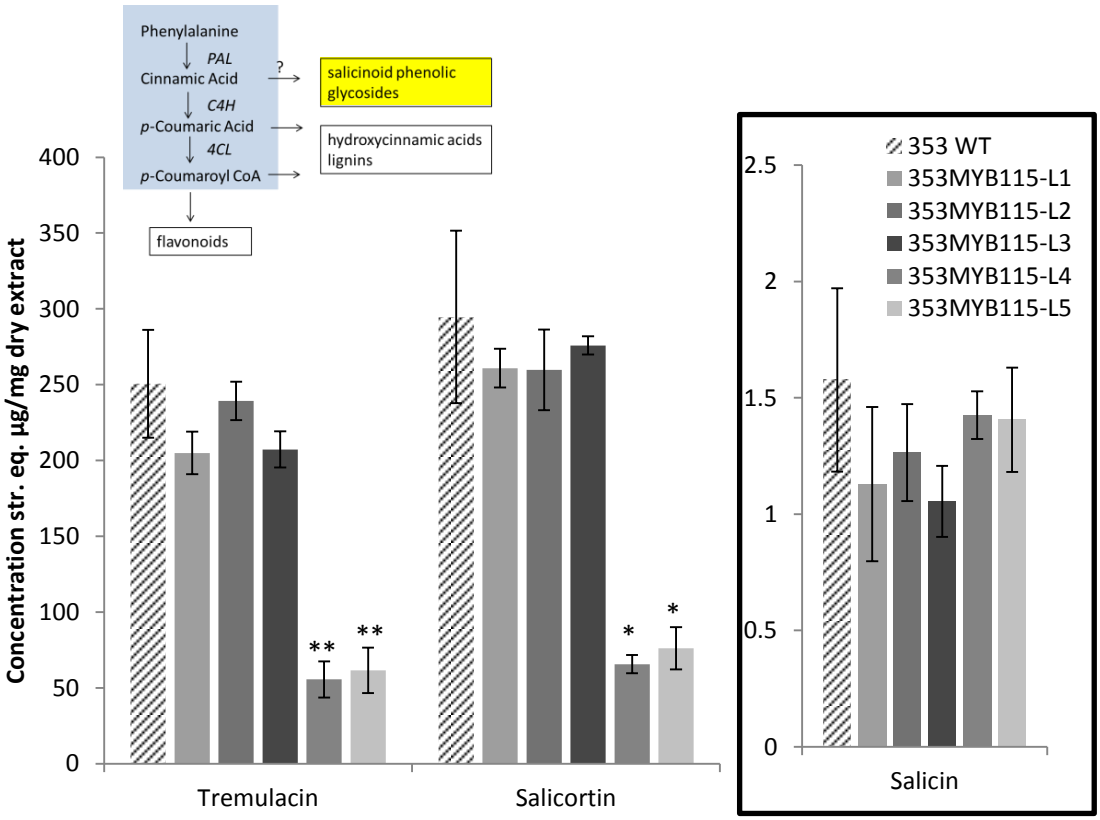

Figure 2-16. Concentration of salicinoid phenolic glycosides in 353-38 wild-type (WT) and 353-38 MYB115 overexpressing poplar as anlayzed by HPLC.
Concentrations are calculated as a salicin standard equivalent. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n = 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). The location of phenolic glycosides in the general phenylpropanoid pathway is also shown (refer to Figure 1-1).

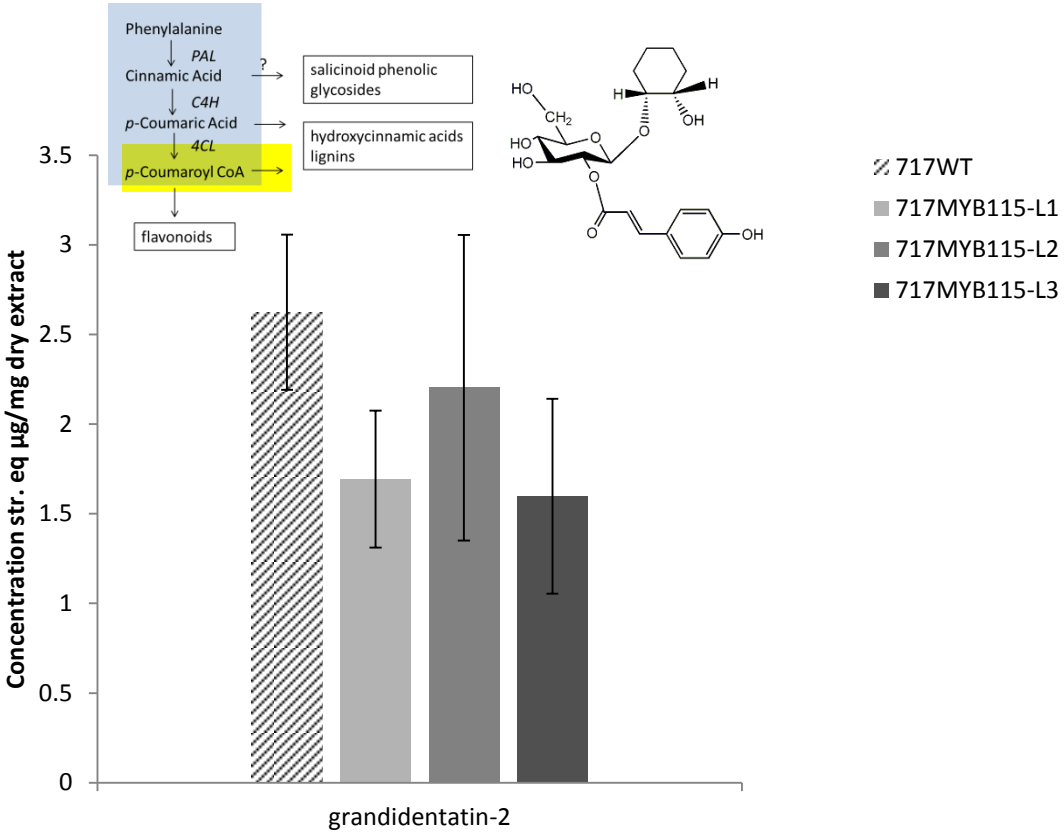

Figure 2-17. Concentration of the phenolic glycoside, grandidentatin (derivative 2), in 717-1-B4 wild-type (WT) and MYB115 overexpressing poplar.

Concentrations are calculated as a chlorogenic acid standard equivalent. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). The biosynthesis of grandidentatin is unknown, however, it is likely partly derived from *p*-coumaroyl CoA which is highlighted in yellow in the general phenylpropanoid pathway shown here. The chemical structure of grandidentatin is also shown.

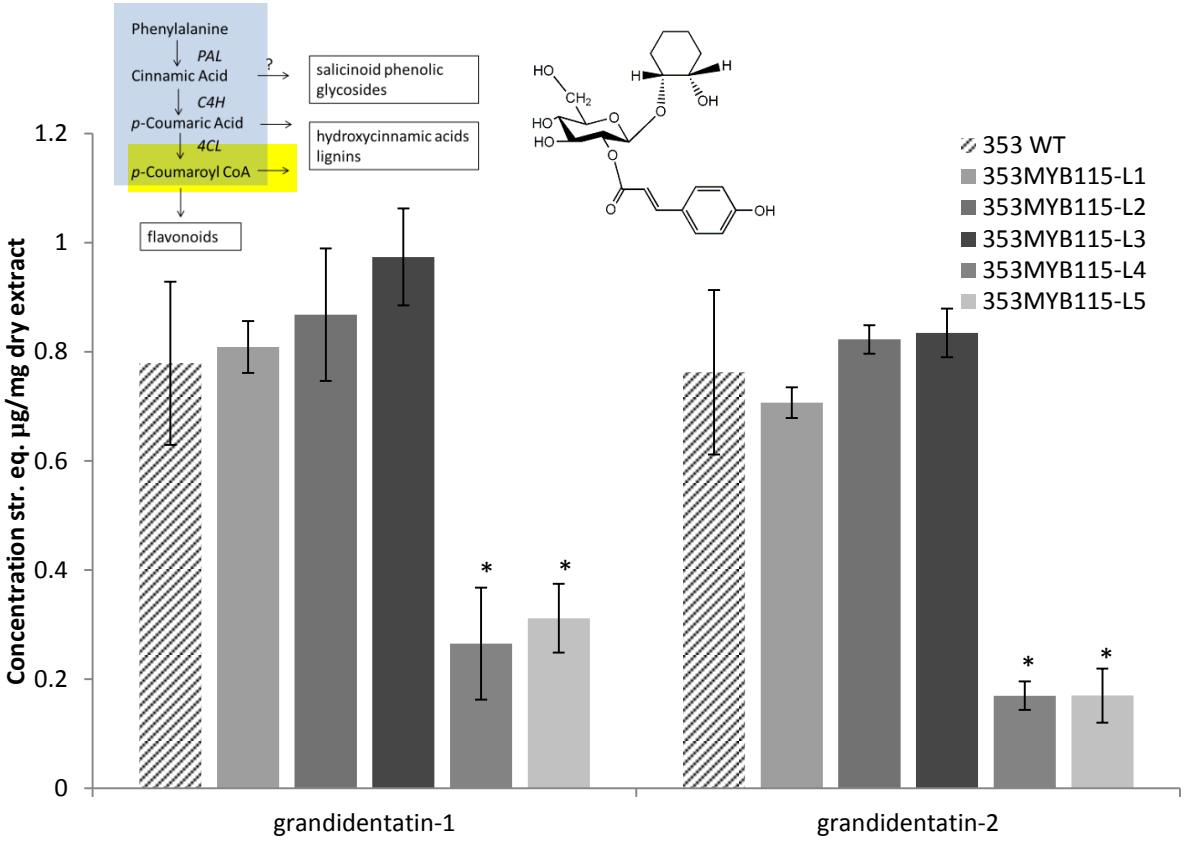

Figure 2-18. Concentration of the phenolic glycosides, grandidentatin derivative 1 and 2, in 353-38 wild-type (WT) and MYB115 overexpressing poplar.

Concentrations are calculated as a chlorogenic acid standard equivalent. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n = 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). The biosynthesis of grandidentatin is unknown, however, it is likely partly derived from *p*-coumaroyl CoA which is highlighted in yellow in the general phenylpropanoid pathway shown here. The chemical structure of grandidentatin is also shown.

2.3.5 Transgenic plants overexpressing MYB115 show changes in accumulation of hydroxycinnamic acid derivatives compared with wild-type plants

Hydroxycinnamic acids (HCA) derivatives are phenolic acids with a C6-C3 skeleton synthesized by hydroxylation and O-methylation of cinnamic acid. Hydroxycinnamic acid composition is of interest due to the potential role of these compounds in plant defense (see Chapter One). HCA derivatives were tentatively identified here based on their UV spectra compared with a chlorogenic acid standard. Compounds were tentatively identified as chlorogenic acid derivatives and coumaroyl-quinic acid derivatives based on their fragmentation pattern from LC/MS analysis (M. Reichelt and C.P. Constabel, unpublished data). Compounds that were not further characterized by LC/MS analysis are designated here as HCA1/2/3. In total, 12 HCA derivatives in 717-1-B4 plants and five HCA derivatives in 353-38 plants were tentatively identified based on their UV spectra and quantified. Only HCA derivatives that showed statistically significant changes in more than one transgenic line compared to WT plants are reported here. The majority of HCAs are increased in transgenic plants overexpressing MYB115 compared to WT plants. Two compounds were tentatively characterized as chlorogenic acid derivatives. Chlorogenic acids are esters of caffeic acid and quinic acid. Chlorogenic acids differ in the position of the caffeoyl moiety. The two chlorogenic acid derivatives identified here were significantly enhanced (ANOVA, $p \leq 0.05$) in 717-1-B4 MYB115 overexpressing lines (Figure 2-19). Chlorogenic acid-2 showed significantly higher concentrations in all 353-38 transgenic lines while chlorogenic acid-1 was only significantly higher in two of the five lines (Figure 2-20). In 717-1-B4 plants, chlorogenic acid-2 shows a strong correlation with condensed tannins (Appendix A, Table A-3).

Two other compounds were tentatively identified as coumaroyl-quinic acid type HCAs. Coumaroyl-quinic acid is a coumaric ester of quinic acid. Coumaroyl-quinic acids differ in their position of the coumaroyl moiety. The coumaroyl-quinic acid compounds accumulated at significantly higher (ANOVA, $p \leq 0.05$) levels in 717-1-B4 and 353-38

MYB115 overexpressing lines compared with WT plants (with the exception of coumaroyl quinic acid-1 in 353-38 MYB115 Line 4) (Figures 2-19, 2-20). In 717-1-B4 plants, coumaroyl-quinic acid-2 shows a strong correlation with condensed tannins (Appendix A, Table A-3). Only one HCA type compound showed a significant reduction in MYB115 overexpressing plants and is designated here as HCA-3. HCA-3 was only identified in 353-38 plants and showed a decrease in high tannin MYB115 overexpressing plants (line 4 and 5) compared to the WT. The concentration of HCA-3 shows a strong negative correlation with the concentration of condensed tannins (Appendix A, Table A-3). Further analysis is needed to determine the structure of HCA-3 as it was only tentatively identified based on its UV spectrum. Therefore, with the exception of HCA-3, the overall trend seen in MYB115 overexpressing plants was that concentrations of HCAs correlated with MYB115 expression.

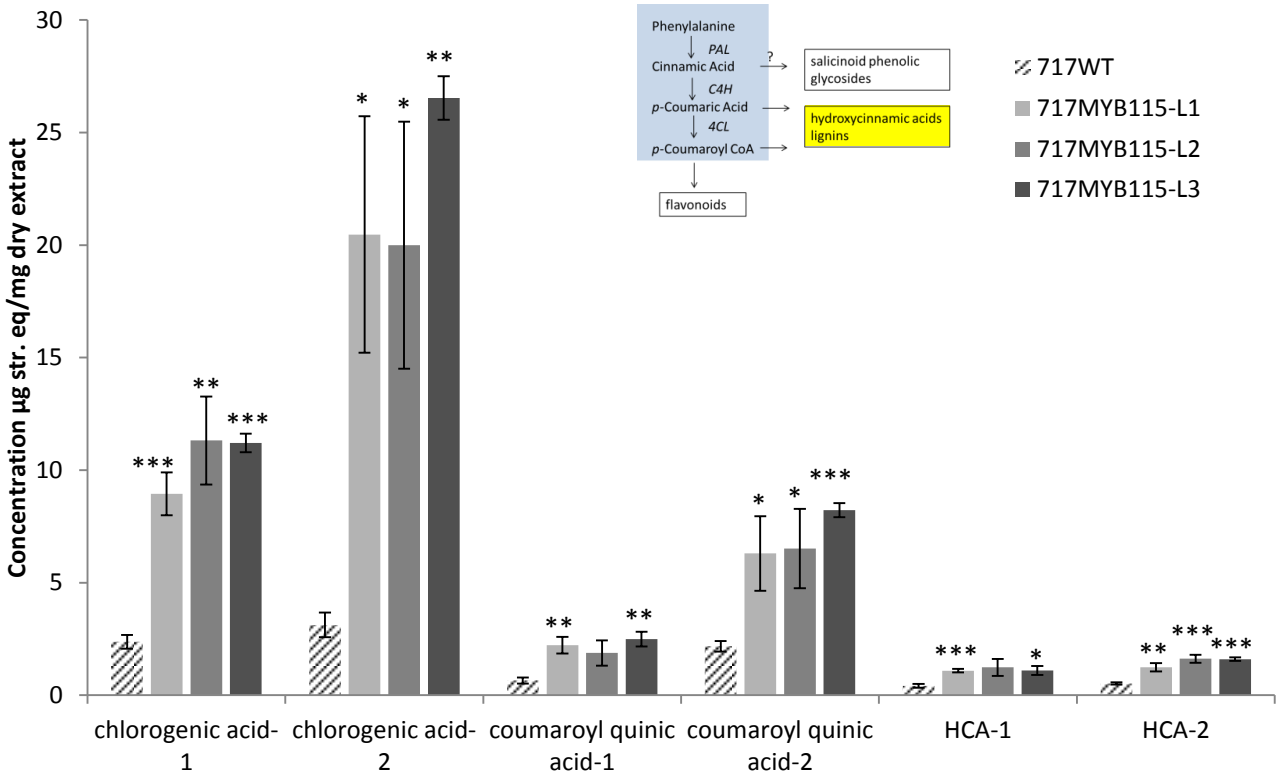

Figure 2-19. Concentration of hydroxycinnamic acid (HCA) derivatives in 717-1-B4 wild-type (WT) and 717-1-B4 MYB overexpressing poplar.
Concentrations are calculated as a chlorogenic acid standard equivalent. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n ≥ 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). The location of hydroxycinnamic acids in the general phenylpropanoid pathway is also shown (refer to Figure 1-1). Chlorogenic acid - 1 (retention time: 13.4 minutes, UV spectrum: 300, 326). Chlorogenic acid - 2 (retention time: 9 minutes, UV spectrum: 300, 326). Coumaroyl quinic acid - 1 (retention time: 11.2 minutes, UV spectrum: sh 300, 308). Coumaroyl quinic acid - 2 (retention time: 11.8 minutes, UV spectrum: sh300, 314). HCA-1 (retention time: 15.7 minutes, UV spectrum: sh 304, 312). HCA-2 (retention time: 22.9 minutes, UV spectrum: sh 300, 322).

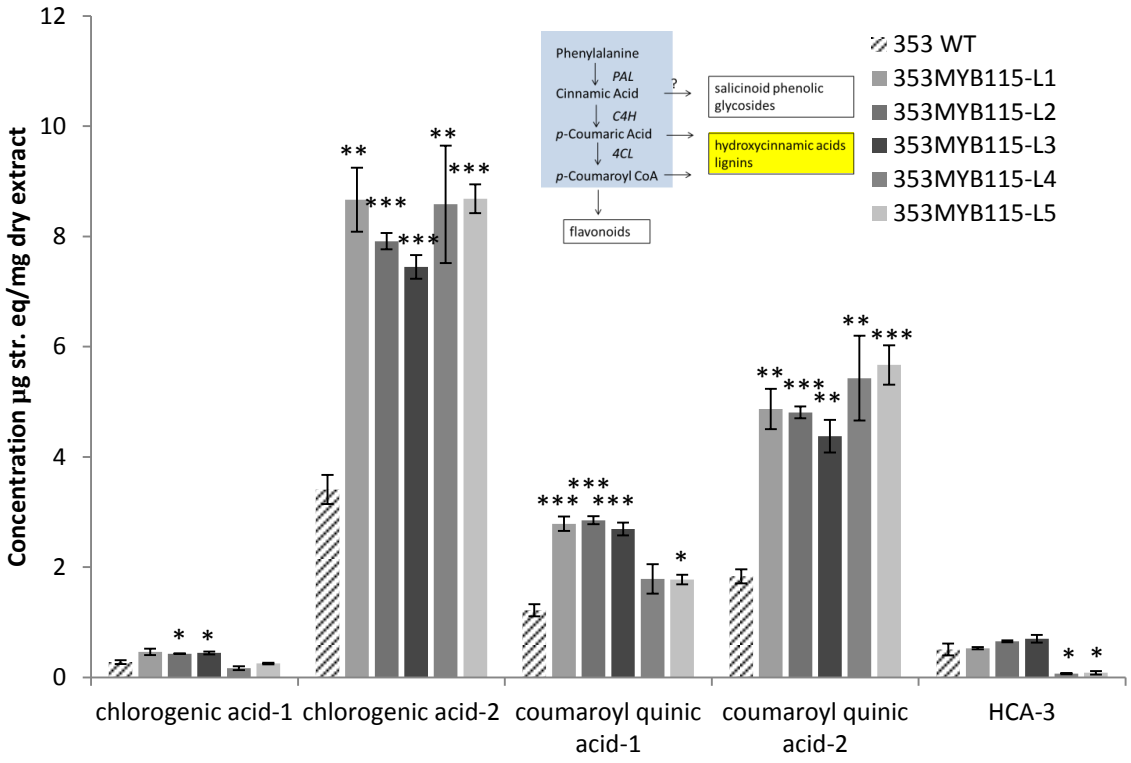

Figure 2-20. Concentration of hydroxycinnamic acid (HCA) derivatives in 353-38 wild-type (WT) and 353-38 MYB overexpressing plants.

Concentrations are calculated as a chlorogenic acid standard equivalent. Each line (L1, L2 etc.) represents an independent transformation event. Error bars indicate standard error (n = 3 plants). Asterisks indicate results of a one-way ANOVA ($p \leq 0.05$, *; $p \leq 0.01$, **; $p \leq 0.001$, ***). The location of hydroxycinnamic acids in the general phenylpropanoid pathway is also shown (refer to Figure 1-1). Chlorogenic acid-1 (retention time: 13 minutes, UV spectrum: 300, 326). Chlorogenic acid-2 (retention time: 9 minutes, UV spectrum: 300, 326). Coumaroyl quinic acid-1 (retention time: 11.1 minutes, UV spectrum: sh 298, 306). Coumaroyl quinic acid-2 (retention time: 11.8 minutes, UV spectrum: sh 300, 312). HCA-3 (retention time: 28.7 minutes, UV spectrum: sh 305, 315).

2.4 Discussion

2.4.1 Biochemical and expression analysis of MYB115 overexpressing plants reveals that MYB115 regulates the synthesis of condensed tannins

MYB115 was predicted to function in regulation of condensed tannin synthesis in poplar as it shows enhanced expression in high tannin transgenic plants and shows sequence similarity to known regulators of condensed tannin synthesis in grape and persimmon. In order to functionally characterize MYB115, transgenic plants overexpressing MYB115 were generated. Transgenic plants showed a dramatic increase in the accumulation of condensed tannins (Figure 2-3). This increase in condensed tannin synthesis was a result of increased expression, compared to WT plants, of flavonoid pathway genes leading to the synthesis of condensed tannins, including the tannin specific genes ANR1 and LAR1/2/3 (Table 2-3). Tannin levels correlated with concentrations of the tannin precursors dihydromyricetin, dihydroquercetin, and catechin (Appendix A, Table A-3). Furthermore, LC/MS analysis showed that flavan-3-ols and flavan-3-ol dimers showed enhanced accumulation in high tannin transgenic lines compared to WT plants (Figures 2-7, 2-8). Importantly, FLS, the branch point enzyme leading to flavonol synthesis, did not show enhanced expression, indicating that MYB115 is specifically regulating condensed tannin synthesis and not the other flavonoid branches. This is consistent with *in silico* co-expression analysis (Table 2-2). Together, these results support the role of MYB115 as a regulator of the flavonoid pathway leading to condensed tannins.

Microarray analysis revealed that both early (CHS1/3/4/6, CHI1, CHIL2, F3H3/6, F3'H1, F3'5'H1/2) and late (DFR1/2, ANS2, ANR1, LAR1/2/3) flavonoid biosynthetic genes leading to tannin biosynthesis show enhanced expression in MYB115 overexpressing plants compared to WT plants (Table 2-3). Some transcripts annotated as modifiers of flavonoids that could function in flavonoid transport showed enhanced expression including two glucosyltransferases (Pang et al., 2013). One of these has been characterized to function in the modification of flavonols and cyanidin but not flavan-3-

ols (UGT78L1) and therefore likely does not function in the synthesis of condensed tannins (Veljanovski and Constabel, 2013). The general phenylpropanoid pathway genes, C4H and 4CL had increased expression which could also contribute to the accumulation of condensed tannins. The results of the microarray analysis are supported by *in silico* co-expression analysis that showed that MYB115 is co-expressed with flavonoid biosynthetic genes (Table 2-2).

Three independently transformed 717-1-B4 transgenic lines and five independently transformed 353-38 transgenic lines overexpressing MYB115 were generated. While all transgenic lines showed increased expression of PtMYB115 compared to WT plants, three of the five 353-38 transgenic lines did not accumulate significantly higher tannins compared to WT plants. These low tannin lines did not show activation of the flavonoid biosynthetic genes DFR2 or F3’5’H1 but did show activation of the tannin specific gene ANR1. DFR2 has been shown to function specifically in the synthesis of condensed tannins and not anthocyanins, through overexpression analysis (Huang et al., 2012). Since DFR2 showed correlation with condensed tannin levels while the tannin specific gene ANR1 did not, DFR2 could be a rate limiting step in condensed tannin synthesis.

Although not all plants showed a high tannin phenotype when greenhouse grown plants were analyzed, when tested *in vitro* with a DMACA stain, all lines showed increased accumulation of tannins compared to WT plants. A similar phenomenon has been observed in MYB134 overexpressing plants (V. Walker and C.P. Constabel, unpublished). Some MYB134 overexpressing plants that show higher expression of MYB134 lack a high tannin phenotype. The regulatory mechanism leading to this is unclear. Future studies could investigate global changes in gene expression in high expressing lines that lack a high tannin phenotype in order to identify potential transcriptional repressors.

The first committed step of the flavonoid pathway is the formation of chalcone from coumaroyl-CoA and three molecules of malonyl-CoA by chalcone synthase. There are six annotated chalcone synthase genes in the *Populus* genome (Tsai et al., 2006).

Here, four of the six CHS isoforms (CHS1/3/4/6) were upregulated in 353-38 MYB115 transgenic plants as shown by microarray analysis. Surprisingly, CHS6 is characterized as a root specific isoform through expression analysis (Tsai et al., 2006). Furthermore, CHS1 and CHS3 are shown to have higher expression in root tips than in leaves (Tsai et al., 2006). CHS4 is the most highly upregulated isoform in MYB115 overexpressing plants and has been shown to be is most highly expressed in leaves and stems (Sun et al., 2011; Tsai et al., 2011). In a study by Sun et al. (2011), CHS4 was shown to be wound-inducible. This supports the hypothesis that MYB115 may regulate the flavonoid pathway in response to wounding (section 2.4.6.).

Two potential condensed tannin transporters show enhanced expression in MYB115 overexpressing plants. In Arabidopsis, a multidrug and toxic efflux (MATE) type transporter, AtTT12, has been characterized to transport flavonoids, primarily flavan-3-ols and anthocyanins into the vacuole (Marinova et al., 2007). The most highly upregulated MATE transporter in MYB115 overexpressing plants showed high sequence similarity to AtTT12. Several MATE family transporters showed reduced expression in MYB115 overexpressing plants compared to WT plants. However, none were closely related to AtTT12. Two showed close similarity to the flavonoid transporter, AtFFT. The substrate specificity of AtFFT has not been determined but early evidence suggests that is likely a transporter of glycosylated flavonols (Thompson et al., 2010). Potentially, the downregulated transporters function to transport other flavonoids not related to condensed tannins.

Overall the results of this study indicate that MYB115 is a key regulator of condensed tannin synthesis in poplar. MYB115 overexpressing plants show a high tannin phenotype that is a result of enhanced expression of all flavonoid biosynthetic genes necessary for tannin synthesis. Furthermore, putative enzymes involved in the transport of condensed tannin precursors were also upregulated. The data from this study also suggest that enhanced accumulation of condensed tannins is dependent on expression of DFR2 since tannin concentrations correlated with expression of DFR2 but not the tannin specific gene, ANR1.

2.4.3 MYB115 overexpression leads to enhanced expression of F3'5'H1 and accumulation of dihydromyricetin

In high tannin 353-38 transgenic plants, one compound accumulated at particularly high levels compared to WT plants. This compound was tentatively identified as dihydromyricetin based on its fragmentation pattern by LC/MS analysis (M. Reichelt and C.P. Constabel, unpublished data). Dihydromyricetin is a precursor to the flavonol, myricetin, and to the leucoanthocyanidin, leucodelphinidin, a precursor of both anthocyanins and condensed tannins. Myricetin was not detected by HPLC but was detected by LC/MS to be higher, as a minor component, in MYB115 overexpressing plants. However, flavonol synthase, was not induced in MYB115 overexpressing plants suggesting MYB115 is not directly regulating myricetin synthesis.

The most highly upregulated gene identified through microarray analysis was a flavonoid 3' 5' hydroxylase (F3'5'H1; Potri.009G069100). F3'5'H is required for the synthesis of dihydromyricetin from dihydrokaempferol. F3'5'H, along with F3'H, determine the hydroxylation pattern of the B ring of flavonoids (Seitz et al., 2006). There are two copies of F3'5'H in the poplar genome and F3'5'H2 was also found to be induced here, but only two fold (Tsai et al., 2006). In 353-38 plants, F3'5'H1 only showed higher expression in plants that showed higher accumulation of dihydromyricetin. This implicates that F3'5'H1 functions in the synthesis of dihydromyricetin and explains the accumulation of dihydromyricetin in MYB115 overexpressing plants.

Importantly, dihydromyricetin is a precursor to leucodelphinidin, a leucoanthocyanidin. Leucodelphinidin is a precursor to anthocyanins and to gallocatechin, a trihydroxylated flavan-3-ol precursor of condensed tannins. Here, gallocatechin and gallocatechin-catechin dimers showed enhanced accumulation in high tannin transgenic 353-38 lines. Furthermore, like dihydromyricetin, gallocatechin-catechin dimers were below the detection limit of LC/MS in WT plants. This increase in trihydroxylated flavan-3-ols is likely a result of the increase in dihydromyricetin and higher expression of F3'5'H1.

An increase in the availability trihydroxlated flavan-3-ols is predicted to alter the subunit composition of condensed tannins. Although not tested here, this could lead to targeted modification of condensed tannin structure. Studies have shown that subunit composition can influence the efficacy of condensed tannins as anti-herbivore compounds (Ayres et al., 1997). Likely, the increased expression of F3’5’H1 in MYB115 overexpressing plants will alter condensed tannin structure and thus their antiherbivore activity.

Dihydromyricetin did not accumulate at higher levels in 717-I-B4 plants overexpressing MYB115 compared to wild-type control plants. Expression of F3’5’H was significantly higher in one line of 717 plants (Line 1) but this did not lead to an increased accumulation of dihydromyricetin. Furthermore, relative expression was lower than in 353-38 high tannin lines that accumulated dihydromyricetin. Potentially, dihydromyricetin is being utilized as a substrate and not accumulating in 717-I-B4 plants due to this lower relative expression of F3’5’H. Further analysis of subunit composition of condensed tannins in 717-I-B4 transgenic plants overexpression MYB115 are necessary to verify this.

DFR catalyzes the formation of leucoanthocyanidins from dihydroflavonols. It is unclear why dihydromyricetin is accumulating at high concentrations in 353-38 MYB115 overexpressing plants and not being utilized as a substrate for DFR. This could be explained by the substrate specificities of PtDFR1 and DFR2 which catalyze the next step in condensed tannin synthesis (Figure 1-2). The substrate specificity of PtDFR1 and 2 have not been determined as of yet; however, a 26 amino acid region in DFR is postulated to determine substrate specificity (Beld et al., 1989; Figure 2-21). More recent studies in petunia and *Medicago* suggest that a single amino acid change can alter the substrate preference (Johnson et al., 2001; Xie et al., 2004). An asparagine at position 133 has been demonstrated to confer specificity for dihydrokaemferol over dihydroquercetin while an aspartic acid at this position can confer preference for dihydroquercetin and dihydromyricetin (Johnson et al., 2001; Xie et al., 2004). Both poplar DFR isozymes have an aspartic acid at this position suggesting that both might

preferentially utilize dihydroquercetin or dihydromyricetin as substrates over dihydrokaemferol. This is consistent with a study by Peters and Constabel (2002) that showed PtDFR1 can utilize dihydroquercetin as a substrate. However, this single amino acid change likely does not fully explain the substrate specificity of DFR. Further studies using recombinant enzymes are necessary to show differences in substrate specificity.

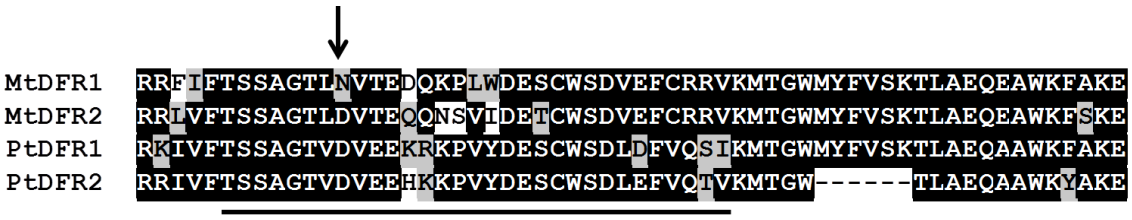

Figure 2-21. Alignment of DFR partial sequences from *Medicago trunculata* and *Populus trichocarapa.*
The underlined region indicates the proposed sequence that determines substrate specificity (Beld et al., 1986). The arrow indicates the Asn or Asp residue that influences specificity for dihydrokaemferol over dihydroquercetin or dihydromyricetin. Identical amino acids are shaded black and similar amino acids are shaded grey.

2.4.4 MYB115 overexpression leads to a reduction in phenolic glycosides

MYB134 overexpressing high tannin plants were previously shown to have a strong reduction in phenolic glycosides (Mellway et al., 2009). Here, a high tannin phenotype in MYB115 overexpressing plants also corresponded with a reduction in phenolic glycosides. Both salicin based phenolic glycosides as well as grandidentatin (*p*-coumaroyl-glucopyranoside) derivatives were inversely correlated with condensed tannins. Mellway et al. (2009) propose three potential reasons for the reduction in phenolic glycosides in MYB134 overexpressing plants: 1) the reduction in phenolic glycosides is due to competition for common co-factors by other MYB regulatory factors and not directly caused by the regulation by MYB134, 2) phenolic glycosides and condensed tannins could share common precursors and there is a flux towards condensed tannin synthesis in MYB134 overexpressing plants, or 3) MYB134 is directly regulating phenolic glycoside synthesis at the transcriptional level.

The pathway leading to phenolic glycosides is unknown. However, in a study by Babst et al. (2010), cinnamic acid was shown to be a precursor to salicin and salicortin by feeding leaf disks stable-isotope labelled cinnamic acid. Cinnamic acid is a central

compound in the general phenylpropanoid pathway and a pre-cursor to many classes of phenylpropanoids including lignins, flavonoids, and hydroxycinnamic acids (Vogt, 2010). Cinnamic acid is synthesized from phenylalanine via deamination by phenylalanine ammonia lyase (PAL) (Vogt, 2010). Microarray analysis of 353-38 MYB115 overexpressing plants that showed a reduction in phenolic glycosides did not show a significant change to the expression of any of the PAL genes. Therefore, a decrease in phenolic glycosides is likely not due to changes in the regulation of cinnamic acid biosynthesis. By contrast, the other enzymes in the general phenylpropanoid pathway, trans-cinnamate 4-monooxygenase (C4H) and 4-coumaroyl:CoA-ligase (4CL) leading to the synthesis of *p*-coumaric acid and *p*-coumaroyl CoA, respectively, showed higher expression in MYB115 transgenic plants. *p*-coumaric acid and *p*-coumaroyl CoA are general precursors to many phenylpropanoid classes. The reduction of phenolic glycosides could be a result of cinnamic acid preferentially being converted to p-coumaric acid by C4H. This would imply that phenolic glycosides and condensed tannins share cinnamate as a common precursor. However, this would not explain the reduction of grandidentatin derivatives which are likely derived from *p*-coumaric acid.

MYB115 may also be actively leading to a reduction in phenolic glycosides through negative regulation of biosynthetic genes involved in the synthesis of phenolic glycosides. Because no genes in the phenolic glycoside pathway are known to date, this cannot be tested directly. However, in a proposed pathway by Babst et al. (2010), benzyl-CoA is suggested to be involved in phenolic glycoside biosynthesis. BAHD acyltransferases require a CoA-thioester molecule as one of their substrates and therefore may be involved in phenolic glycoside biosynthesis (D'Auria, 2006; R. Chedgy and C.P. Constabel, unpublished). Two putative BAHD acyltransferases pACT47 and pACT49 (Potri.013G074500 and Potri.019G043600) were significantly downregulated in MYB115 overexpressing plants (Table 2-3). These acyltransferases are also downregulated in MYB134 transgenic plants (Mellway, 2009). This is consistent with the idea that both poplar regulators of condensed tannin synthesis down regulate phenolic glycoside synthesis at the transcriptional level.

Transcriptional regulators specific to phenolic glycoside biosynthesis have not yet been identified. Potentially, phenolic glycoside biosynthesis is transcriptionally controlled by R2R3 MYB transcription factors that share common co-factors, such as bHLH co-factors, with regulators of condensed tannin biosynthesis. By overexpressing MYB115, these common co-factors could be competitively bound leading to loss of positive regulation of phenolic glycoside biosynthesis. However, this does not explain the lower expression of putative phenolic glycoside pathway genes, pACT47 and pACT49, in MYB115 overexpressing plants. Identification of phenolic glycoside pathway genes as well as regulators of phenolic glycoside biosynthesis is necessary in order to determine how MYB115 overexpression is leading to a reduction in phenolic glycoside accumulation.

2.4.5 MYB115 overexpression leads to changes in concentrations of hydroxycinnamic acid derivatives

Hydroxcinnamic acids are a highly diverse class of phenylpropanoids implicated in plant defense that are synthesized by hydroxylation, and in some cases O-methylation, of cinnamic acid. In 717-1-B4 MYB115 overexpressing plants, levels of hydroxycinnamic acid derivatives were increased as well as chlorogenic acid derivatives. Chlorogenic acid is synthesized via 3' hydroxylation of esters of quinic acids by a cytochrome P450 protein, *p*-coumaroyl-shikimate/quinate 3'-hydroxylase (CYP98A3) (Schoch et al., 2001). Microarray analysis of 353-38 MYB115 overexpressing plants showed a 2.6 and 2.4 fold induction of two putative *p*-coumaroyl-shikimate 3'-hydroxylases, C3H3 and CYP98A23, which could be involved in chlorogenic acid biosynthesis. However, unlike 717-1-B4 transgenic lines, the 353-38 MYB115 overexpressing line used for microarray analysis did not show significantly higher accumulation of chlorogenic acid compared to wild-type plants and so the mechanism by which chlorogenic acid and related hydroxycinnamic acid derivatives are upregulated is unclear.

Importantly, concentrations of chlorogenic acid derivatives and coumaroyl quinic acid derivatives were higher in 353-38 MYB115 overexpressing plants regardless of the tannin content. This was not observed for any other class of compounds which correlated with tannin levels. Potentially, the inhibition of MYB115 activity is specific for flavonoid and phenolic glycoside biosynthesis. Another possible reason for the increases seen in hydroxycinnamic acid concentrations could be an artifact of infection with *Agrobacterium* since wild-type control plants were not infected. Analysis of wild-type plants compared to infected but not transformed plants is necessary to determine if changes in biochemistry are due to overexpression of the transgene alone. One hydroxycinnamic acid derivative (HCA-3) did correlate with tannin levels and was reduced in 353-38 MYB115 overexpressing plants high tannin plants. Potentially, the reduction of this hydroxycinnamic acid in high tannin plants was a result of preferential use of common precursors, such as cinnamic acid, for synthesis of flavonoids. However, with the exception of HCA-3, the overall trend was that HCA derivatives correlated with expression of MYB115 suggesting that MYB115 may be regulating their synthesis.

2.4.2 MYB115 expression leads to changes in expression of other regulatory factors

Microarray analysis of MYB115 overexpressing plants showed induced expression of many biosynthetic genes compared to WT plants. In addition to these biosynthetic genes, a number of regulatory factors showed differential expression. These included putative positive and negative MYB regulators of condensed tannin synthesis as well as putative co-factors of MYB transcription factors. In addition to MYB transcription factors, other classes of transcription factors showed reduced expression in MYB115 overexpressing plants compared to WT plants.

MYB134 showed the highest fold change expression of all regulatory factors as analyzed by microarray. However, it does not appear that MYB115 is strongly influencing its expression since when all 717-1-B4 and 353-38 transgenic lines are analyzed, MYB115 expression does not correlate with MYB134 expression. Furthermore, MYB134 did not show higher expression in any of the 717-I-B4 transgenic lines. This

supports the results of promoter activation analysis (Chapter 3) which shows that MYB115 does not activate the MYB134 promoter. However, likely the higher expression of MYB134 in 353-38 MYB115 overexpressing Lines 4 and 5 is contributing to the strong, high tannin phenotype.

MYB182 is a putative negative regulator of condensed tannin synthesis (K. Yoshida and C.P. Constabel, unpublished). MYB182 showed 3.7 fold higher expression in MYB115 Line 5 compared to WT plants through microarray analysis (Table 2-4). Only 717-1-B4 Line 2 and 353-38 Line 5 showed a fold change in expression of MYB182 greater than 2 when analyzed by qPCR. Expression of MYB182 did not correlate with MYB115. However, in 353-38 plants MYB182 strongly correlated with MYB134. This could suggest that MYB134 expression has a stronger influence than MYB115 on expression of MYB182.

Two genes encoding for WDR type factors (Potri.006G209000 and Potri.016G075800) and one gene encoding for a bHLH type factor (Potri.005G208600) showed higher expression in MYB115 overexpressing plants compared to WT plants. Potentially, these act as co-factors for MYB115. As shown in Arabidopsis, the MYB regulator of condensed tannin synthesis, AtTT2, forms a ternary complex with a bHLH factor (AtTT8) and a WDR factor (AtTTG1) to regulate transcription (Baudry et al., 2004). Here, a TT8 like bHLH factor was induced. The bHLH type factor (called here PtbHLH131) was tested as a co-factor with MYB115 in transient promoter activation assays and does function to co-activate the ANR1 promoter (Figure 3-7). Two WDR factors that show sequence similarity to AtLWD1, also known as AtAN11, were induced in MYB115 overexpressing plants. One of these (Potri.016G075800) was tested as a co-factor for MYB115 in a transient promoter activation assay but did not influence the level of activation (Figures 3-1, 3-4). LWD1 is known to act as a regulator of flowering time (Wu et al., 2008). AtLWD1 is also named AtAN11 due to its similarity to the regulator or anthocyanin synthesis in petunia (de Vetten et al., 1997). In petunia, AN11 acts along with a MYB transcription factor, AN2, and a bHLH factor, AN1, to regulate anthocyanin biosynthesis (de Vetten et al., 1997). However, it has not been shown if AN11 directly

binds with AN2 or AN1. Potentially, the WDR factors identified here function as co-factors for MYB115 to regulate the flavonoid pathway.

2.4.6. MYB115 expression is induced by wounding

PtMYB134 has been previously found to be induced by wounding, high-intensity light, UV light, and pathogen stress. In order to determine if PtMYB115 is also induced by stress, leaves of 353-38 plants were wounded by crushing. PtMYB115 was induced by wounding along with PtMYB134 and flavonoid biosynthetic genes (Figure 2-1). PtMYB134 is predicted to act upstream of PtMYB115 (see Chapter 3). Because of this, activation of PtMYB115 may be a result of the increased expression of PtMYB134 in response to wounding.

The PtMYB134 homolog in persimmon, DkMYB2, shows induction one hour after wounding along with many of the flavonoid biosynthetic genes leading to condensed tannin synthesis (Akagi et al., 2010). The PtMYB115 homolog in persimmon, DkMYB4, does not show significant induction until three hours after treatment. This suggests that DkMYB2 is upstream of DkMYB4 and supports the model of PtMYB134 regulating the expression of PtMYB115 in response to wounding. Future experiments will test for induction of PtMYB115 by UV-B light stress. Ideally, a time-course experiment with multiple time points would help to elucidate if PtMYB134 expression precedes expression of PtMYB115 in response to stress.

2.4.7 Conclusions

The results of this chapter demonstrate a role for MYB115 in the synthesis of condensed tannins. MYB115 overexpressing poplar showed high accumulation of condensed tannins and increased expression of tannin biosynthetic genes compared to WT plants. Condensed tannin precursors also accumulated at higher concentrations in transgenic poplar. A dihydroflavonol, dihydromyricetin, showed a dramatic increase in 353-38 MYB115 overexpressing plants which is likely due to a strong increase in expression of F3’5’H1. As the structure of condensed tannins is partially dependent on the hydroxylation pattern of its subunits, likely the increased expression of F3’5’H1

altered condensed tannin structure. Other less dramatic changes in phenolic metabolism included reductions in concentrations of phenolic glycosides and increases in concentrations of hydroxycinnamic acids.

3. Chapter Three: Analysis of transcription activation by PtMYB115 using the dual-luciferase promoter activation assay

3.1 Introduction

R2R3 MYB transcription factors are key regulators of the flavonoid pathway in higher plants. This transcription factor family is greatly expanded in plants compared to animals with 125 annotated members in the Arabidopsis genome and 192 in poplar (Dubos et al., 2010). R2R3 MYB transcription factors have been shown to regulate many branches of the flavonoid pathway including flavonols, anthocyanins and proanthocyanidins (Feller et al., 2011). R2R3 MYB regulators of condensed tannin synthesis have been characterized in Medicago (Verdier et al., 2012), Arabidopsis (Nesi et al., 2001), grape (Bogs et al., 2007; Terrier et al., 2009), persimmon (Akagi et al., 2009, 2010), and poplar (Mellway et al., 2009). These transcription factors have been characterized using analysis of transgenic plants (either mutants or overexpressors) and also through promoter activation assays. With the exception of the condensed tannin MYB regulator in Arabidopsis, AtTT2, all these transcription factors show regulation of representative genes in both the early and late flavonoid biosynthetic pathway along with condensed tannin specific genes. By contrast, AtTT2 has been shown to regulate only the late flavonoid biosynthetic genes leading to condensed tannin synthesis (Nesi et al., 2001).

Many R2R3 MYB transcription factors have been shown to form a complex with bHLH and WDR factors (Ramsay and Glover, 2005). Such ternary complexes are involved in cellular differentiation and plant metabolism. The WDR factor is considered to function to stabilize the interaction between bHLH and MYB transcription factors while the MYB and bHLH proteins specifically bind to the promoters of genes. One well studied example is anthocyanin synthesis in petunia, which is regulated by a set of MYB, bHLH and WDR factors (Albert et al., 2011; de Vetten et al., 1997). In Arabidopsis seed coats, condensed tannin synthesis is regulated by a MYB-bHLH-WDR complex consisting

of TT2, TT8 (bHLH) and TTG1 (WDR). As shown by a yeast-two hybrid assay, the WDR factor TTG1 binds to the MYB factor TT2 and the bHLH factor TT8 to form a ternary complex that regulates condensed tannin synthesis (Baudry et al., 2004).

The bHLH factors are less specific in their regulatory activity than MYB factors. For example, a bHLH factor in maize, ZmLc, and a bHLH factor in Arabidopsis, AtGL3, can regulate both anthocyanin synthesis and trichome development (Ramsay and Glover, 2005). This contrasts with most MYB type factors that show a regulatory role in only one pathway (Hichri et al., 2010). Furthermore, studies assaying the activity of MYB transcription factors *in vitro* have used bHLH co-factors isolated from different plants. For example, Akagi et al. (2010), studied the direct targets of persimmon MYB transcription factors, DkMYB2 and DkMYB4, a bHLH co-factor from Arabidopsis, AtEGL3, was shown to be necessary for activation of genes involved in condensed tannin synthesis. AtEGL3 was also used as a co-factor for the grape regulator of condensed tannin synthesis in transient promoter activation analysis (Bogs et al., 2007).

Transcription factors regulate the expression of genes by binding to promoter regions to activate or repress transcription. There are several methods to test the target genes of transcription factors. Most of these assay for *in vitro* binding activity such as gel-shift assays or yeast-one hybrid assays. The dual-luciferase promoter activation assay demonstrates activation or repression by transcription factors following transient transformation of plant tissues. The dual-luciferase assay shows transcription factor binding by overexpressing a transcription factor along with a promoter fused to a gene coding for a luciferase enzyme. By measuring the luciferase enzyme activity, one can estimate the level of promoter activation.

In this assay, poplar cells or Arabidopsis leaves are bombarded with gold beads coated with plasmid DNA. Here, three vectors are used for each bombardment. The first two vectors are overexpression constructs carrying the transcription factor of interest and its necessary co-factor. The third vector is the reporter vector which carries the promoter of interest fused to the gene coding for the firefly luciferase enzyme. This reporter construct also carries an internal control of the *Renilla* luciferase enzyme under

the control of the 35S promoter. Thus, the firefly enzyme activity can be normalized against the activity of the *Renilla* enzyme. This controls for differences in transformation efficiency between bombardments. If the transcription factor is able to activate the promoter, the amount of enzyme produced will be detected upon the addition of the corresponding substrate. A unique substrate is used for each luciferase enzyme allowing measurement of both enzymes from the same extract.

The dual-luciferase promoter activation assay has been used previously to test the activity of condensed tannin regulators. Yoshida et al. (2010) identified an AtTT2 homolog in *Lotus japonicas* and showed it could activate the promoters of DFR, ANS and ANR. They also found that the addition of bHLH and WDR co-factors increased the activation activity. LjTT2 was also directly shown to interact with both bHLH and WDR co-factors using a yeast-two hybrid assay (Yoshida et al., 2008). Likewise, in a study by Akagi et al. (2010), the R2R3 MYB transcription factors in persimmon, DkMYB2 and DkMYB4 could activate the promoters of DkANR and DkLAR when co-bombarded with a bHLH factor. These studies not only demonstrate that the dual-luciferase promoter activation assay is a useful tool for testing the activity of transcription factors but they also demonstrate the importance of transcriptional co-factors.

The objectives of this study were to further demonstrate the regulation by MYB115 of key biosynthetic genes of the flavonoid pathway leading to condensed tannin synthesis. The predicted targets of MYB115 were tested using the dual luciferase promoter activation assay. Two promoters of tannin specific enzymes, ANR1 and LAR3, were tested for activation by MYB115 as well as the other poplar condensed tannin regulator, MYB134. The data showed both MYB115 and MYB134 could activate the ANR1 promoter while only MYB134 could activate the LAR3 promoter.

Since MYB115 shows higher expression in MYB134 overexpressing plants, it was predicted that the MYB115 promoter is a target of MYB134. Here I found that MYB134 activated the promoter of MYB115. MYB115 could activate its own promoter but neither MYB115 nor MYB134 could activate the MYB134 promoter. Together, the results of these experiments further suggest that both MYB115 and MYB134 regulate at

least some genes in the condensed tannin pathway and that MYB115 is a downstream target of MYB134.

3.2 Methods

3.2.1 Particle bombardment for transient plant transformation

Arabidopsis thaliana (Columbia-1) plants were grown in incubators at 22 ^{o}C with a day-length of 16-hr light/8-hr dark. Leaves were excised after 6-8 weeks of growth. Particle bombardment of Arabidopsis leaves was identical as for poplar suspension cells.

Populus trichocarpa x deltoides (H11-11) cells were available in the Constabel lab. Cells were maintained in Murashige and Skoog (MS) liquid media (Murashige and Skoog, 1962) and sub-cultured every 8 – 10 days. Cells were bombarded after 7 days of growth. Gold beads were coated with 250 ng of each plasmid as follows. 250 ng of each vector (activator, co-factor, and reporter vectors) were added to 25 μL of 0.6 μm gold or 0.7 μm tungsten particles (Bio-Rad, Hercules, CA, USA) under constant vortexing, followed by 25 μL of 2.5 M CaCl2 and 10 μL of 0.1 M spermidine. Samples were further vortexed for an additional 10 minutes at 4 ^{o}C. The beads were washed with 70 % and 100 % ethanol and resuspended in 20 μL of 100% ethanol. The beads were pipetted onto presterilized flying disks (Bio-Rad, Hercules, CA, USA) and allowed to air dry. Rupture disks (900 psi) were used for the bombardments in a Bio-Rad model PDS-1000/He Biolistic Particle Delivery System.

Prior to bombardment, poplar suspension cultures were allowed to settle for 15 minutes and liquid media was removed from cells. Poplar suspension cell aliquots were taken from four independently grown cultures and pipetted onto four individual filter papers for each bombardment. The filter paper with the cells was placed on MS with 0.5 M mannitol solid media and the cells were incubated for one hour prior to bombardment. All four replicates were bombarded simultaneously 13 cm from the rupture disc. Following bombardment the cells were maintained in darkness for 48 hours before assaying for luciferase activity.

3.2.3 Luciferase assay

pMDC32/pGREENSK62 overexpression vectors (Curtis and Grossniklaus, 2003; Hellens et al., 2005) and the pGREEN800 reporter vector (Hellens et al., 2005) were used for the dual-luciferase reporter assay. pGREENSK62 and pGREEN800LUC were obtained from Dr. Roger P. Hellens (HortResearch, Palmerston North, New Zealand). Construction of vectors was perfomed by Dr. Andreas Gesell with the exception of pMDC32:PtGL3 which was cloned by Dr. Kazuko Yoshida and pGREEN800LUC:PtMYB134 which was cloned for the purposes of this thesis (Zifkin et al., 2012; A. Gesell and C.P. Constabel, unpublished). Successful insertion of promoter sequences and coding sequences was confirmed by sequencing within the vector. For all constructs except for pGREEN800LUC:PtMYB134 and pMDC32:PtGL3, sequences were cloned into pGEM T – Easy (Promega, Madison, WI, USA) then subcloned into the *Not1* sites of pGREEN800LUC or pGREENSK62. For pMDC32, sequences were cloned into the Not1 sites of pENTR then recombined into pMDC32 vectors using the Gateway system (LR Clonase, Invitrogen, Carlsbad, CA, USA). Promoter regions were between 1 kb and 2 kb upstream of the gene of interest (see Table 3-1). For PtGL31 (Potri.001G103600), the coding sequence was cloned directly into *Kpn1* and *Spe1* restriction enzyme sites of pMDC32. For the MYB134 promoter, a 1.2 kb region upstream of PtMYB134 was PCR amplified from DNA extracted from *P. trichcarpra* (Nisqually-1) leaf tissue and cloned into *Xho1* and *BamH1* restriction sites in the multi-cloning site of pGREEN800LUC using the In Fusion Cloning Kit (Clontech, Mountain View, CA, USA).

Table 3-1. Primers for cloning regulatory factors and promoter sequences for the dual-luciferase promoter activation assay.

	Forward Primer	Reverse Primer	Promoter length (kb)
PtMYB115	ATGGGAAGGGCTCCTTG	ATTGTCATACCAGCAGTGAC	NA
PtMYB134	ATGGGAAGGAGTCCATG	TCATGGCCACTCTTCAG	
PtbHLH131	AATGGCTACCCCGCCTAG	GTAAGAGTCTCTCTGCTCAATC	NA
PtbHLH79	ATGGCTACTAAGCTCCACAACC	TCAACATTTCCCAGCAACTCTT	NA
AtGL3	CAATGGCTACCGGACA	CTTCAACAGATCCATGCA	NA
PtWD40-1	GGGGGATCCATGGAAAGCTCTTCACAGAAAGG	GGGGAGCTCCTAAACCTTCAAGAGTTGTA	
PtANR1 promoter	GGTCACTGGAGGCTTAC	CAACTGGGATGCCATGC	1.1
PtLAR3 promoter	CATGCATGCCTGTTATAAG	ACACCCTCTTATCCGGA	1.9
PtMYB115 promoter	CAAACTCATGACCCGAGTC	GCTTATGTATATTTTTTCCCTCTC	1.1
PtMYB134 promoter	AAGACTCTGAACCACATCAC	CATCTCTCTCTACGATCACC	1.2

Promoter activation was assayed using the Dual-Luciferase® Reporter (DLR™) Assay System (Promega Madison, WI, USA). 48 hours after bombardment, Arabidopsis leaves were ground in 500 uL lysis buffer for 15 seconds using a Precellys® 24 tissue homogenizer (Bertin Technologies, Aix-en-Provence, France) or H11-11 poplar suspension cells were collected from the filter paper and ground with a micro-pestle in 100 µL of lysis buffer in 1.5 mL microcentrifuge tubes. The ground leaf tissue or cells were incubated at 4 ^{o}C for 10 minutes then centrifuged for 10 minutes at 15 871 rcf. 10 µL of lysate was added to 40 µL of firefly luciferase substrate (chemiluminescence buffer). Luminescence was measured in a 96-well plate using a Perkin Elmer Victor™ X5 multilabel plate reader. 40 µL of the *Renilla* substrate (Stop and Glo^{TM} buffer) was then added and a second chemiluminscence reading was taken. Three of the four replicates with the highest reading for the *Renilla* substrate were used for further analysis, a higher reading indicating a higher efficiency of transformation.

Statistical tests were performed in R (http://www.R-project.org). To perform analysis of variance the aov function from the *stats* version 2.15.0 package was used. To

perform Tukey's test for honest significant difference, the TukeyHSD function from the *stats* version 2.15.0 package was used.

3.3 Results

3.3.1 Activation of the ANR promoter by MYB115 and MYB134

Activation of the promoters of key flavonoid biosynthetic enzymes leading to the synthesis of tannins by MYB115 was tested in a transient transformation system. First, the anthocyanidin reductase (PtANR1) promoter was tested for activation. There are two ANR genes in the poplar genome (Tsai et al., 2006), but ANR1 was chosen because it showed higher activation than ANR2 in high tannin MYB134 overexpressing plants (Mellway, 2009). ANR1 also shows a higher fold change in MYB115 overexpressing plants. Furthermore, PtANR1 has been shown to function in the biosynthesis of condensed tannins through analysis of overexpressing transgenic plants (Wang et al., 2013). ANR is a tannin specific enzyme and represents the final characterized step in condensed tannin biosynthesis (refer to Section 1.3). Poplar cells and Arabidopsis leaves were transformed with MYB115 or MYB134 along with a bHLH co-factor (AtGL3 or PtbHLH131) and the reporter vector. For negative controls, cells and leaves were also transformed with the co-factor alone.

In poplar cells, MYB115 activated the ANR1 promoter approximately 20 fold relative to the bHLH only control (Figure 3-1; Table 3-1). The experiment was repeated on three different days. A Tukey's honest significant difference test showed that MYB115 with AtGL3 resulted in significantly higher activation than either AtGL3 or MYB115 alone. Addition of a WDR co-factor did not significantly influence activation.

In order to verify the results of the bombardment in a different biological system and also to compare ANR1 activation by MYB115 with activation by MYB134, the experiment was repeated using Arabidopsis leaves. Transient transformation of Arabidopsis leaves showed an activation of the ANR1 promoter by MYB115 on average 17.2 times higher than the AtGL3 only control (Figure 3-2). This experiment was repeated on 6 different dates (n = 18) and includes the data from Figure 3-7. MYB134

had been previously shown to activate the ANR1 promoter (Zifkin et al., 2011). Here, MYB134 activated the ANR1 promoter on average 7.2 fold relative to controls (Figure 3-2). This experiment was repeated on two different dates. An analysis of variance with date as a random factor showed that the activation by MYB134 was significant relative to the control (n = 6, p = 0.002).

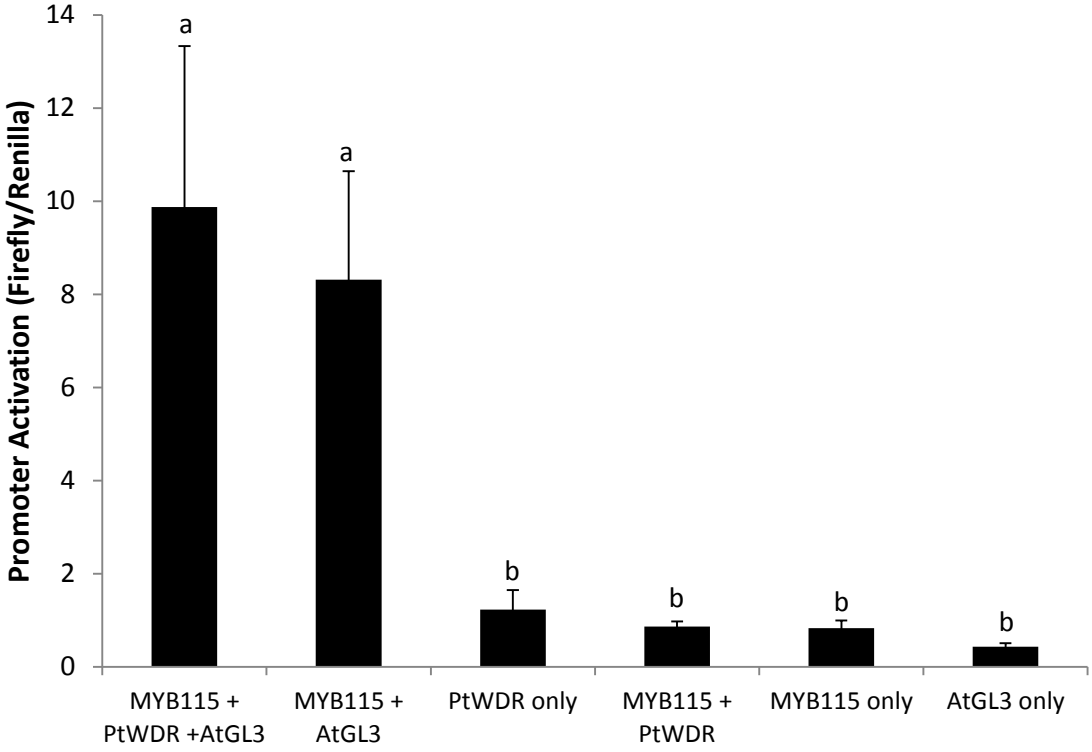

Figure 3-1. Dual-luciferase promoter assay following transient transformation of H11-11 poplar suspension cells showing activation of the anthocyanidin reductase promoter (ANR1) by MYB115.

AtGL3 (a bHLH protein) and PtWDR (PtWD40-1) were co-expressed with PtMYB115 or transformed alone with the reporter vector. Bombarded activators and co-factors are represented on the x-axis. Data is representative of three bombardments on three different days. Different letters indicate significant differences (Tukey's Honest Significant Differences, $p \leq 0.05$). Error bars indicate standard error (n = 9).

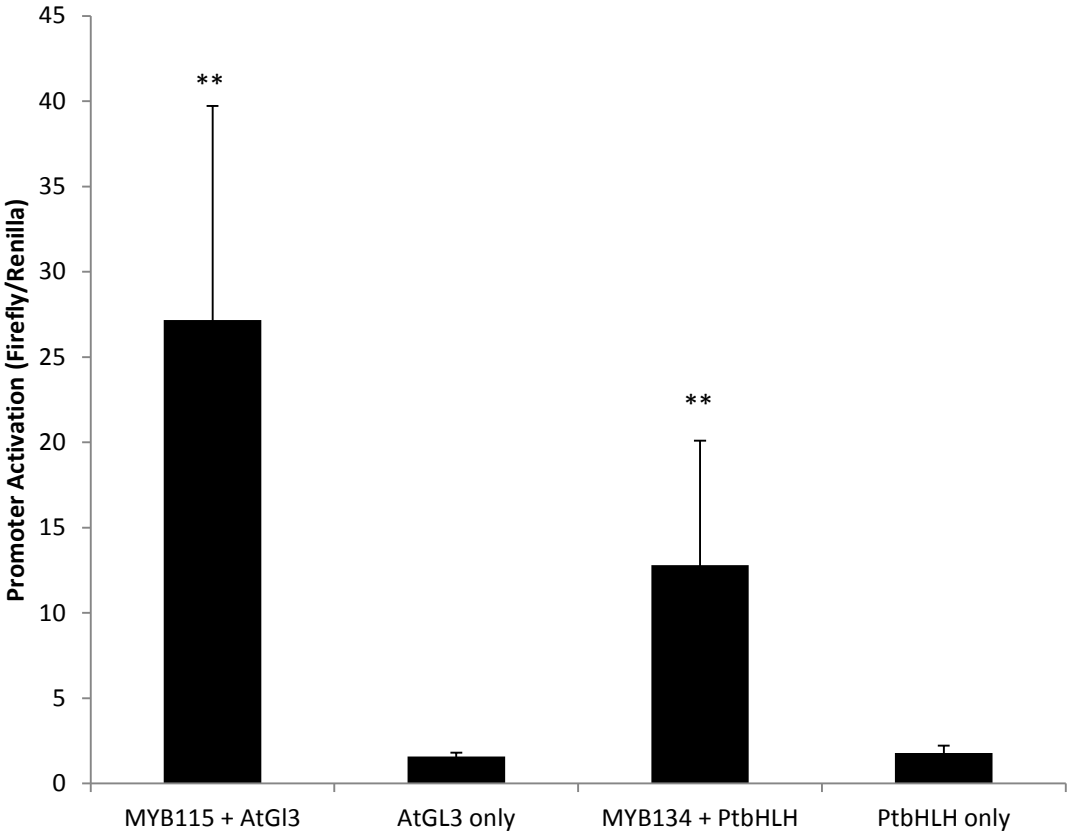

Figure 3-2. Dual-luciferase promoter assay following transient transformation of Arabidopsis leaves showing activation of the anthocyanidin reductase (ANR1) promoter by MYB134 and MYB115.

AtGL3 or PtbHLH (PtbHLH131) was co-expressed with PtMYB115 or PtMYB134. bHLH co-factors were expressed alone as a control. Bombarded activators and co-factors are represented on the x-axis. Error bars indicate standard error (for MYB115 and AtGL3, n = 18 [representing bombardments on 6 days]; for MYB134 and PtbHLH, n = 6 [representing bombardments on 2 days]). Asterisks indicate results of a one-way ANOVA with date as a residual factor ($p \leq 0.05$, *; $p \leq 0.01$, **).

3.3.2 Testing activation of the LAR3 promoter by MYB115 and MYB134

Activation of the leucoanthocyanidin reductase (LAR3) promoter by MYB115 and MYB134 was tested in transiently transformed Arabidopsis leaves. LAR is a key enzyme specific to the synthesis of flavan-3-ol precursors of condensed tannins. LAR3 was chosen because it shows a greater fold change in expression in MYB134 overexpressing plants compared to LAR1 or LAR2. While LAR1 shows a higher fold change in MYB115 overexpressing plants, the LAR3 promoter was chosen prior to analysis of MYB115 overexpressing plants. LAR3 was also chosen as it has been directly shown to function in the synthesis of condensed tannins (Yuan et al., 2012). MYB115 activated the LAR3 promoter on average 1.3 fold relative to controls (Figure 3-3). The experiment was repeated on four different dates. An analysis of variance with date as a random factor showed no significant activation of LAR3 by MYB115 with its co-factor compared to the co-factor only control (n = 12, p = 0.08). For comparison, MYB134 induced transcriptional activation of the LAR3 promoter on average 2.2 times higher than the controls (Figure 3-3). For these experiments, MYB134 was co-expressed with PtbHLH131 in Arabidopsis leaves on three different dates. An analysis of variance with date as a random factor showed a significant activation by MYB134 compared to the co-factor only control (n=9, p = 0.002). Thus, these results show that MYB134, but not MYB115, can activate the LAR3 promoter higher than controls. However, the activation is relatively modest relative to activation of the ANR1 promoter.

To test if addition of a WDR factor would result in increased activation of the LAR3 promoter by MYB115, the poplar ortholog of the TT2 WDR co-factor TTG1, PtWD40-1 (Potri.016G075800), was co-bombarded with MYB115 and AtGL3 into poplar cells (Figure 3-4). This also tested if the host tissue might explain the low level of LAR3 activation. Addition of the WDR factor with AtGL3 showed no significant difference in activation by MYB115 compared to addition of the AtGL3 alone (one-way ANOVA, n = 3, $p \geq 0.05$). This supports the results of the assay previously performed in Arabidopsis leaves (Figure 3-3) that showed no activation of LAR3 by MYB115 with AtGL3. However,

the assay was only repeated once and co-factor only controls were not included, as they were not expected to activate the LAR3 promoter from the previous experiment. These data confirm that the LAR3 promoter is not strongly activated by MYB115, and that addition of a WDR cofactor does not improve activation.

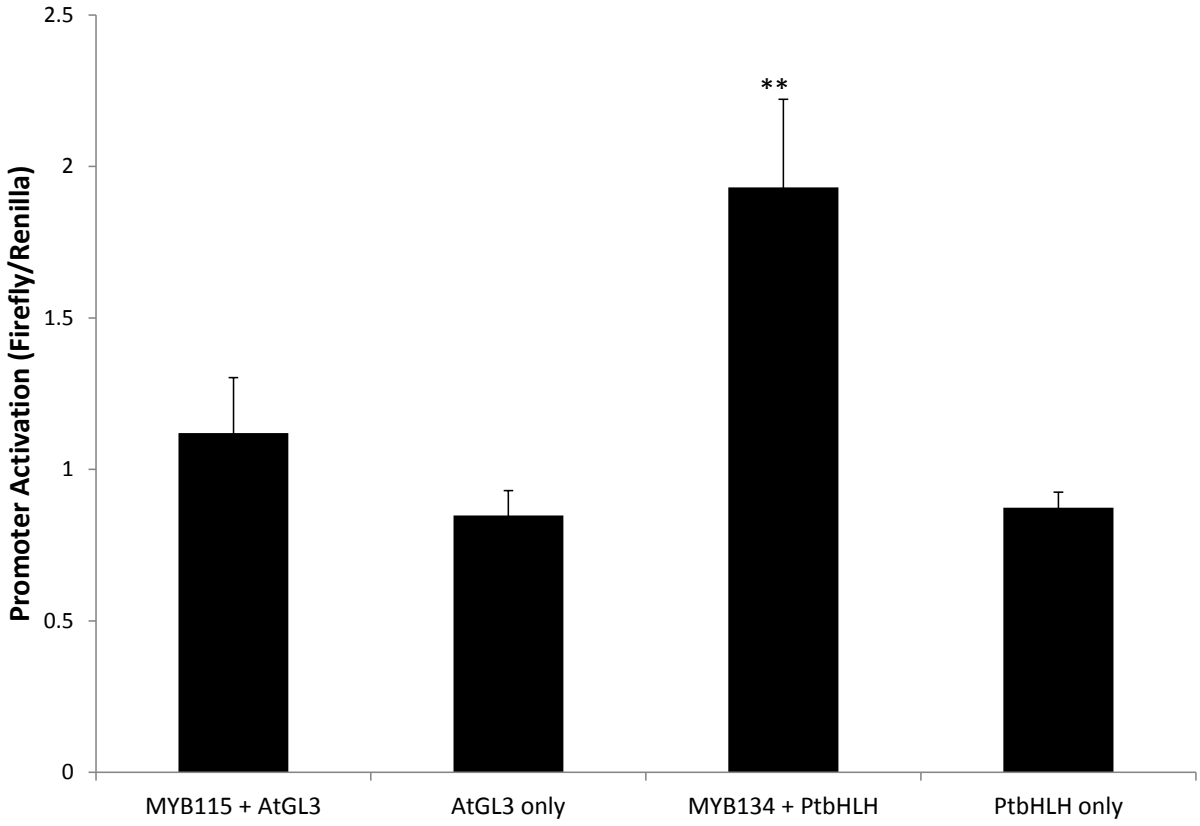

Figure 3-3. Dual-luciferase promoter assay following bombardment of Arabidopsis leaves to test for activation of the leucoanthocyanidin reductase (LAR3) promoter.
AtGL3 and PtbHLH (PtbHLH131) were co-expressed with PtMYB115 or PtMYB134 or expressed alone as a control. Bombarded activators and co-factors are represented on the x-axis. Error bars indicate standard error (for MYB115 and AtGL3, n = 12 [representing bombardments on four days], For MYB134 and PtbHLH, n=9 [representing bombardments on 3 days]). Asterisks indicate results of a one-way ANOVA with date as a residual factor ($p \leq 0.05$, *; $p \leq 0.01$, **).

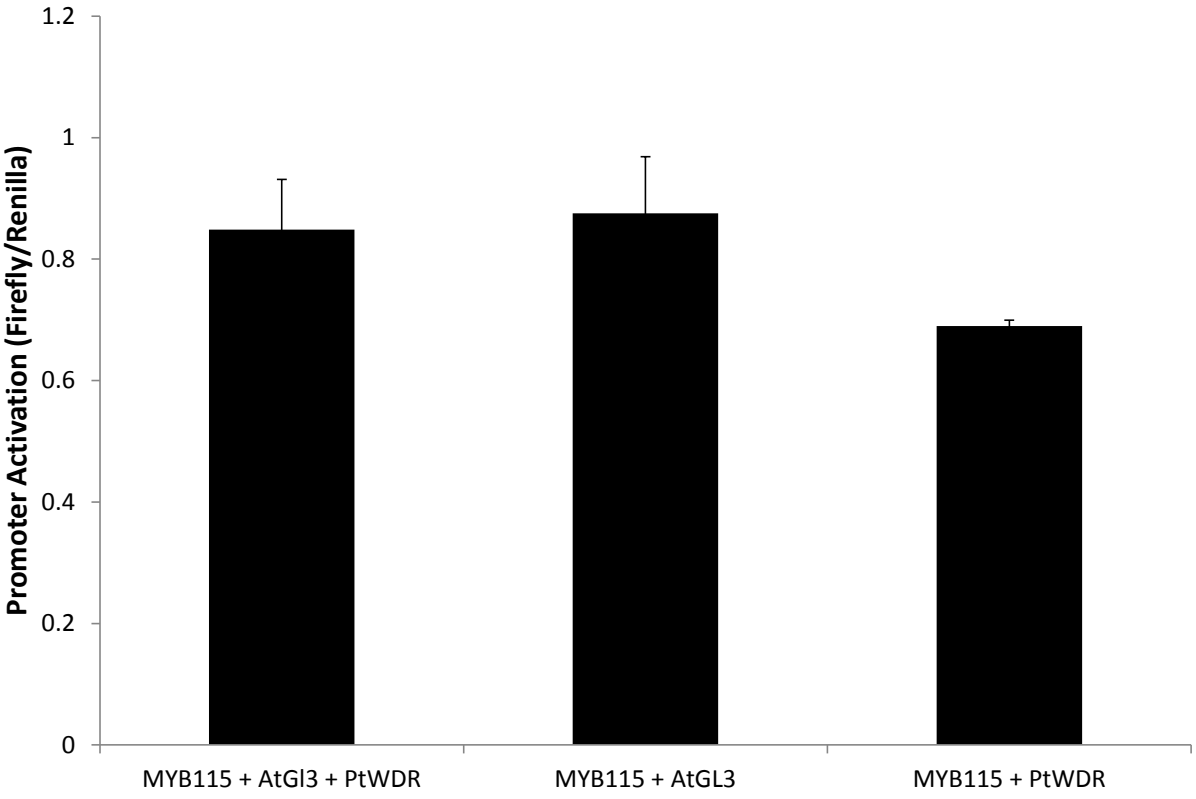

Figure 3-4. Activation of the leucoanthocyanidin reductase (LAR3) promoter using the dual-luciferase promoter assay in H11-11 poplar suspension cells to test the potential effect of a WDR factor.

AtGL3, PtbHLH (PtbHLH131) and PtWDR (PtWD40-1) were co-expressed with PtMYB115 or PtMYB134 or expressed alone along with the reporter vector as a control. Bombarded activators and co-factors are represented on the x-axis. Data is representative of bombardment on one day. Error bars indicate standard error (n = 3).

3.3.3 Activation of the MYB115 promoter by MYB115 and MYB134

MYB134 is predicted to act upstream of MYB115, since the latter is highly upregulated (35.3 fold) in MYB134 overexpressing plants (Mellway, 2009). Activation of the MYB115 promoter by MYB134 was tested using the transient luciferase assay. MYB134 activated the promoter of MYB115 on average 3.7 times relative to the co-factor only control, based on assays conducted on three separate days (Figure 3-5). An analysis of variance with date as a random factor showed a significant activation by MYB134 of the MYB115 promoter compared to the co-factor only control (n = 9, p = 0.002). Furthermore, MYB115 could activate its own promoter with an average activation of 2.2 fold above the controls (Figure 3-5). Activation of the MYB115 promoter by MYB115 was replicated on two different days. An analysis of variance with date as a random factor showed a significantly higher activation by MYB115 with its co-factor compared to the co-factor only control (n = 6, p = 0.02). These results show that both MYB115 and MYB134 are able to activate the MYB115 promoter suggesting that MYB115 expression is regulated by its gene product and by MYB134.

To test if the MYB134 gene is also under the regulation of either MYB134 or MYB115, activation of a MYB134 promoter-luciferase construct was also tested. Neither MYB115 nor MYB134 could activate the MYB134 promoter in this assay (Figure 3-6). Activation of the MYB134 promoter was tested on two different days. A Tukey's test for honest significant differences showed no significant difference in activation. These results suggest that the MYB134 gene is not under the regulation of either MYB115 or MYB134.

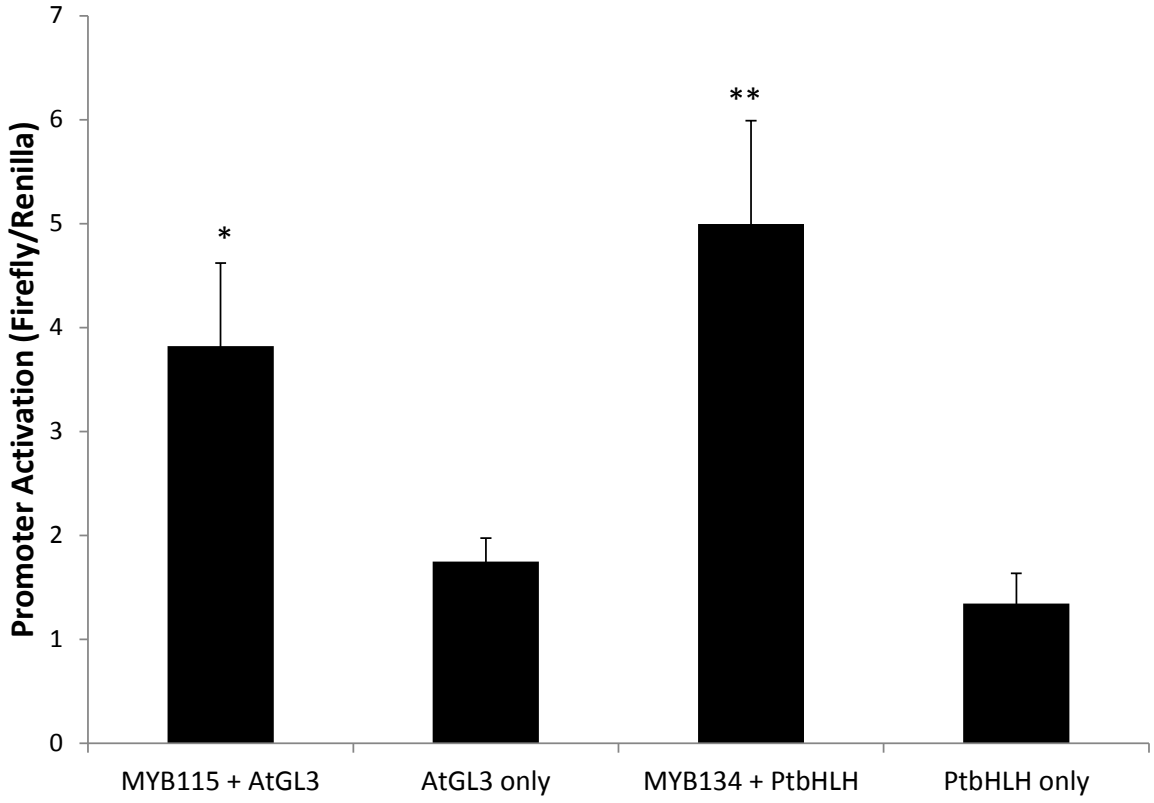

Figure 3-5. Activation of the MYB115 promoter by MYB115 and MYB134 using the dual-luciferase promoter assay following particle bombardment of H11-11 poplar suspension cells. AtGL3 and PtbHLH (PtbHLH131) were co-expressed with PtMYB115 or PtMYB134 or expressed alone as a control. Bombarded activators and co-factors are represented on the x-axis. Error bars indicate standard error (For MYB115 and AtGL3, n = 6 [representing bombardments on two different days], For MYB134 and PtbHLH, n = 9 [representing bombardments on three different days]).

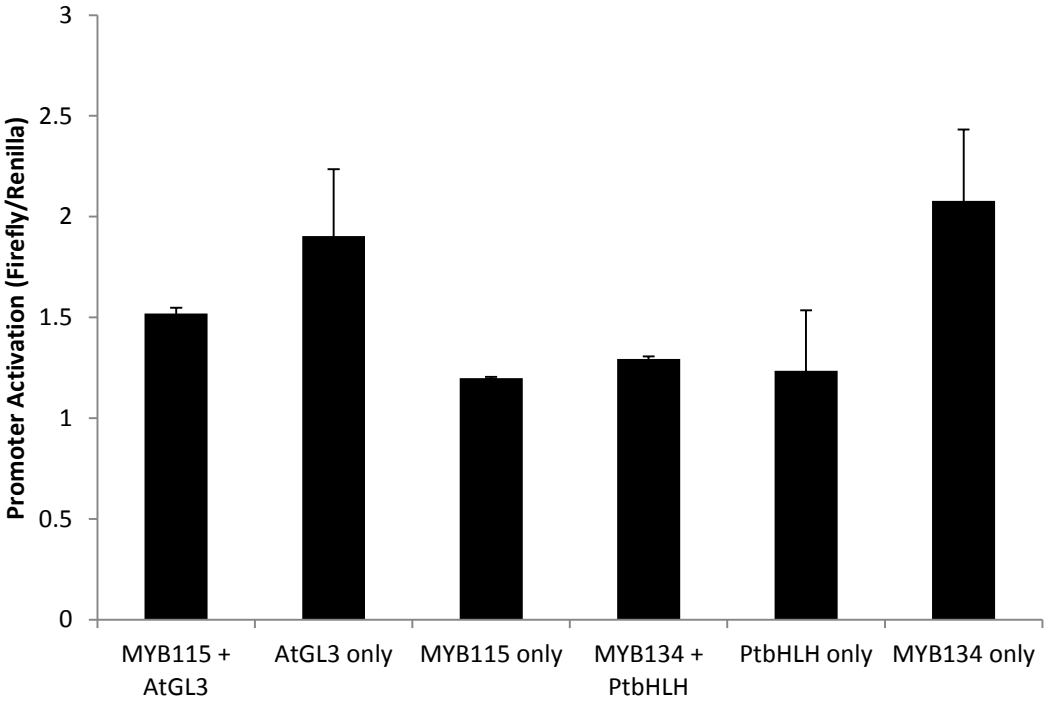

Figure 3-6. Testing activation of the MYB134 promoter by MYB115 and MYB134 using the dual-luciferase promoter assay following particle bombardment of H11-11 poplar suspension cells.

AtGL3 and PtbHLH (PtbHLH131) were co-expressed with PtMYB115 or PtMYB134 or expressed alone along with the reporter vector as a control. The MYB transcription factors were expressed alone as a second control. Error bars indicate standard error (n = 9).

3.3.4 AtGL3 as a bHLH co-factor for MYB115

Previous transient assays with MYB134 have shown that a bHLH co-factor is necessary for activation of the PtANR1 promoter, using the dual-luciferase promoter activation assay in Arabidopsis leaves (Zifkin et al, 2012; A. Gesell and C.P. Constabel, unpublished). Two bHLH co-factors, PtbHLH131 (Potri.005G208600) and AtGL3 (AT5G41315), showed similar activation of PtANR1 when co-expressed with MYB134 (Appendix A, Figure A-3). Here, I tested activation of PtANR1 with these two bHLH co-factors with MYB115 (Figure 3-1). PtbHLH131 is an ortholog of the bHLH, AtTT8, which is a co-factor of AtTT2 (Baudry et al., 2004). AtGL3 is a bHLH factor that is involved in the regulation of trichome development but has also been shown to induce anthocyanin synthesis in *Matthiola incana* petals and functions as a co-factor for the anthocyanin regulator, AtPAP1, in transient assays (Payne et al., 2000; Ramsay et al., 2003; Zimmermann et al., 2004). *Arabidopsis* leaves were bombarded on two different dates. Both PtbHLH131 and AtGL3 were able to co-activate the promoter of ANR1 with MYB115. AtGL3 as a co-factor showed a higher activation with MYB115 (27.7 fold-change) compared to PtbHLH131 (9.8 fold-change) (Figure 3-7). Subsequently to these experiments, the poplar ortholog of AtGL3, PtGL3 (PtbHLH79 [Potri.001G103600]), became available and was also tested as a co-factor with MYB115 and showed similar co-activation as AtGL3 (Appendix A, Figure A-4).

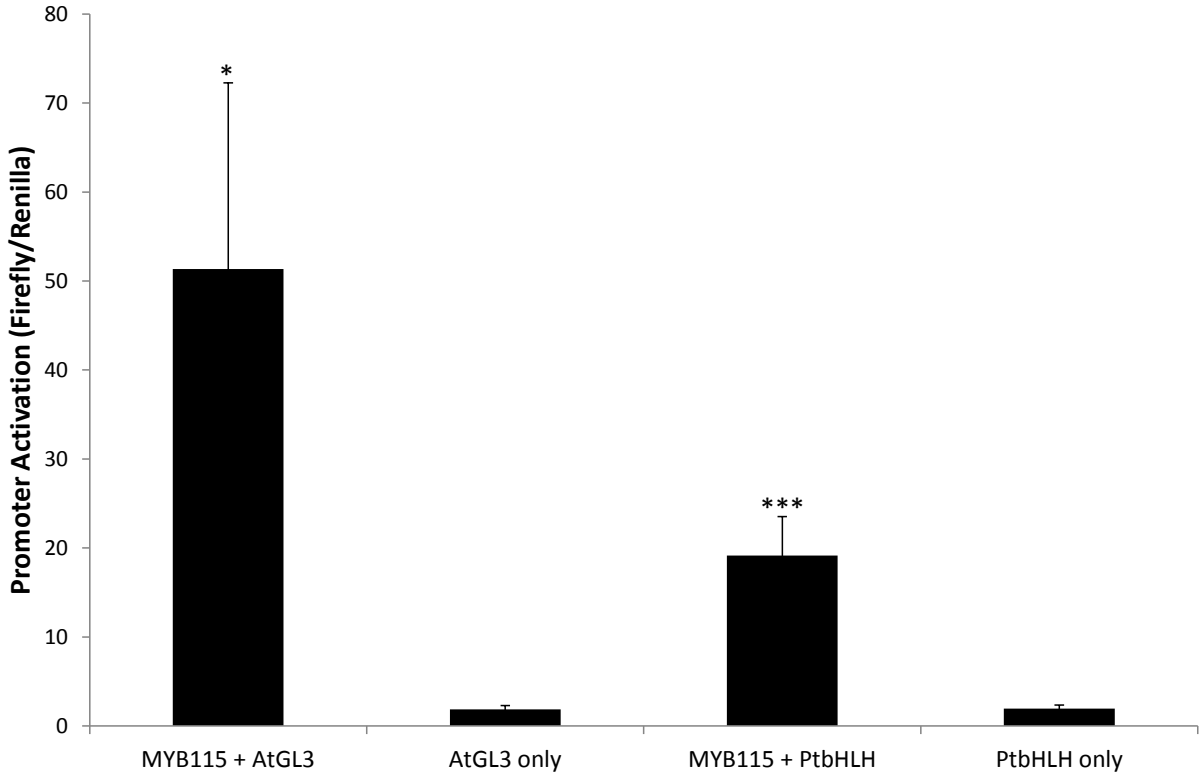

Figure 3-7. Analysis of two bHLH factors as co-factors for MYB115.
Activation of the anthocyanidin reductase (ANR1) promoter by the dual-luciferase promoter assay following bombardment of Arabidopsis leaves. AtGL3 and PtbHLH (PtbHLH131) were co-expressed with PtMYB115 or expressed alone as a control. Bombarded activators and co-factors are represented on the x-axis. Data is representative of two bombardments on two different dates. Error bars indicate standard error (n = 6). Asterisks indicate results of a one-way ANOVA with date as a residual factor ($p \leq 0.05$, *; $p \leq 0.01$, ** $p \leq 0.001$, ***).

3.4 Discussion

3.4.1 Activation of the promoters of flavonoid biosynthetic genes by MYB115 and MYB134

MYB115 overexpressing plants showed a high tannin phenotype as well as other changes in phytochemistry (Chapter 2). These changes in phytochemistry suggest a role for MYB115 in the regulation of flavonoid biosynthesis. To further demonstrate the role of MYB115 as a regulator in the biosynthesis of tannins, I used the dual-luciferase promoter activation assay in both poplar suspension cells and Arabidopsis leaves. The promoters of two key biosynthetic genes involved in the synthesis of condensed tannins were tested for activation, PtANR1 and PtLAR3.

PtANR1 was chosen primarily because it showed higher fold change than PtANR2 in MYB134 overexpressing plants and was thus implicated in condensed tannin synthesis (Mellway, 2009). Furthermore, PtANR1 has been shown to contribute to the synthesis of condensed tannins through analysis of PtANR1 overexpressing plants (Wang et al., 2013). Here, MYB115 was shown to activate the ANR1 promoter (Figure 3-1). These results support gene expression analysis of MYB115 overexpressing plants that show higher expression of PtANR1 compared to wild-type plants (Chapter Two). Together these results demonstrate that MYB115 is likely regulating the condensed tannin specific gene, ANR1, in conjunction with MYB134. These results are similar to studies in persimmon (Akagi et al., 2010) where both MYB transcription factors tested could activate the promoter of DkANR.

LAR is the second known enzyme specific to the synthesis of the flavan-3-ol precursors of condensed tannins. LAR3 has been shown to function in condensed tannin synthesis and shows a higher increase in transcript abundance in MYB134 plants compared to LAR1 or LAR2 (Mellway, 2009; Yuan et al., 2012). In the current work, MYB134 activated the promoter of LAR3 (Figure 3-3). However, activation of LAR3 by MYB134 was only 2.2 times above controls compared to a 7.2 fold activation of ANR1. This suggests that MYB134 may have a stronger preference for ANR1 relative to LAR3.

MYB115 along with its co-factor showed apparent weakly activation of the tannin biosynthetic gene LAR3 but the effect was not significant (Figure 3-3). LAR3 does show higher expression in MYB115 overexpressing plants (Chapter Two). However, induction of LAR3 was possibly a result of the induced expression of MYB134.

The weak activation of the LAR3 promoter by MYB115 and MYB134 could potentially be due to the promoter region tested. Since the DNA binding site is unknown, a promoter region of an arbitrary length of 1.9 kb upstream of the start codon is used for this assay. Potentially, the binding site is not within this region. Furthermore, in a study by Yin et al. (2005) the size of the insert was inversely correlated with activation. For LAR3 a 1.9 kb region was isolated and for ANR1 a 1.1 kb region was isolated. The larger size of the LAR3 promoter region analyzed may have negatively influenced the level of activation. Another possibility is that a different co-factor is necessary for activation. However, addition of the AtTTG1 homolog from poplar, PtWD40-1, or the bHLH factor AtGl3 had no influence on activity.

The results of this study correlate with the activity of the homologous regulators in persimmon. DkMYB2, the MYB134 ortholog, is able to activate both DkANR and DkLAR while DkMYB4, the MYB115 ortholog, can activate ANR but not LAR (Akagi et al., 2010). Since MYB115 was unable to activate LAR3, it is likely that it has a different *cis*-binding motif than MYB134. Mellway et al. (2009) found that recombinant MYB134 protein could bind to a 180 bp region in three flavonoid biosynthetic genes (PAL1, ANR2, and DFR1). The 180 bp region of each promoter contained a conserved AC element which is a putative binding site for MYB transcription factors (Lai et al., 2013). In persimmon, DkMYB2 binds to the AC element while DkMYB4 binds to the MYBCORE motif (Akagi et al., 2009, 2010). Akagi et al. (2010) propose that MYB transcription factors belonging to the PA-2 clade bind to the AC element and MYB transcription factors belonging to the PA-1 clade bind to the MYBCORE *cis*-motif. Since MYB115 belongs to the PA-1 clade, it potentially binds to the MYBCORE *cis*-motif. All promoter sequences tested here have both MYBCORE *cis*-motifs and AC elements so further analysis would be necessary to determine the DNA binding motif of MYB115.

3.4.2 Activation of the promoters of tannin regulatory factors by MYB115

MYB134 overexpressing plants show an increase in transcript abundance of MYB115 (Mellway, 2009). As expected, in the current experiments, MYB134 activated the MYB115 promoter above controls (Figure 3-5). This suggests that MYB134 acts up-stream of MYB115. In addition, MYB115 could activate its own promoter indicating that it is can regulate its own expression. Both MYB115 and MYB134 failed to activate the promoter of MYB134 (Figure 3-6). However, this is consistent with the hypothesis that MYB134 is a "master regulator" of condensed tannin synthesis regulating both the expression of biosynthetic genes as well as other regulatory factors (Mellway et al., 2009). These results correlate with the regulatory system in grape. Overexpression of the MYB134 homolog VvMYBPA2 leads to enhanced expression of the MYB115 homolog VvMYBPA1 and not *vice versa* (Terrier et al., 2009). Furthermore, as is seen with MYB115 and MYB134 transgenic plants, VvMYBPA1 or VvMYBPA2 overexpression results in highly similar phenotypes (Terrier et al., 2009). This raises questions about how MYB115 and MYB134 differ in their function as regulators of condensed tannin synthesis.

Mellway et al. (2009) demonstrated that MYB134 was induced following exposure to UV light, wounding, and infection by *Melampsora* sp. rust. Since MYB134 is not regulated by itself or MYB115, induction of MYB134 following stress may precede activation of MYB115. As MYB115 is a positive regulator of its own promoter, its role could be to maintain the expression of the flavonoid pathway genes following activation by MYB134 leading to strong activation of the flavonoid pathway and enhanced accumulation of tannins. This is further supported by analysis of MYB134 overexpressing plants that showed a much higher (35.3 fold) increase in MYB115 expression than the MYB134 transgene itself (7.97 fold) (Mellway, 2009). In apple, autoregulation by the regulator of anthocyanin synthesis, MYB10, leads to ectopic accumulation of anthocyanins (Espley et al., 2009). This autoregulatory ability was found to be sufficient for enhanced accumulation of anthocyanins as a minisatellite polymorphism in the promoter of MYB10 necessary for autoregulation was only present in red-fleshed fruit

(Espley et al., 2009). This supports the hypothesis that MYB115 autoregulation could lead to the strong high tannin phenotype observed in MYB115 overexpressing plants.

3.4.3 PtbHLH79 (PtGL3) is the putative bHLH co-factor of MYB115

As predicted based on previous transient promoter activation assays, a bHLH co-factor was necessary for activation of ANR1 by MYB115 (Figure 3-2). This has been observed in other promoter activation studies with R2R3 MYB regulators of condensed tannin synthesis (Akagi et al., 2010; Yoshida et al., 2010). Furthermore, a yeast two hybrid experiment showed that the regulator of condensed tannin synthesis in Arabidopsis, AtTT2, can directly interact with a bHLH factor (Baudry et al., 2004).

PtbHLH131 and AtGL3 were tested as co-factors for MYB115. PtbHLH131 is one of two poplar orthologs of the condensed tannin regulator, AtTT8, and AtGL3 is a bHLH factor involved in epidermal cell fate that belongs to a different clade. Both co-factors were able to co-activate the promoter of ANR along with MYB115. AtGL3, when co-bombarded with MYB115, showed a higher fold change activation compared to PtbHLH131. The poplar homolog of AtGL3, PtbHLH79, showed similar co-activation with MYB115. This suggests that PtbHLH79 may be the *in vivo* co-factor for MYB115; however, PtbHLH131, but not PtbHLH79, showed higher expression in MYB115 overexpressing plants (Table 2-3). To determine the *in vivo* co-factor for MYB115, a co-immunoprecipitation study could be performed.

Previous studies have shown that in promoter activation studies with MYB transcription factors, the addition of several different bHLH genes will result in increased co-activation suggesting that the bHLH protein has a less specific regulatory function than the MYB factor (Lin-Wang et al., 2010; Yoshida et al., 2010; Zimmermann et al., 2004). However, these studies have also shown differences in promoter activation by a MYB factor depending on the co-expressed bHLH co-factor which could indicate a degree of specificity for the interaction between the two factors (Lin-Wang et al., 2010; Yoshida et al., 2010; Zimmermann et al., 2004). For example, AtPAP1 and AtPAP2, the MYB regulators of anthocyanin synthesis, both show higher activation of the AtDFR

promoter with AtGL3 compared to AtTT8 while AtTT2 shows similar activation with both co-factors (Zimmermann et al., 2004). An amino acid motif has been identified in the R3 MYB domain that contributes to the interaction with bHLH factors (Grotewold et al., 2000; Zimmermann et al., 2004). AtPAP1, AtPAP2 and MYB115 all show an Asp residue at position 12 of the R3 domain while MYB134 and AtTT2 show a Glu (Figure 3-8). AtPAP1/2 and MYB115 show higher promoter activation with GL3 type factors than TT8 type factors while MYB134 and TT2 show similar activation with either. Potentially this residue is influencing the affinity of these MYB factors for a specific bHLH. Further protein-protein interaction tests with site specific mutations are necessary to verify this.

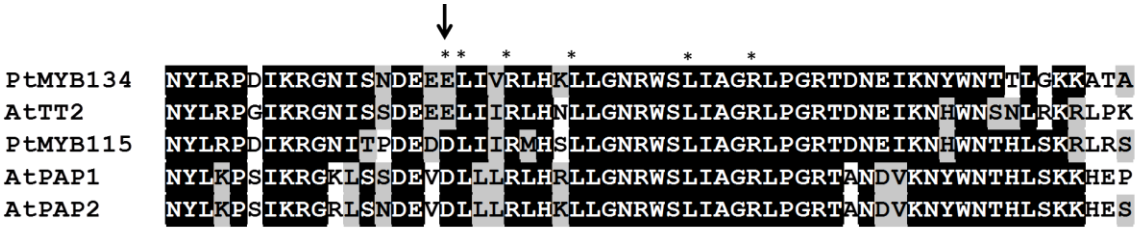

Figure 3-8. Sequence alignments of the conserved bHLH binding domain in the R3 region of MYB transcription factors involved in flavonoid biosynthesis.
Asterisks indicate conserved amino acids residues proposed to define interactions with bHLH co-factors (Zimmermann et al., 2004)**.** The arrow indicates the only amino acid of the conserved amino acid residues that differs between the MYB factors analyzed. Identical amino acids are shaded black while similar amino acids are shaded grey.

WDR factors act to stabilize interactions between proteins. For example, a combinatorial complex of MYB, WDR and bHLH factors is conserved between plants and is involved in the regulation of epidermal cell fate and secondary metabolism (Baudry et al., 2004; Ramsay and Glover, 2005). Here, the addition of a poplar ortholog of AtTTG1 (PtWD40-1), a WDR co-factor of AtTT2, was tested for enhanced activation of ANR1 and LAR3 promoters. Addition of a PtWD40-1 did not increase promoter activation by MYB115. This is comparable with experiments in Arabidopsis that show that TTG1 is not necessary for activation by TT2 and TT8 in a transient assay (Baudry et al., 2004). The lack of effect could be due to the endogenous expression of a WDR in the plants cells. If a WDR factor is already expressed at high levels in the cells, then an increase in expression might not lead to increased activation. Since WDR factors have been shown

to be less specific in their interactions, it is possible that a number of WDR factors could act as co-factors for MYB115. For example, AtTTG1 functions as a co-factor for multiple MYB proteins to regulate synthesis of flavonoids, seed coat development/pigmentation and trichome formation (Gonzalez et al., 2008, 2009; Zhao et al., 2008). Furthermore, the AtTTG1 mutant phenotype in Arabidopsis can be rescued by overexpression of a WDR factor from maize (Carey et al., 2004). Therefore, it is difficult to conclude if MYB115 forms a complex with a WDR factor from this analysis.

3.4.4 Conclusions

MYB115 overexpressing plants have been shown to accumulate condensed tannins at high levels (see Chapter Two). In this study, the dual-luciferase promoter assay was used to show that MYB115 can activate the promoters of condensed tannin biosynthetic genes and the promoters of tannin regulators. MYB115 can transcriptionally activate the ANR1 promoter, as did MYB134. MYB134 but not MYB115 could activate the promoter of LAR3. These results indicate that while MYB134 and MYB115 are both regulators of condensed tannin synthesis, they have different roles in regulating the pathway.

4. Chapter Four: Overall conclusions and future directions

Condensed tannins are stress-inducible plant polyphenolic compounds that have been shown to function in plant defense. Analysis of high tannin poplar plants overexpressing the known regulator of condensed tannin, MYB134, showed 35.3 fold higher expression of another MYB transcription factor, MYB115, compared to control plants. MYB115 shows high sequence similarity to characterized regulators of condensed tannin synthesis in grape (VvMYBPA1) and in persimmon (DkMYB4) suggesting it could function in regulation of condensed tannin synthesis. The overall objective of this study was to test MYB115 for such a role through the generation of transgenic poplar plants and by using a transient promoter activation assay. Both approaches demonstrated a role of MYB115 in condensed tannin synthesis. First, MYB115 overexpressing plants accumulate condensed tannins at higher levels than WT plants. Second, MYB115 overexpressing plants show enhanced expression of the flavonoid biosynthetic genes leading to condensed tannin synthesis as shown by microarray analysis and further verified by qPCR. Third, MYB115 activates the promoter of the tannin specific gene, ANR, as shown by the dual-luciferase promoter activation assay. Fourth, MYB115 expression is induced by wounding and coincides with the expression of flavonoid biosynthetic genes.

A second objective of this study was to compare the functions of MYB115 and MYB134. MYB115 was predicted to be downstream of MYB134 since it is induced in MYB134 overexpressing plants. The dual-luciferase promoter activation assay was used to demonstrate that MYB134 activates the MYB115 promoter confirming that MYB134 acted upstream of MYB115. The MYB115 gene product could also activate its own promoter, while neither MYB134 nor MYB115 could activate the promoter of MYB134. Therefore, once its expression is induced by MYB134, MYB115 could activate its own promoter and thus maintain condensed tannin synthesis. Another difference was that MYB134, and not MYB115, could activate the LAR3 promoter in a transient promoter

activation assay. These results also support the role of MYB134 as a “master regulator” of condensed tannin synthesis, regulating both biosynthetic genes and other regulatory genes.

While these findings show that the expression of MYB134 and MYB115 are regulated differently, ultimately MYB134 and MYB115 share a highly similar regulatory function in the regulation of tannin synthesis. It is still unclear if these two factors differ greatly in their function as tannin regulators. *In silico* analysis does not show that MYB134 is highly co-expressed with MYB115 suggesting they are differentially expressed through development. In persimmon, DkMYB4 but not DkMYB2 is regulated by seasonal temperature (Akagi et al., 2011). Potentially, MYB134 and MYB115 are also regulated by different environmental stimuli. Further expression analysis could help to determine how MYB115 and MYB134 differ in their regulatory function.

Gaining a greater understanding of how plant metabolism is regulated is of interest for genetic engineering of plants with improved characteristics such as increased resistance to pests. Here, MYB115 was characterized as an important regulator of condensed tannin synthesis as well as other phenolic pathways. However, MYB115 overexpressing plants also showed unexpected changes in several classes of plant defense compounds including flavonoids, hydroxycinnamic acids and phenolic glycosides. These unexpected changes in phenolic metabolism indicate additional levels of regulation by MYB115. These overall changes to plant chemistry will likely influence plant resistance to various stresses and may constrain the use of MYB transcription factors in the generation of plants with specific metabolic changes. Future work could utilize MYB115 transgenic plants to correlate chemical composition with the level of resistance to different stresses.

Bibliography

Abdel-Aal, E.S.M., Hucl, P., Sosulski, F.W., Graf, R., Gillott, C., and Pietrzak, L. (2001). Screening spring wheat for midge resistance in relation to ferulic acid content. J. Agric. Food Chem. *49*, 3559–3566.

Akagi, T., Ikegami, A., Tsujimoto, T., Kobayashi, S., Sato, A., Kono, A., and Yonemori, K. (2009). DkMyb4 Is a Myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. Plant Physiol. *151*, 2028 –2045.

Akagi, T., Ikegami, A., and Yonemori, K. (2010). DkMyb2 wound-induced transcription factor of persimmon (*Diospyros kaki* Thunb.), contributes to proanthocyanidin regulation. Planta *232*, 1045–1059.

Akagi, T., Tsujimoto, T., Ikegami, A., and Yonemori, K. (2011). Effects of seasonal temperature changes on DkMyb4 expression involved in proanthocyanidin regulation in two genotypes of persimmon (*Diospyros kaki* Thunb.) fruit. Planta *233*, 883–894.

Akagi, T., Katayama-Ikegami, A., Kobayashi, S., Sato, A., Kono, A., and Yonemori, K. (2012). Seasonal abscisic acid signal and a basic leucine zipper transcription factor, DkbZIP5, regulate proanthocyanidin biosynthesis in persimmon fruit. Plant Physiol. *158*, 1089–1102.

Albert, N.W., Lewis, D.H., Zhang, H., Schwinn, K.E., Jameson, P.E., and Davies, K.M. (2011). Members of an R2R3-MYB transcription factor family in Petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. Plant J. *65*, 771–784.

Arnold, T.M., Tanner, C.E., Rothen, M., and Bullington, J. (2008). Wound-induced accumulations of condensed tannins in turtlegrass, *Thalassia testudinum*. Aquat. Bot. *89*, 27–33.

Ayres, M.P., Clausen, T.P., MacLean, S.F., Redman, A.M., and Reichardt, P.B. (1997). Diversity of structure and antiherbivore activity in condensed tannins. Ecology *78*, 1696–1712.

Azaiez, A., Boyle, B., Levée, V., and Séguin, A. (2009). Transcriptome profiling in hybrid poplar following interactions with *Melampsora* rust fungi. Mol. Plant. Microbe Interact. *22*, 190–200.

Babst, B.A., Harding, S.A., and Tsai, C.-J. (2010). Biosynthesis of phenolic glycosides from phenylpropanoid and benzenoid precursors in *Populus*. J. Chem. Ecol. *36*, 286–297.

Barbehenn, R.V., and Peter Constabel, C. (2011). Tannins in plant–herbivore interactions. Phytochemistry *72*, 1551–1565.

Barbehenn, R., Cheek, S., Gasperut, A., Lister, E., and Maben, R. (2005). Phenolic compounds in Red Oak and Sugar Maple leaves have prooxidant activities in the midgut fluids of *Malacosoma disstria* and *Orgyia leucostigma* caterpillars. J. Chem. Ecol. *31*, 969–988.

Barbehenn, R., Weir, Q., and Salminen, J.-P. (2008). Oxidation of ingested phenolics in the tree-feeding caterpillar *Orgyia leucostigma* depends on foliar chemical composition. J. Chem. Ecol. *34*, 748–756.

Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant J. *39*, 366–380.

Baxter, I.R., Young, J.C., Armstrong, G., Foster, N., Bogenschutz, N., Cordova, T., Peer, W.A., Hazen, S.P., Murphy, A.S., and Harper, J.F. (2005). A plasma membrane H+-ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U. S. A. *102*, 2649–2654.

Beld, M., Martin, C., Huits, H., Stuitje, A., and Gerats, A. (1989). Flavonoid synthesis in *Petunia hybrida* - Partial characterization of dihydroflavonol-4-reductase genes. Plant Mol. Biol. *13*, 491–502.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Wheeler, D.L. (2005). GenBank. Nucleic Acids Res. *33*, D34–38.

Bernays, E.A., Chamberlain, D.J., and Leather, E.M. (1981). Tolerance of acridids to ingested condensed tannin. J. Chem. Ecol. *7*, 247–256.

Bieza, K., and Lois, R. (2001). An Arabidopsis mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. Plant Physiol. *126*, 1105–1115.

Boeckler, G.A., Gershenzon, J., and Unsicker, S.B. (2011). Phenolic glycosides of the Salicaceae and their role as anti-herbivore defenses. Phytochemistry *72*, 1497–1509.

Bogs, J., Downey, M.O., Harvey, J.S., Ashton, A.R., Tanner, G.J., and Robinson, S.P. (2005). Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves. Plant Physiol. *139*, 652–663.

Bogs, J., Jaffé, F.W., Takos, A.M., Walker, A.R., and Robinson, S.P. (2007). The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. Plant Physiol. *143*, 1347 –1361.

Bradshaw, H.D., Ceulemans, R., Davis, J., and Stettler, R. (2000). Emerging model systems in plant biology: Poplar (*Populus*) as a model forest tree. J. Plant Growth Regul. *19*, 306–313.

Brunner, A.M., Busov, V.B., and Strauss, S.H. (2004). Poplar genome sequence: functional genomics in an ecologically dominant plant species. Trends Plant Sci. *9*, 49–56.

Burbulis, I.E., and Winkel-Shirley, B. (1999). Interactions among enzymes of the Arabidopsis flavonoid biosynthetic pathway. Proc. Natl. Acad. Sci. U. S. A. *96*, 12929–12934.

Cabrera, H.M., Muñoz, O., Zúñiga, G.E., Corcuera, L.J., and Argandoña, V.H. (1995). Changes in ferulic acid and lipid content in aphid-infested barley. Phytochemistry *39*, 1023–1026.

Carey, C.C., Strahle, J.T., Selinger, D.A., and Chandler, V.L. (2004). Mutations in the pale aleurone color1 regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar TRANSPARENT TESTA GLABRA1 gene in *Arabidopsis thaliana*. Plant Cell *16*, 450–464.

Clausen, T.P., Koller, J.W., and Reichardt, P.B. (1990). Aglycone fragmentation accompanies β-glucosidase catalyzed hydrolysis of salicortin, a naturally-occurring phenol glycoside. Tetrahedron Lett. *31*, 4537–4538.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformaics *21*, 3674–3676.

Curtis, M.D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol. *133*, 462–469.

D'Auria, J.C. (2006). Acyltransferases in plants: a good time to be BAHD. Curr. Opin. Plant Biol. *9*, 331–340.

Datla, R.S., Hammerlindl, J.K., Panchuk, B., Pelcher, L.E., and Keller, W. (1992). Modified binary plant transformation vectors with the wild-type gene encoding NPTII. Gene *122*, 383–384.

Devic, M., Guilleminot, J., Debeaujon, I., Bechtold, N., Bensaude, E., Koornneef, M., Pelletier, G., and Delseny, M. (1999). The BANYULS gene encodes a DFR-like protein and is a marker of early seed coat development. Plant J. *19*, 387–398.

Ding, H., Lamb, R.J., and Ames, N. (2000). Inducible production of phenolic acids in wheat and antibiotic resistance to *Sitodiplosis mosellana*. J. Chem. Ecol. *26*, 969–985.

Donaldson, J.R., Stevens, M.T., Barnhill, H.R., and Lindroth, R.L. (2006). Age-related shifts in leaf chemistry of clonal aspen (*Populus tremuloides*). J. Chem. Ecol. *32*, 1415–1429.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. Trends Plant Sci. *15*, 573–581.

Espley, R.V., Brendolise, C., Chagné, D., Kutty-Amma, S., Green, S., Volz, R., Putterill, J., Schouten, H.J., Gardiner, S.E., Hellens, R.P., et al. (2009). Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. Plant Cell *21*, 168–183.

Feeny, P. (1968). Effect of oak leaf tannins on larval growth of the winter moth *Operophtera brumata*. J. Insect Physiol. *14*, 805–817.

Feeny, P. (1969). Inhibitory effect of oak leaf tannins on the hydrolysis of proteins by trypsin. Phytochemistry *8*, 2119–2126.

Feeny, P. (1970). Seasonal changes in oak leaf tannins and nutrients as a cause of spring feeding by winter moth caterpillars. Ecology *51*, 565–581.

Feller, A., Machemer, K., Braun, E.L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J. *66*, 94–116.

Ferrer, J.-L., Jez, J.M., Bowman, M.E., Dixon, R.A., and Noel, J.P. (1999). Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat. Struct. Mol. Biol. *6*, 775–784.

Feucht, W., and Treutter, D. (1990). Flavan-3-ols in trichomes, pistils and phelloderm of some tree species. Ann. Bot. *65*, 225–230.

Fillatti, J.J., Sellmer, J., McCown, B., Haissig, B., and Comai, L. (1987). Agrobacterium mediated transformation and regeneration of *Populus*. Mol. Gen. Genet. *206*, 192–199.

Fischer, D., Stich, K., Britsch, L., and Grisebach, H. (1988). Purification and characterization of (+)dihydroflavonol (3-hydroxyflavanone) 4-reductase from flowers of *Dahlia variabilis*. Arch. Biochem. Biophys. *264*, 40–47.

Gonzalez, A., Zhao, M., Leavitt, J.M., and Lloyd, A.M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. Plant J. *53*, 814–827.

Gonzalez, A., Mendenhall, J., Huo, Y., and Lloyd, A. (2009). TTG1 complex MYBs, MYB5 and TT2, control outer seed coat differentiation. Dev. Biol. *325*, 412–421.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., et al. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. *40*, D1178–1186.

Grotewold, E., Sainz, M.B., Tagliani, L., Hernandez, J.M., Bowen, B., and Chandler, V.L. (2000). Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R. Proc. Natl. Acad. Sci. *97*, 13579–13584.

Hagerman, A.E., Rice, M.E., and Ritchard, N.T. (1998). Mechanisms of protein precipitation for two tannins, pentagalloyl glucose and epicatechin $_{16}$ (4→8) catechin (procyanidin). J. Agric. Food Chem. *46*, 2590–2595.

Harborne, J.B., and Williams, C.A. (2000). Advances in flavonoid research since 1992. Phytochemistry *55*, 481–504.

Haruta, M., Pedersen, J.A., and Constabel, C.P. (2001). Polyphenol oxidase and herbivore defense in trembling aspen (*Populus tremuloides*): cDNA cloning, expression, and potential substrates. Physiol. Plant. *112*, 552–558.

Hellens, R., Edwards, E., Leyland, N., Bean, S., and Mullineaux, P. (2000). pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol. Biol. *42*, 819–832.

Hichri, I., Heppel, S.C., Pillet, J., Léon, C., Czemmel, S., Delrot, S., Lauvergeat, V., and Bogs, J. (2010). The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. Mol. Plant *3*, 509–523.

Holton, T.A., Brugliera, F., and Tanaka, Y. (1993). Cloning and expression of flavonol synthase from *Petunia hybrida*. Plant J. *4*, 1003–1010.

Huang, Y., Gou, J., Jia, Z., Yang, L., Sun, Y., Xiao, X., Song, F., and Luo, K. (2012). Molecular cloning and characterization of two genes encoding dihydroflavonol-4-reductase from *Populus trichocarpa*. PLoS ONE *7*, e30364.

Hwang, S.-Y., and Lindroth, R.L. (1997). Clonal variation in foliar chemistry of aspen: Effects on gypsy moths and forest tent caterpillars. Oecologia *111*, 99–108.

Ikonen, A., Tahvanainen, J., and Roininen, H. (2001). Chlorogenic acid as an antiherbivore defence of willows against leaf beetles. Entomol. Exp. Appl. *99*, 47–54.

Jassbi, A.R. (2003). Secondary metabolites as stimulants and antifeedants of *Salix integra* for the leaf beetle *Plagiodera versicolora*. Z. Naturforschung C- J. Biosci. *58*, 573–579.

Jez, J.M., Bowman, M.E., Dixon, R.A., and Noel, J.P. (2000). Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. Nat. Struct. Mol. Biol. *7*, 786–791.

Johnson, E.T., Ryu, S., Yi, H.K., Shin, B., Cheong, H., and Choi, G. (2001). Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. Plant J. *25*, 325–333.

Kim, B.-G., Lee, E.-R., and Ahn, J.-H. (2012). Analysis of flavonoid contents and expression of flavonoid biosynthetic genes in *Populus euramericana* Guinier in response to abiotic stress. J. Korean Soc. Appl. Biol. Chem. *55*, 141–145.

Kitamura, S., Shikazono, N., and Tanaka, A. (2004). TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in Arabidopsis. Plant J. *37*, 104–114.

Lai, Y., Li, H., and Yamagishi, M. (2013). A review of target gene specificity of flavonoid R2R3-MYB transcription factors and a discussion of factors contributing to the target gene selectivity. Front. Biol. 1–22.

Leiss, K.A., Maltese, F., Choi, Y.H., Verpoorte, R., and Klinkhamer, P.G.L. (2009). Identification of chlorogenic acid as a resistance factor for thrips in Chrysanthemum. Plant Physiol. *150*, 1567–1575.

Li, J., Ou-Lee, T.-M., Raba, R., Amundson, R.G., and Last, R.L. (1993). Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. Plant Cell *5*, 171–179.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402–408.

Lukačin, R., Gröning, I., Schiltz, E., Britsch, L., and Matern, U. (2000). Purification of recombinant flavanone 3β-hydroxylase from *Petunia hybrida* and assignment of the primary site of proteolytic degradation. Arch. Biochem. Biophys. *375*, 364–370.

Maher, E., Bate, N., Ni, W., Elkind, Y., Dixon, R., and Lamb, C. (1994). Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products. Proc. Natl. Acad. Sci. U. S. A. *91*, 7802–7806.

Marinova, K., Pourcel, L., Weder, B., Schwarz, M., Barron, D., Routaboul, J.-M., Debeaujon, I., and Klein, M. (2007). The Arabidopsis MATE transporter TT12 acts as a vacuolar flavonoid/H+ -antiporter active in proanthocyanidin-accumulating cells of the seed coat. Plant Cell *19*, 2023–2038.

Martin, M.M., Rockholm, D.C., and Martin, J.S. (1985). Effects of surfactants, pH, and certain cations on precipitation of proteins by tannins. J. Chem. Ecol. *11*, 485–494.

Matus, J., Aquea, F., and Arce-Johnson, P. (2008). Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across *Vitis* and *Arabidopsis* genomes. BMC Plant Biol. *8*, 83.

McArt, S.H., Spalinger, D.E., Collins, W.B., Schoen, E.R., Stevenson, T., and Bucho, M. (2009). Summer dietary nitrogen availability as a potential bottom-up constraint on moose in south-central Alaska. Ecology *90*, 1400–1411.

Mellway, R. (2009). The regulation of stress-induced proanthocyanidin metabolism in poplar. PhD Thesis. University of Victoria, Victoria, BC, Canada.

Mellway, R.D., Tran, L.T., Prouse, M.B., Campbell, M.M., and Constabel, C.P. (2009). The wound-, pathogen-, and ultraviolet B-responsive MYB134 gene encodes an R2R3 MYB transcription factor that regulates proanthocyanidin synthesis in Poplar. Plant Physiol. *150*, 924 –941.

Miranda, M., Ralph, S.G., Mellway, R., White, R., Heath, M.C., Bohlmann, J., and Constabel, C.P. (2007). The transcriptional response of hybrid poplar (*Populus trichocarpa* x *P. deltoides*) to infection by *Melampsora medusae* leaf rust involves induction of flavonoid pathway genes leading to the accumulation of proanthocyanidins. Mol. Plant. Microbe Interact. *20*, 816–831.

Mueller, L.A., Goodman, C.D., Silady, R.A., and Walbot, V. (2000). AN9, a petunia glutathione S-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. Plant Physiol. *123*, 1561–1570.

Muoki, R.C., Paul, A., Kumari, A., Singh, K., and Kumar, S. (2012). An improved protocol for the isolation of RNA from roots of tea (*Camellia sinensis* (L.) O. Kuntze). Mol. Biotechnol. *52*, 82–88.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. *15*, 473–497.

Nesi, N., Jond, C., Debeaujon, I., Caboche, M., and Lepiniec, L. (2001). The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. Plant Cell *13*, 2099 –2114.

Osier, T.L., and Lindroth, R.L. (2004). Long-term effects of defoliation on quaking aspen in relation to genotype and nutrient availability: plant growth, phytochemistry and insect performance. Oecologia *139*, 55–65.

Osier, T.L., Hwang, S.-Y., and Lindroth, R.L. (2000). Effects of phytochemical variation in quaking aspen *Populus tremuloides* clones on gypsy moth *Lymantria dispar* performance in the field and laboratory. Ecol. Entomol. *25*, 197–207.

Pang, Y., Peel, G.J., Wright, E., Wang, Z., and Dixon, R.A. (2007). Early steps in proanthocyanidin biosynthesis in the model legume *Medicago truncatula*. Plant Physiol. *145*, 601–615.

Pang, Y., Cheng, X., Huhman, D.V., Ma, J., Peel, G.J., Yonekura-Sakakibara, K., Saito, K., Shen, G., Sumner, L.W., Tang, Y., et al. (2013). Medicago glucosyltransferase UGT72L1: potential roles in proanthocyanidin biosynthesis. Planta *238*, 139–154.

Pasch, H., Pizzi, A., and Rode, K. (2001). MALDI–TOF mass spectrometry of polyflavonoid tannins. Polymer *42*, 7531–7539.

Payne, C.T., Zhang, F., and Lloyd, A.M. (2000). GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. Genetics *156*, 1349–1362.

Pearl, I.A., and Darling, S.F. (1962). Studies on the barks of the ramily salicaceae. V.1 Grandidentatin, a new glucoside from the bark of *Populus grandidentata*. J. Org. Chem. *27*, 1806–1809.

Pearl, I.A., and Darling, S.F. (1970). Phenolic extractives of *Salix purpurea* bark. Phytochemistry *9*, 1277–1281.

Penuelas, J., and Estiarte, M. (1997). Trends in plant carbon concentration and plant demand for N throughout this century. Oecologia *109*, 69–73.

Peters, D.J., and Constabel, C.P. (2002). Molecular analysis of herbivore-induced condensed tannin synthesis: cloning and expression of dihydroflavonol reductase from trembling aspen (*Populus tremuloides*). Plant J. *32*, 701–712.

Porter, L.J. (1992). Structure and Chemical Properties of the Condensed Tannins. In Plant Polyphenols, R.W. Hemingway, and P.E. Laks, eds. (Springer US), pp. 245–258.

Porter, L.J., Hrstich, L.N., and Chan, B.G. (1985). The conversion of procyanidins and prodelphinidins to cyanidin and delphinidin. Phytochemistry *25*, 223–230.

Prudic, K.L., Khera, S., Sólyom, A., and Timmermann, B.N. (2007). Isolation, identification, and quantification of potential defensive compounds in the viceroy butterfly and its larval host-plant, Carolina willow. J. Chem. Ecol. *33*, 1149–1159.

Ramsay, N.A., and Glover, B.J. (2005). MYB–bHLH–WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci. *10*, 63–70.

Ramsay, N.A., Walker, A.R., Mooney, M., and Gray, J.C. (2003). Two basic-helix-loop-helix genes (MYC-146 and GL3) from Arabidopsis can activate anthocyanin biosynthesis in a white-flowered *Matthiola incana* mutant. Plant Mol. Biol. *52*, 679–688.

Ryan, K.G., Swinny, E.E., Winefield, C., and Markham, K.R. (2001). Flavonoids and UV photoprotection in Arabidopsis mutants. Z. Naturforschung C- J. Biosci. *56*, 745–754.

Saikumar, P., Murali, R., and Reddy, E.P. (1990). Role of tryptophan repeats and flanking amino acids in MYB-DNA interactions. Proc. Natl. Acad. Sci. *87*, 8452 –8456.

Sannigrahi, P., Ragauskas, A.J., and Tuskan, G.A. (2010). Poplar as a feedstock for biofuels: A review of compositional characteristics. Biofuels Bioprod. Biorefining *4*, 209–226.

Saslowsky, D., and Winkel-Shirley, B. (2001). Localization of flavonoid enzymes in Arabidopsis roots. Plant J. *27*, 37–48.

Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., and Werck-Reichhart, D. (2001). CYP98A3 from *Arabidopsis thaliana* is a 3 ’-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J. Biol. Chem. *276*, 36566–36574.

Schweitzer, J.A., Madritch, M.D., Bailey, J.K., LeRoy, C.J., Fischer, D.G., Rehill, B.J., Lindroth, R.L., Hagerman, A.E., Wooley, S.C., Hart, S.C., et al. (2008). From genes to ecosystems: The genetic basis of condensed tannins and their role in nutrient regulation in a *Populus* model system. Ecosystems *11*, 1005–1020.

Scioneaux, A.N., Schmidt, M.A., Moore, M.A., Lindroth, R.L., Wooley, S.C., and Hagerman, A.E. (2011). Qualitative variation in proanthocyanidin composition of *Populus* species and hybrids: Genetics is the key. J. Chem. Ecol. *37*, 57–70.

Seitz, C., Eder, C., Deiml, B., Kellner, S., Martens, S., and Forkmann, G. (2006). Cloning, functional identification and sequence analysis of flavonoid 3’-hydroxylase and flavonoid 3’,5’-hydroxylase cDNAs reveals independent evolution of flavonoid 3’,5’-hydroxylase in the Asteraceae Family. Plant Mol. Biol. *61*, 365–381.

Shadle, G.L., Wesley, S.V., Korth, K.L., Chen, F., Lamb, C., and Dixon, R.A. (2003). Phenylpropanoid compounds and disease resistance in transgenic tobacco with altered expression of L-phenylalanine ammonia-lyase. Phytochemistry *64*, 153–161.

Sharma, S.B., and Dixon, R.A. (2005). Metabolic engineering of proanthocyanidins by ectopic expression of transcription factors in Arabidopsis thaliana. Plant J. *44*, 62–75.

Shimada, T. (2006). Salivary proteins as a defense against dietary tannins. J. Chem. Ecol. *32*, 1149–1163.

Shirley, B.W., Kubasek, W.L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F.M., and Goodman, H.M. (1995). Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. Plant J. *8*, 659–671.

Skadhauge, B., Thomsen, K.K., and von Wettstein, D. (1997). The role of the barley testa layer and its flavonoid content in resistance to *Fusarium* infections. Hereditas *126*, 147–160.

Solovchenko, A., and Schmitz-Eiberger, M. (2003). Significance of skin flavonoids for UV-B-protection in apple fruits. J. Exp. Bot. *54*, 1977–1984.

Sun, Y., Tian, Q., Yuan, L., Jiang, Y., Huang, Y., Sun, M., Tang, S., and Luo, K. (2011). Isolation and promoter analysis of a chalcone synthase gene PtrCHS4 from *Populus trichocarpa*. Plant Cell Reports *30*, 1661–1671.

Tanner, G.J., Francki, K.T., Abrahams, S., Watson, J.M., Larkin, P.J., and Ashton, A.R. (2003). Proanthocyanidin biosynthesis in plants. Purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. J. Biol. Chem. *278*, 31647–31656.

Terrier, N., Torregrosa, L., Ageorges, A., Vialet, S., Verriès, C., Cheynier, V., and Romieu, C. (2009). Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. Plant Physiol. *149*, 1028 – 1041.

Thompson, E.P., Wilkins, C., Demidchik, V., Davies, J.M., and Glover, B.J. (2010). An Arabidopsis flavonoid transporter is required for anther dehiscence and pollen development. J. Exp. Bot. *61*, 439–451.

Tsai, C.-J., Harding, S.A., Tschaplinski, T.J., Lindroth, R.L., and Yuan, Y. (2006). Genome-wide analysis of the structural genes regulating defense phenylpropanoid metabolism in Populus. New Phytol. *172*, 47–62.

Tsai, C.-J., Ranjan, P., DiFazio, S., Tuskan, G.A., and Johnson, V. (2011). Poplar genome microarrays. In Genetics, Genomics and Breeding of Poplars, (Science Publishers, Enfield, NH), pp. 112–127.

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

Veljanovski, V., and Constabel, C.P. (2013). Molecular cloning and biochemical characterization of two UDP-glycosyltransferases from poplar. Phytochemistry *91*, 148–157.

Verdier, J., Zhao, J., Torres-Jerez, I., Ge, S., Liu, C., He, X., Mysore, K.S., Dixon, R.A., and Udvardi, M.K. (2012). MtPAR MYB transcription factor acts as an on switch for proanthocyanidin biosynthesis in *Medicago truncatula*. Proc. Natl. Acad. Sci. *109*, 1766–1771.

De Vetten, N., Quattrocchio, F., Mol, J., and Koes, R. (1997). The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev. *11*, 1422–1434.

Vogt, T. (2010). Phenylpropanoid Biosynthesis. Mol. Plant *3*, 2–20.

Wallis, C.M., and Chen, J. (2012). Grapevine phenolic compounds in xylem sap and tissues are significantly altered during infection by *Xylella fastidiosa*. Phytopathology *102*, 816–826.

Wang, J., and Constabel, C.P. (2004). Polyphenol oxidase overexpression in transgenic *Populus* enhances resistance to herbivory by forest tent caterpillar (*Malacosoma disstria*). Planta *220*, 87–96.

Wang, L., Jiang, Y., Yuan, L., Lu, W., Yang, L., Karim, A., and Luo, K. (2013). Isolation and characterization of cDNAs encoding leucoanthocyanidin reductase and anthocyanidin reductase from *Populus trichocarpa*. PLoS ONE *8*, e64664.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*, 57–63.

Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T.K., Espley, R.V., Hellens, R.P., and Allan, A.C. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC Plant Biol. *10*, 50.

Warren, J.M., Bassman, J.H., Fellman, J.K., Mattinson, D.S., and Eigenbrode, S. (2003). Ultraviolet-B radiation alters phenolic salicylate and flavonoid composition of *Populus trichocarpa* leaves. Tree Physiol. *23*, 527–535.

Whitham, T.G., Bailey, J.K., Schweitzer, J.A., Shuster, S.M., Bangert, R.K., LeRoy, C.J., Lonsdorf, E.V., Allan, G.J., DiFazio, S.P., Potts, B.M., et al. (2006). A framework for community and ecosystem genetics: from genes to ecosystems. Nat. Rev. Genet. *7*, 510–523.

Wilkins, O., Nahal, H., Foong, J., Provart, N.J., and Campbell, M.M. (2009). Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. Plant Physiol. *149*, 981 –993.

Wu, J.-F., Wang, Y., and Wu, S.-H. (2008). Two new clock proteins, LWD1 and LWD2, regulate Arabidopsis photoperiodic flowering. Plant Physiol. *148*, 948–959.

Xie, D.-Y., Jackson, L.A., Cooper, J.D., Ferreira, D., and Paiva, N.L. (2004). Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol-4-reductase from *Medicago truncatula*. Plant Physiol. *134*, 979–994.

Yin, W., Xiang, P., and Li, Q. (2005). Investigations of the effect of DNA size in transient transfection assay using dual luciferase system. Anal. Biochem. *346*, 289–294.

Yoshida, K., Iwasaka, R., Kaneko, T., Sato, S., Tabata, S., and Sakuta, M. (2008). Functional differentiation of *Lotus japonicus* TT2s, R2R3-MYB transcription factors comprising a multigene family. Plant Cell Physiol. *49*, 157–169.

Yoshida, K., Kume, N., Nakaya, Y., Yamagami, A., Nakano, T., and Sakuta, M. (2010). Comparative analysis of the triplicate proathocyanidin regulators in *Lotus japonicus*. Plant Cell Physiol. *51*, 912 –922.

Young, B., Wagner, D., Doak, P., and Clausen, T. (2010). Within-plant distribution of phenolic glycosides and extrafloral nectaries in trembling aspen (*Populus tremuloides*; Salicaceae). Am. J. Bot. *97*, 601–610.

Yu, X.-H., Gou, J.-Y., and Liu, C.-J. (2009). BAHD superfamily of acyl-CoA dependent acyltransferases in *Populus* and *Arabidopsis*: bioinformatics and gene expression. Plant Mol. Biol. *70*, 421–442.

Yuan, L., Wang, L., Han, Z., Jiang, Y., Zhao, L., Liu, H., Yang, L., and Luo, K. (2012). Molecular cloning and characterization of PtrLAR3, a gene encoding leucoanthocyanidin reductase from *Populus trichocarpa*, and its constitutive expression enhances fungal resistance in transgenic plants. J. Exp. Bot. *63*, 2513–2524.

Zhang, Z.-Z., Che, X.-N., Pan, Q.-H., Li, X.-X., and Duan, C.-Q. (2013). Transcriptional activation of flavan-3-ols biosynthesis in grape berries by UV irradiation depending on developmental stage. Plant Sci. *208*, 64–74.

Zhao, M., Morohashi, K., Hatlestad, G., Grotewold, E., and Lloyd, A. (2008). The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. Development *135*, 1991–1999.

Zifkin, M., Jin, A., Ozga, J.A., Zaharia, L.I., Schernthaner, J.P., Gesell, A., Abrams, S.R., Kennedy, J.A., and Constabel, C.P. (2012). Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. Plant Physiol. *158*, 200–224.

Zimmermann, I.M., Heim, M.A., Weisshaar, B., and Uhrig, J.F. (2004). Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. Plant J. *40*, 22–34.

Appendix A: Supplementary Figures and Tables

This appendix contains the following supplementary material:

Figure A-1 presents expression of MYB115 in wounded 353-38 plants. It is a replicate experiment of the experiment shown in Figure 2-1.

Figure A-2 presents an example HPLC chromatogram for MYB115 overexpressing plants and wild-type plants with both 353-38 and 717-1-B4 genetic backgrounds.

Table A-1 presents a complete list of upregulated genes in 353-38 MYB115 overexpressing plants (Line 5) compared to wild-type plants as analyzed by Affymetrix GeneChip® Poplar Genome Array.

Table A-2 presents a complete list of downregulated genes in 353-38 MYB115 overexpressing plants (Line 5) compared to wild-type plants as analyzed by Affymetrix GeneChip® Poplar Genome Array.

Table A-3 presents correlations in MYB115 overexpressing plants and wild-type plants between tannin concentrations and concentrations of other phenolics as well as expression of genes as analyzed by qPCR. The data used for correlation analysis is presented in Chapter 2.

Figure A-3 presents activation of the ANR1 promoter by MYB134 with two different bHLH cofactors (PtbHLH131 and AtGL3).

Figure A-4 presents activation of the MYB115 promoter by MYB115 with two different bHLH cofactors (AtGL3 and PtbHLH79)

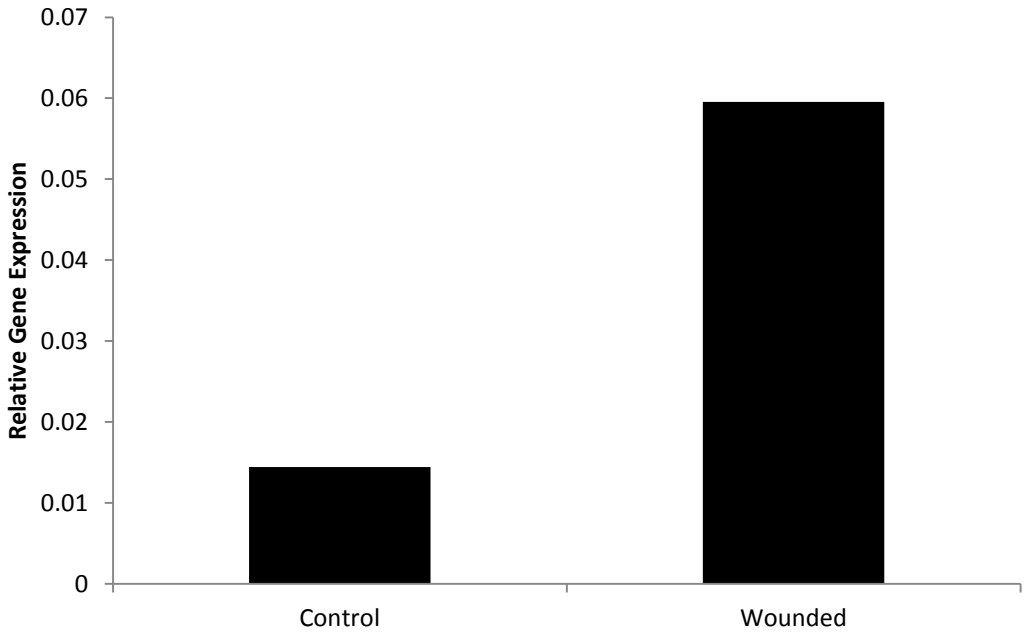

Figure A-1. MYB115 expression in control and wounded 353-38 plant leaves.
The wounding experiment, performed as described in section 2.2.1, was repeated to confirm MYB115 expression is induced by wounding. (n = 1 plant). The results of the other wounding experiment are presented in Figure 2-1.

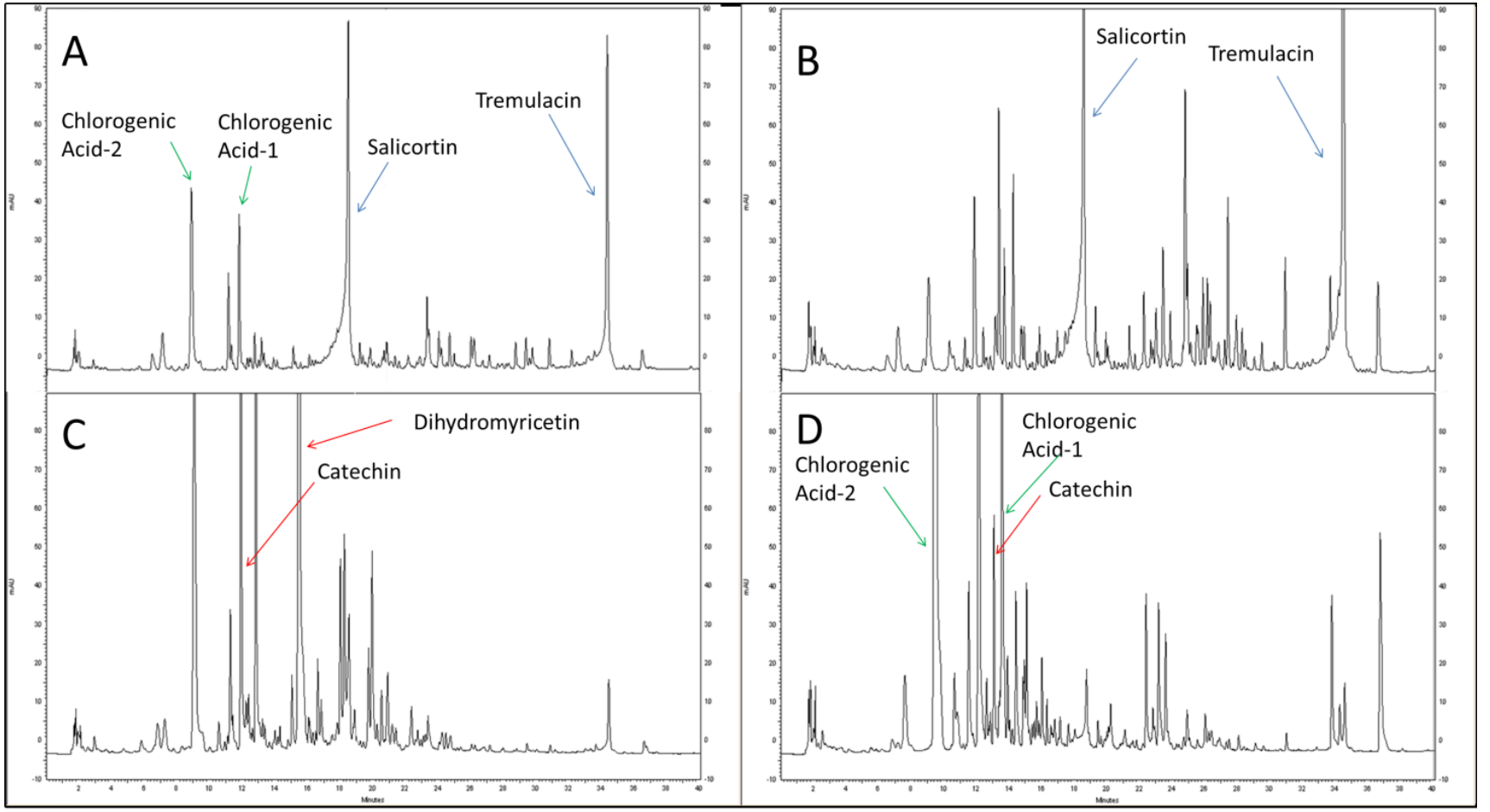

Figure A-2. HPLC analysis of soluble phenolic metabolites from representative MYB115 overexpressing poplar leaf tissue (C [353 - 38background] and D [717-1-B4 background]) and wild-type (WT) leaf tissue (A [353-38 background] and B [717-1-B4 background]). A subset of compounds are labelled with coloured arrows. Blue arrows indicate phenolic glycosides. Green arrows indicate chlorogenic acid derivatives. Flavonoids leading to the synthesis of condensed tannins are labelled with red arrows. These peaks were quantified in all transgenic lines and the data is presented in following figures. Compounds were identified based on comparison of UV spectra of representative standards as well as from the fragmentation pattern from LC/MS analysis.

Table A-1. Complete list of upregulated genes (fold change > 2) in 353 MYB115 overexpressing plants (Line 5). Genes are listed from highest to lowest fold change. Fold change is relative to 353 WT plants (n = 3 plants). All functional annotations are from Affymetrix annotation files with the exception of the probes that presented in Table 2-4.

Probeset	Gene Model	Gene Title	Fold change	P-value
PtpAffx.83404.1.A1_at	Potri.009G069100	flavonoid 3',5'-hydroxylase (CYP75A12|F3'5'H1)	138.3866	1.45E-07
PtpAffx.224252.1.S1_at	Potri.002G173900	MYB115	121.7093	5.27E-10
PtpAffx.30659.1.A1_at	Potri.002G173900	MYB115	101.6081	1.73E-09
Ptp.8030.1.S1_s_at	Potri.014G019200	cytochrome b5	81.52364	1.39E-09
PtpAffx.7896.2.S1_at	Potri.003G176700	chalcone synthase(CHS4)	64.90673	9.49E-11
PtpAffx.7896.4.A1_a_at	Potri.003G176900	chalcone synthase (CHS6)	50.9888	2.07E-11
PtpAffx.224485.1.S1_s_at	Potri.005G207600	(MATE) family transporter-related (AtTT12 like [82.6 % identity])	48.93626	4.81E-10
Ptp.6057.1.S1_at	Potri.001G113100	anthocyanidin synthase (ANS2)	41.48594	2.15E-09
PtpAffx.7896.3.S1_a_at	Potri.001G051600	chalcone synthase (CHS3)	39.651	1.80E-11
PtpAffx.37082.1.A1_at	Potri.002G033600	dihydroflavonol 4-reductase (DFR1)	38.50957	1.21E-09
PtpAffx.161181.1.S1_at	Potri.006G178700	cinnamoyl-CoA reductase-like protein (KOG:1502)	35.7962	3.53E-08
Ptp.6711.1.S1_s_at	Potri.014G145100	chalcone synthase (CHS1)	33.89923	3.20E-09
PtpAffx.6065.2.S1_at	Potri.008G116500	leucoanthocyanidin reductase (LAR1)	33.23252	1.56E-09
Ptp.8030.1.S1_at	Potri.014G019200	cytochrome b5	30.92049	2.52E-11
Ptp.3138.2.A1_a_at	Potri.009G133300	UDP-glucose:flavonoid 3-O-glucosyltransferase-like protein (GO:0047213/0080043)	29.8633	1.67E-06
RPTR-Ptp-U57609-2_s_at	---	Reporter	24.25004	1.32E-08
PtpAffx.142603.1.A1_s_at	Potri.013G073300	flavonoid 3'-hydroxylase (F3'H1)	23.07233	1.71E-10
PtpAffx.8131.4.A1_a_at	Potri.006G221800	MYB134	22.59711	1.93E-10
PtpAffx.120325.1.S1_s_at	Potri.013G073300	flavonoid 3'-hydroxylase (F3'H1)	22.02849	3.36E-09
Ptp.4863.1.S1_s_at	Potri.013G073300	flavonoid 3'-hydroxylase (F3'H1)	21.87519	3.27E-10
PtpAffx.18705.2.A1_a_at	Potri.015G050200	leucoanthocyanidin reductase (LAR3)	21.52007	2.85E-07
PtpAffx.5092.1.A1_at	Potri.004G030700	anthocyanidin reductase (ANR1)	20.29951	6.12E-09
Ptp.4458.1.S1_s_at	Potri.T035000	glutathione s-transferase f13	20.29729	7.53E-06
Ptp.5716.1.S1_at	Potri.002G095900	sugar transport protein 5-like	19.893	4.01E-09

Table A-1 continued.

Probeset	Gene Model	Gene Title	Fold change	P-value
PtpAffx.5092.2.S1_a_at	Potri.004G030700	anthocyanidin reductase (ANR1)	19.81152	2.29E-10
Ptp.323.1.S1_s_at	Potri.005G113900	flavanone 3-hydroxylase (F3H6)	19.61242	1.62E-10
Ptp.6753.1.S1_s_at	Potri.006G027000	Serine hydrolase	17.94192	3.58E-08
PtpAffx.8131.6.A1_a_at	Potri.006G221800	MYB134	17.36366	4.79E-10
PtpAffx.127289.1.A1_at	Potri.006G209000	WD40 repeat-containing protein (AtLWD1 like [72.8 % identity])	16.70971	2.10E-07
PtpAffx.25553.1.A1_at	Potri.005G229500	dihydroflavonol 4-reductase (DFR2)	15.60789	2.33E-08
Ptp.4912.2.A1_at	---	---	12.61972	3.65E-08
PtpAffx.160113.1.A1_s_at	Potri.001G006700	acyl-activating enzyme	11.57284	1.87E-09
PtpAffx.7740.2.A1_a_at	Potri.018G100500	cinnamoyl-CoA reductase-like protein (KOG:1502)	11.3066	1.27E-08
Ptp.3754.1.A1_at	Potri.011G097300	transcription elongation factor Elf1 like	10.87061	0.00029069
PtpAffx.157838.2.S1_at	---	---	10.2875	1.06E-05
Ptp.1362.1.S1_at	Potri.001G256900.1	similar to photosystem II 5 kD protein	10.16059	1.27E-05
PtpAffx.6065.3.A1_a_at	Potri.010G129800	leucocyanidin reductase	10.10974	5.35E-09
Ptp.1080.1.S1_s_at	Potri.010G129800	leucoanthocyanidin reductase (LAR2)	9.148366	5.12E-09
PtpAffx.212895.1.S1_at	Potri.015G100600	chitinase 1	8.641889	2.69E-05
Ptp.6632.1.S1_at	Potri.019G130700	cinnamate 4-hydroxylase (C4H1|CYP73A43)	8.199498	3.11E-07
PtpAffx.92231.1.S1_at	---	---	8.148303	1.24E-08
PtpAffx.150025.1.S1_s_at	Potri.013G157900	cinnamate 4-hydroxylase (C4H2|CYP73A42)	7.907291	8.37E-08
Ptp.1080.1.S1_at	Potri.010G129800	leucoanthocyanidin reductase (LAR2)	7.835983	5.49E-08
PtpAffx.156161.1.A1_at	Potri.001G157600	mtn21-like protein	7.723566	1.23E-08
PtpAffx.200035.1.S1_at	Potri.001G006700	acyl-activating enzyme	7.546858	1.92E-09
PtpAffx.94822.1.A1_at	Potri.002G055100	(MATE) family transporter-related (AtTT12 like [81.5 % identity])	7.284551	1.57E-07
PtpAffx.222060.1.S1_at	---	---	7.223664	0.00012774
Ptp.1512.1.S1_s_at	Potri.019G057800	---	7.201737	5.13E-08
Ptp.6057.1.S1_s_at	Potri.001G113100	anthocyanidin synthase (ANS2)	7.027916	5.33E-08

Table A-1 continued.

Probeset	Gene Model	Gene Title	Fold change	P-value
PtpAffx.30128.1.S1_at	Potri.001G140700	cinnamoyl-CoA reductase-like protein * (KOG:1502)	6.969607	1.60E-05
PtpAffx.12443.4.A1_s_at	Potri.013G157900	cinnamate 4-hydroxylase (C4H2|CYP73A42)	6.967922	2.20E-07
Ptp.336.1.S1_at	Potri.019G130700	cinnamate 4-hydroxylase (C4H1|CYP73A43)	6.234563	4.74E-07
PtpAffx.7795.1.A1_at	Potri.019G024600	hypothetical protein	6.00181	0.00771087
PtpAffx.12353.1.A1_at	Potri.018G146100	trans-cinnamate 4-monooxygenase (C4H3)	5.992858	5.54E-05
PtpAffx.7243.1.A1_a_at	Potri.002G212300	---	5.925439	1.98E-06
PtpAffx.12056.3.S1_a_at	Potri.T071600	4-Coumarate:CoA ligase (4CL4)	5.802742	2.97E-07
PtpAffx.133554.1.A1_s_at	Potri.005G167400	similar to acetyl-CoA carboxylase	5.629325	5.15E-08
PtpAffx.46281.1.S1_at	Potri.019G024600	hydrolase activity, acting on ester bonds	5.359818	0.00437883
PtpAffx.249.86.A1_a_at	---	---	5.031714	1.46E-05
PtpAffx.215434.1.S1_at	Potri.017G110500	nad -binding rossmann-fold superfamily protein	4.837022	1.19E-07
PtpAffx.34213.1.S1_at	Potri.013G083600	peroxidase (lignin biosynthesis) (GO:0004601)	4.7162	3.79E-06
PtpAffx.211755.1.S1_at	Potri.014G127500	root hair elongation	4.686047	7.86E-05
PtpAffx.162989.1.S1_at	Potri.009G133300	UDP-glucose:flavonoid 3-O-glucosyltransferase-like protein (GO:0047213/0080043)	4.626936	3.01E-05
PtpAffx.205491.1.S1_at	Potri.005G167400	similar to acetyl-CoA carboxylase	4.48875	1.59E-07
PtpAffx.157838.2.S1_a_at	---	---	4.461568	6.45E-05
PtpAffx.249.102.A1_x_at	Potri.018G114800	Hypothetical protein	4.456062	2.00E-07
PtpAffx.249.83.A1_x_at	Potri.018G114800	hypothetical protein	4.391304	2.21E-07
PtpAffx.104180.1.S1_at	Potri.010G075600	serine protease inhibitor activity	4.238565	9.54E-05
Ptp.4912.1.S1_at	---	---	4.19839	1.08E-06
Ptp.6847.1.S1_a_at	Potri.013G103700	phenylcoumaran benzylic ether reductase -like protein	4.190316	0.00038945
PtpAffx.4850.1.A1_s_at	Potri.010G213000	chalcone isomerase (CHI1)	4.163411	2.73E-08
Ptp.6847.2.S1_s_at	Potri.013G103700	phenylcoumaran benzylic ether reductase -like protein	4.108134	0.00028581
PtpAffx.159543.1.A1_at	Potri.006G101100	similar to putative ANS in *Ricinus communis* (86.7% identity)	4.057872	0.00011175
PtpAffx.160901.1.S1_s_at	Potri.004G057700	cysteine protease	4.001451	4.20E-05
PtpAffx.5685.1.S1_x_at	Potri.005G045300	hypothetical protein	3.945677	2.21E-07

Table A-1. continued

Probeset	Gene Model	Gene Title	Fold change	P-value
PtpAffx.211115.1.S1_at	Potri.013G069200	(MATE) family transporter-related (PANTHER:11206)	3.945352	6.92E-07
PtpAffx.224943.1.S1_s_at	Potri.019G036400	MYB061	3.897258	2.63E-07
Ptp.6482.1.S1_s_at	Potri.002G092700	similar to acetyl-CoA carboxylase	3.874861	5.92E-06
Ptp.1442.1.S1_x_at	---	---	3.86488	2.47E-05
PtpAffx.215434.1.S1_s_at	Potri.017G110500	Flavonol reductase/cinnamoyl-CoA reductase	3.851884	6.89E-06
Ptp.5773.1.S1_at	Potri.016G111000	probable anion transporter chloroplastic-like	3.845173	3.16E-06
PtpAffx.108534.1.A1_at	Potri.014G145100	chalcone synthase (CHS1)	3.767658	2.93E-07
PtpAffx.137746.1.S1_at	Potri.004G088100	MYB182	3.744134	6.71E-05
PtpAffx.5685.1.S1_a_at	Potri.005G045300	hypothetical protein	3.62709	1.37E-07
PtpAffx.48329.2.S1_at	---	---	3.613938	0.00049483
PtpAffx.47237.2.A1_a_at	Potri.001G181700	---	3.606718	6.23E-07
PtpAffx.201822.1.S1_at	Potri.002G072600	bidirectional sugar transporter sweet1-like	3.570947	1.21E-05
Ptp.6610.1.S1_at	---	---	3.467474	5.07E-06
PtpAffx.224746.1.S1_s_at	Potri.011G066800	cysteine protease	3.433912	9.76E-06
PtpAffx.256.2.S1_at	Potri.005G208600	bHLH type protein (TT8 like [50.5 % identity])	3.407876	8.15E-07
PtpAffx.161412.1.S1_at	Potri.004G054600	phospholipase a1-iigamma-like	3.388368	6.57E-05
PtpAffx.19596.1.S1_at	---	---	3.299199	9.04E-07
PtpAffx.65824.1.A1_at	Potri.001G001600	pyruvate kinase	3.27352	5.00E-07
PtpAffx.224650.1.S1_s_at	Potri.010G114000	MYB165	3.237289	0.00073541
PtpAffx.50866.2.A1_s_at	Potri.004G041800	allantoinase	3.233605	0.00063159
PtpAffx.249.498.A1_at	Potri.018G114800	hypothetical protein	3.226604	5.81E-05
PtpAffx.112345.1.S1_at	Potri.002G059200	outward rectifying potassium channel	3.206058	6.00E-05
PtpAffx.160504.1.A1_s_at	Potri.002G092700	similar to acetyl-CoA carboxylase	3.182847	7.71E-06
PtpAffx.224878.1.S1_at	Potri.014G100800	MYB201	3.168808	1.03E-06
PtpAffx.225583.1.S1_s_at	Potri.014G033400	Patatin-like phospholipase	3.136809	2.85E-06

Table A-1. continued

Probeset	Gene Model	Gene Title	Fold change	P-value
Ptp.8024.1.S1_at	Potri.005G208600	bHLH type protein (TT8 like [50.5 % identity])	3.114279	7.06E-05
PtpAffx.157838.1.S1_at	---	---	3.0776	1.37E-05
PtpAffx.205684.1.S1_at	Potri.005G208600	bHLH type protein (TT8 like [50.5 % identity])	3.059646	4.67E-05
PtpAffx.83148.1.S1_s_at	Potri.010G145800	ATP citrate (pro-S)-lyase	3.057902	9.50E-06
PtpAffx.201776.1.S1_at	Potri.002G064100	tryptophan aminotransferase-related protein 4-like	3.043018	0.01069032
PtpAffx.83148.2.S1_s_at	Potri.010G145800	ATP citrate (pro-S)-lyase	3.005611	4.02E-05
PtpAffx.210644.1.S1_at	Potri.012G096800	cytochrome P450	2.961663	0.00589359
PtpAffx.30743.1.A1_a_at	Potri.013G031900	hypothetical protein	2.9584	2.57E-06
PtpAffx.101397.1.A1_at	Potri.013G069200	(MATE) family transporter-related (PANTHER:11206)	2.943119	5.08E-06
PtpAffx.86545.1.S1_at	Potri.010G223700	kinesin like protein	2.891662	1.75E-05
PtpAffx.43288.1.A1_at	Potri.004G054600	phospholipase a1-iigamma-like	2.869454	0.00022971
PtpAffx.204062.1.S1_at	Potri.004G105000	nad -binding rossmann-fold superfamily protein	2.852432	9.36E-06
Ptp.2571.1.A1_at	Potri.002G092700	similar to acetyl-CoA carboxylase	2.808251	1.73E-05
Ptp.5281.1.S1_s_at	Potri.008G149200	alcohol dehydrogenase (NADP+) activity	2.803008	0.00131177
PtpAffx.2158.1.S1_at	---	---	2.757472	0.0002098
Ptp.6940.1.S1_at	---	---	2.754735	0.01996794
PtpAffx.158517.1.S1_at	Potri.005G004900	ATP citrate (pro-S)-lyase	2.748815	2.74E-06
PtpAffx.249.338.A1_a_at	---	---	2.748724	1.26E-05
PtpAffx.159543.2.A1_a_at	Potri.006G101100	oxidoreductase, 2OG-Fe(II) oxygenase family protein	2.73983	0.00180596
PtpAffx.18588.1.S1_at	Potri.013G143300	ribosomal protein l2	2.731795	3.05E-05
PtpAffx.216838.1.S1_at	Potri.015G018300	hypothetical protein	2.71574	0.00013475
PtpAffx.213439.1.S1_at	Potri.016G075800	WD40 repeat-containing protein (AtLWD1 like [67.5 % identity])	2.690502	0.0002305
Ptp.3045.1.S1_at	Potri.004G188100	arogenate/prephenate dehydratase	2.687724	0.00014765
PtpAffx.57339.1.A1_at	Potri.016G052200	---	2.671401	5.80E-05
PtpAffx.2048.1.S1_a_at	Potri.003G131600	Hypothetical protein	2.666177	0.00012852

Table A-1. continued

Probeset	Gene Model	Gene Title	Fold change	P-value
Ptp.5023.1.S1_s_at	Potri.001G293400	adenylyl-sulfate kinase chloroplastic-like	2.657122	0.02051256
Ptp.4670.1.S1_s_at	Potri.008G105300	ATP citrate (pro-S)-lyase	2.642401	2.27E-05
PtpAffx.161399.1.S1_at	Potri.012G057500	enolase	2.635902	0.00231887
PtpAffx.76291.2.A1_a_at	Potri.002G059200	outward rectifying potassium channel	2.614822	0.00568046
Ptp.6352.1.A1_s_at	Potri.002G242500	cytochrome b5	2.600035	1.15E-05
PtpAffx.4.4.A1_a_at	Potri.009G141800	Chitinase related	2.594996	5.14E-06
Ptp.3045.1.S1_s_at	Potri.004G188100	arogenate/prephenate dehydratase	2.594379	5.40E-05
Ptp.2979.1.S1_at	Potri.010G124700	---	2.593676	0.02169766
PtpAffx.2048.1.S1_x_at	Potri.003G131600	hypothetical protein	2.590999	8.12E-05
PtpAffx.225544.1.S1_s_at	Potri.016G031000	p-coumaroyl shikimate 3'- hydroxylase (C3H3)	2.560835	8.08E-06
Ptp.2358.1.S1_at	Potri.012G057500	enolase	2.557254	0.00127478
PtpAffx.106733.1.A1_at	Potri.002G057700	enoyl-CoA hydratase/isomerase family protein	2.555981	5.47E-06
PtpAffx.154968.1.S1_at	Potri.015G125100	membrane steroid-binding protein 2-like	2.538523	7.17E-06
PtpAffx.33535.2.A1_at	Potri.002G029300	purple acid phosphatase	2.527966	3.72E-05
PtpAffx.224602.1.S1_at	Potri.008G128500	MYB194	2.503994	0.00048538
PtpAffx.18295.2.A1_s_at	Potri.010G055400	Pt-GAPDH.3	2.503507	8.35E-06
PtpAffx.138813.1.A1_s_at	Potri.013G004400	ATP citrate (pro-S)-lyase	2.480356	8.81E-05
PtpAffx.30278.1.S1_at	Potri.003G061200	hypothetical protein	2.473269	0.00203166
PtpAffx.163838.2.S1_a_at	Potri.008G125900	fructose-bisphosphate aldolase	2.471575	4.93E-05
PtpAffx.224381.1.S1_s_at	Potri.004G071000	quercetin O-glucosyltransferase-like protein (GO:0080045)	2.466033	0.00839078
PtpAffx.2048.2.S1_a_at	Potri.003G131600	hypothetical protein	2.44478	0.00026489
PtpAffx.220100.1.S1_at	Potri.001G133000	anthocyanidin 3-O-glucosyltransferase activity	2.440327	0.00116377
PtpAffx.208112.1.S1_s_at	Potri.008G143600	hypothetical protein	2.435925	7.23E-05
PtpAffx.148416.1.S1_s_at	Potri.001G214500	adipocyte plasma membrane-associated	2.413321	0.0007896

Table A-1. continued

Probeset	Gene Model	Gene Title	Fold change	P-value
PtpAffx.39332.1.A1_s_at	Potri.004G057700	cysteine protease	2.412716	7.02E-06
PtpAffx.55376.1.S1_at	Potri.014G143200	peroxidase (lignin biosynthesis) (GO:0004601)	2.409356	0.00439342
PtpAffx.225544.1.S1_x_at	Potri.016G031100	p-coumaroyl shikimate 3'- hydroxylase (CYP98A23)	2.381891	1.49E-05
PtpAffx.47237.3.S1_a_at	Potri.003G054000	hypothetical protein	2.366469	0.00010976
PtpAffx.53470.1.S1_at	Potri.006G009200	fumarylacetoacetase	2.359404	7.81E-05
Ptp.1405.1.S1_at	Potri.014G128500	blue-light photoreceptor phr2-like	2.35134	0.00014237
Ptp.5738.2.S1_s_at	Potri.006G050400	hypothetical protein	2.349355	0.03853812
PtpAffx.30743.5.A1_at	Potri.013G031900	hypothetical protein	2.339317	8.75E-06
PtpAffx.3539.1.S1_a_at	Potri.001G415100	hypothetical protein	2.338787	0.00118429
PtpAffx.28572.2.S1_at	Potri.001G001600	---	2.33414	1.23E-06
Ptp.2303.1.S1_at	Potri.001G214500	adipocyte plasma membrane-associated	2.327823	0.00060105
PtpAffx.6463.1.S1_at	Potri.004G188100	arogenate/prephenate dehydratase	2.312865	3.83E-05
PtpAffx.24245.1.A1_at	Potri.006G173500	---	2.297136	0.00069208
PtpAffx.202703.1.S1_at	Potri.002G252400	hypothetical protein	2.286861	0.04692555
PtpAffx.161644.1.S1_s_at	Potri.001G083900	cytochrome P450	2.285017	0.0244352
PtpAffx.144130.1.S1_s_at	Potri.001G083900	cytochrome P450	2.256183	0.03075088
Ptp.2573.1.S1_at	Potri.015G034700	glutamine synthetase	2.255553	4.12E-06
PtpAffx.75249.1.A1_at	Potri.001G161400	autoinhibited H+ ATPase	2.252031	1.74E-05
PtpAffx.138656.1.A1_at	Potri.018G152100	aquaporin, MIP family, TIP subfamily	2.217306	0.0002432
PtpAffx.208118.1.S1_at	Potri.008G143600	hypothetical protein	2.208126	0.00111715
PtpAffx.204049.1.S1_s_at	Potri.004G071000	quercetin O-glucosyltransferase-like protein (GO:0080045)	2.207924	0.00272634
PtpAffx.154968.1.S1_s_at	Potri.015G125100	membrane steroid-binding protein 2-like	2.205416	1.16E-05
PtpAffx.80486.1.A1_at	Potri.001G212900	serine-type peptidase activity	2.200679	0.01018903
PtpAffx.25980.1.S1_at	Potri.012G032700	flavonol 3-sulfotransferase-like protein (GO:0047364)	2.191155	0.01206353
PtpAffx.224381.1.S1_at	Potri.004G069600	quercetin O-glucosyltransferase-like protein (GO:0080045)	2.177728	0.03405199
PtpAffx.82668.1.A1_at	Potri.013G157400	cytochrome P450	2.168596	0.03448489

Table A-1. continued

Probeset	Gene Model	Gene Title	Fold change	P-value
PtpAffx.52548.1.A1_at	Potri.009G148800	arogenate/prephenate dehydratase	2.030137	0.00049645
Ptp.4667.1.S1_s_at	Potri.001G059900	glucose-6-phosphate 1-dehydrogenase	2.029568	0.01073479
Ptp.1017.1.A1_at	Potri.016G062700	Predicted hydrolase/acyltransferas	2.02316	0.00024495
PtpAffx.254.1.S1_s_at	Potri.014G115200	hypothetical protein	2.018906	8.70E-05
PtpAffx.12443.3.A1_at	Potri.013G157900	cinnamate 4-hydroxylase (C4H2|CYP73A42)	2.018016	0.00143483
PtpAffx.51055.1.S1_at	Potri.003G215200	calcium calmodulin-dependent protein kinase kinase	2.015053	4.16E-05
PtpAffx.2542.1.S1_at	Potri.003G158500	Hypothetical protein	2.011455	1.21E-05
PtpAffx.23401.1.S1_at	Potri.010G177700	nadp-dependent alkenal double bond reductase p2-like	2.009995	4.03E-05
Ptp.323.1.S1_at	Potri.005G113700	flavonone 3-hydroxylase (F3H3)	2.000638	0.03003307

Table A-2. Complete list of down-regulated genes (fold change < 0.5) in 353 MYB115 overexpressing plants (Line 5). Genes are listed from lowest to highest fold change. Fold change is relative to 353 WT plants (n = 3 plants). All functional annotations are from Affymetrix annotation files with the exception of the probes that presented in Table 2-4.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.14459.1.S1_at	Potri.010G083600	homocysteine s-methyltransferase	0.06210493	0.02739958
PtpAffx.218344.1.S1_s_at	Potri.002G240700	hypothetical protein	0.08622309	0.0338121
PtpAffx.16402.2.S1_a_at	Potri.004G093800	hypothetical protein	0.08805288	7.79E-05
PtpAffx.215126.1.S1_s_at	Potri.003G088100	mlp-like protein	0.08874642	5.77E-06
PtpAffx.69240.1.A1_at	Potri.015G132900	hypothetical protein	0.1105906	2.28E-05
Ptp.370.1.S1_at	---	---	0.1233098	5.36E-05
Ptp.5064.1.S1_at	---	---	0.1253671	2.29E-06
PtpAffx.16402.2.S1_at	---	---	0.1257025	0.000116559
PtpAffx.43257.1.S1_s_at	Potri.010G068900	peroxisomal membrane protein pmp22-like	0.1263541	1.76E-05
PtpAffx.2286.4.S1_a_at	Potri.011G140700	---	0.1272973	1.41E-06
PtpAffx.32180.1.A1_at	Potri.001G395700	hypothetical protein	0.1339939	5.84E-07
Ptp.4381.2.A1_a_at	Potri.003G066800	flavonoid 3'5'-hydroxylase-like (CYP92A20v2) (GO:0033772)	0.1357423	3.28E-05
PtpAffx.114060.1.A1_at	Potri.006G039500	---	0.1362252	0.02909995
PtpAffx.163963.1.S1_s_at	Potri.001G167800	cytochrome P450 (flavonoid 3',5'-hydroxylase activity)	0.1372156	0.000558362
PtpAffx.12059.1.S1_at	Potri.008G106400	alpha-dioxygenase	0.1414672	0.000575581
PtpAffx.225268.1.S1_s_at	Potri.001G395900	hypothetical protein	0.1422247	2.58E-07
PtpAffx.57754.1.A1_at	Potri.013G100400	---	0.1430283	0.00045757
PtpAffx.109440.1.S1_at	Potri.004G146000	2-oxoglutarate-dependent dioxygenase	0.1442555	0.04148853
PtpAffx.77318.1.S1_at	Potri.009G142200	class i chitinase	0.1446768	0.000183428
Ptp.5757.1.S1_x_at	Potri.015G132900	hypothetical protein	0.1517323	5.13E-05
PtpAffx.308.5.S1_s_at	Potri.003G066800	cytochrome P450 (flavonoid 3',5'-hydroxylase activity)	0.1529059	0.000131372
PtpAffx.203472.1.S1_s_at	Potri.003G181800	hypothetical protein	0.1540881	0.004916353
PtpAffx.3594.1.S1_at	Potri.002G018300	cinnamyl alcohol dehydrogenase-like protein	0.1573671	8.41E-05
Ptp.5757.1.S1_at	Potri.015G132900	hypothetical protein	0.1580132	9.07E-05

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.77318.1.S1_x_at	Potri.009G142200	hypothetical protein	0.163445	0.000181322
PtpAffx.103190.1.S1_at	Potri.003G066400	flavonoid 3'5'-hydroxylase-like (CYP92A20v2) (GO:0033772)	0.1649159	0.000283986
PtpAffx.1767.1.S1_a_at	Potri.010G000600	hypothetical protein	0.1726796	0.000580499
PtpAffx.129663.1.S1_at	Potri.009G107700	2-oxoglutarate-dependent dioxygenase	0.173197	0.01857348
PtpAffx.209673.1.S1_at	Potri.011G023300	hypothetical protein	0.176965	0.04896658
PtpAffx.3847.3.A1_at	Potri.001G167900	flavonoid 3'5'-hydroxylase-like (CYP92A17v1) (GO:0033772)	0.1801926	5.85E-05
Ptp.3709.1.S1_at	Potri.016G083500	Ribulose-bisphosphate carboxylase small chain	0.1816271	0.01401197
Ptp.4101.1.S1_s_at	Potri.011G140400	---	0.1831508	0.000481424
Ptp.4154.1.S1_s_at	Potri.003G066400	cytochrome P450	0.185724	7.51E-05
PtpAffx.211140.1.S1_s_at	Potri.013G074500	hypothetical protein	0.1934102	6.38E-05
Ptp.2463.2.S1_at	Potri.013G041600	hypothetical protein	0.1939198	0.000868945
PtpAffx.84185.1.S1_a_at	Potri.001G003100	cytochrome P450	0.1940096	0.01894785
PtpAffx.77318.2.S1_at	Potri.004G182000	Hypothetical protein	0.2011004	0.00016244
PtpAffx.249.47.A1_s_at	Potri.001G102400	---	0.2025128	0.001171598
PtpAffx.15128.1.A1_at	Potri.016G031500	hypothetical protein	0.2030832	9.67E-05
Ptp.6751.1.A1_s_at	Potri.016G014500	hypothetical protein	0.2057859	1.43E-05
Ptp.3038.1.A1_at	Potri.004G082000	Hydrolase	0.2070651	0.000414852
Ptp.6831.1.S1_s_at	Potri.016G014500	hypothetical protein	0.2074018	0.000112512
Ptp.5132.1.S1_at	Potri.001G255100	hypothetical protein	0.2128047	0.000165411
Ptp.4185.1.S1_at	Potri.006G275900	MYB097	0.2146141	2.93E-05
PtpAffx.249.361.A1_at	Potri.014G019700	hypothetical protein	0.2147372	0.000106309
PtpAffx.69240.1.A1_s_at	Potri.015G132900	hypothetical protein	0.2158598	0.000642675
PtpAffx.29151.1.S1_a_at	Potri.009G103800	hypothetical protein	0.2175704	2.21E-05
PtpAffx.14425.2.S1_at	Potri.001G321600	hypothetical protein	0.2193854	0.03484923

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.77318.1.S1_x_at	Potri.009G142200	hypothetical protein	0.163445	0.000181322
PtpAffx.103190.1.S1_at	Potri.003G066400	flavonoid 3'5'-hydroxylase-like (CYP92A20v2) (GO:0033772)	0.1649159	0.000283986
PtpAffx.1767.1.S1_a_at	Potri.010G000600	hypothetical protein	0.1726796	0.000580499
PtpAffx.129663.1.S1_at	Potri.009G107700	2-oxoglutarate-dependent dioxygenase	0.173197	0.01857348
PtpAffx.209673.1.S1_at	Potri.011G023300	hypothetical protein	0.176965	0.04896658
PtpAffx.3847.3.A1_at	Potri.001G167900	flavonoid 3'5'-hydroxylase-like (CYP92A17v1) (GO:0033772)	0.1801926	5.85E-05
Ptp.3709.1.S1_at	Potri.016G083500	Ribulose-bisphosphate carboxylase small chain	0.1816271	0.01401197
Ptp.4101.1.S1_s_at	Potri.011G140400	---	0.1831508	0.000481424
Ptp.4154.1.S1_s_at	Potri.003G066400	cytochrome P450	0.185724	7.51E-05
PtpAffx.211140.1.S1_s_at	Potri.013G074500	hypothetical protein	0.1934102	6.38E-05
Ptp.2463.2.S1_at	Potri.013G041600	hypothetical protein	0.1939198	0.000868945
PtpAffx.84185.1.S1_a_at	Potri.001G003100	cytochrome P450	0.1940096	0.01894785
PtpAffx.77318.2.S1_at	Potri.004G182000	Hypothetical protein	0.2011004	0.00016244
PtpAffx.249.47.A1_s_at	Potri.001G102400	---	0.2025128	0.001171598
PtpAffx.15128.1.A1_at	Potri.016G031500	hypothetical protein	0.2030832	9.67E-05
Ptp.6751.1.A1_s_at	Potri.016G014500	hypothetical protein	0.2057859	1.43E-05
Ptp.3038.1.A1_at	Potri.004G082000	Hydrolase	0.2070651	0.000414852
Ptp.6831.1.S1_s_at	Potri.016G014500	hypothetical protein	0.2074018	0.000112512
Ptp.5132.1.S1_at	Potri.001G255100	hypothetical protein	0.2128047	0.000165411
Ptp.4185.1.S1_at	Potri.006G275900	MYB097	0.2146141	2.93E-05
PtpAffx.249.361.A1_at	Potri.014G019700	hypothetical protein	0.2147372	0.000106309
PtpAffx.69240.1.A1_s_at	Potri.015G132900	hypothetical protein	0.2158598	0.000642675
PtpAffx.29151.1.S1_a_at	Potri.009G103800	hypothetical protein	0.2175704	2.21E-05
PtpAffx.14425.2.S1_at	Potri.001G321600	hypothetical protein	0.2193854	0.03484923

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
Ptp.4252.1.S1_at	Potri.001G240300	2-oxoglutarate-dependent dioxygenase	0.2201536	2.44E-05
PtpAffx.209240.1.S1_at	Potri.010G181800	amino acid transporter	0.2212405	0.007390773
PtpAffx.122905.1.S1_at	Potri.016G014400	hypothetical protein	0.2213142	1.92E-05
PtpAffx.202651.1.S1_at	Potri.002G240700	hypothetical protein	0.2229981	0.03922669
Ptp.6598.1.S1_at	Potri.006G141400	flavonoid 3'5'-hydroxylase-like (CYP92A24) (GO:0033772)	0.2257996	5.85E-05
Ptp.5947.1.S1_at	Potri.014G019700	hypothetical protein	0.2262959	0.000138564
Ptp.5106.1.S1_s_at	Potri.018G098200	hypothetical protein	0.2272256	0.000238321
PtpAffx.249.35.S1_s_at	Potri.018G089400	hypothetical protein	0.2280753	0.02673312
Ptp.3143.2.S1_a_at	Potri.006G109100	hypothetical protein	0.2281896	3.24E-05
Ptp.1163.1.A1_at	Potri.003G047200	Hypothetical protein	0.2285845	0.000266394
Ptp.3143.1.S1_at	Potri.006G109100	hypothetical protein	0.2289357	4.06E-05
Ptp.6598.1.S1_s_at	Potri.018G051300	flavonoid 3'5' hydroxylase -like (CYP92A25) (GO:0033772)	0.2289472	2.82E-05
Ptp.5321.1.S1_at	Potri.001G255100	hypothetical protein	0.2337953	0.000149068
PtpAffx.35408.1.S1_at	Potri.007G141000	hypothetical protein	0.235259	0.000346053
Ptp.2463.1.S1_s_at	Potri.013G041600	hypothetical protein	0.2367718	0.000685575
PtpAffx.225573.1.S1_at	Potri.003G028700	hypothetical protein	0.2384242	5.36E-05
PtpAffx.6696.2.S1_at	Potri.013G074500	PtACT47*	0.2396261	0.000196127
PtpAffx.60092.1.S1_at	Potri.006G275900	MYB097	0.2408461	0.000147211
PtpAffx.7783.1.S1_a_at	Potri.017G028100	cytochrome P450	0.2418408	0.000601298
PtpAffx.117367.1.A1_s_at	Potri.010G181800	amino acid transporter	0.2465105	0.01939872
Ptp.5070.1.S1_s_at	Potri.009G083300	hypothetical protein	0.2467849	0.001895392
PtpAffx.14425.1.S1_a_at	Potri.017G060000	hypothetical protein	0.2481145	0.04173925
PtpAffx.64603.1.S1_at	Potri.012G033900	Hypothetical protein	0.2518477	0.001237235
PtpAffx.64184.1.S1_at	Potri.018G051300	flavonoid 3'5' hydroxylase -like (CYP92A25) (GO:0033772)	0.2543887	0.000446657
PtpAffx.43573.2.A1_at	Potri.008G159000	hypothetical protein	0.2582326	0.02469488

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.213178.1.S1_at	Potri.016G020000	hypothetical protein	0.2647937	3.65E-06
PtpAffx.4462.1.S1_at	Potri.010G142800	hypothetical protein	0.2652075	0.000864712
PtpAffx.203341.1.S1_s_at	Potri.003G150700	AP2/ERF domain-containing transcription factor	0.2659575	0.002664115
PtpAffx.72041.1.S1_at	Potri.001G240300	triose-phosphate isomerase	0.26612	0.000219965
Ptp.3040.1.A1_at	Potri.011G002500	(MATE) family transporter-related PANTHER:11206	0.2692467	0.000128532
PtpAffx.35542.1.S1_at	Potri.018G088100	---	0.2709715	8.45E-06
Ptp.2209.1.S1_at	Potri.006G141400	Hypothetical protein	0.2740487	2.21E-05
PtpAffx.32468.1.S1_at	Potri.001G107800	hypothetical protein	0.2784603	0.000198789
PtpAffx.2271.1.S1_s_at	Potri.T094200	hypothetical protein	0.2800086	0.001221114
PtpAffx.4337.2.S1_at	Potri.011G049900	---	0.2832657	0.000184812
Ptp.999.1.A1_at	Potri.002G091500	Hypothetical protein	0.2870782	0.01600978
PtpAffx.79787.1.S1_at	Potri.004G192600	hypothetical protein	0.2882653	3.46E-06
PtpAffx.1195.1.S1_at	Potri.003G194700	hypothetical protein	0.2952124	2.13E-05
PtpAffx.572.3.S1_at	Potri.010G129200	AP2 ERF and B3 domain-containing (RAV1)	0.2968766	5.64E-05
PtpAffx.23456.1.S1_at	Potri.005G188300	---	0.3022281	2.15E-05
PtpAffx.29151.1.S1_at	Potri.009G103800	---	0.30304	1.92E-05
PtpAffx.25029.1.S1_at	Potri.009G104600	hypothetical protein	0.3031148	4.73E-05
PtpAffx.225101.1.S1_x_at	Potri.T094800	hypothetical protein	0.3031971	0.001605064
PtpAffx.203518.1.S1_s_at	Potri.003G205900	hypothetical protein	0.3045647	0.001062029
Ptp.6838.2.S1_a_at	Potri.001G221900	hypothetical protein	0.3078576	0.01332349
Ptp.2164.1.S1_s_at	Potri.018G019800	hypothetical protein	0.3079743	0.000235367
Ptp.4069.1.A1_at	Potri.019G093800	hypothetical protein	0.3083075	0.000571625
PtpAffx.205219.1.S1_at	Potri.005G085200	hypothetical protein	0.3102491	7.19E-05
Ptp.4222.1.S1_at	Potri.T094200	hypothetical protein	0.3106688	0.002976543
PtpAffx.224677.1.S1_at	Potri.011G002400	(MATE) family transporter-related PANTHER:11206	0.3143611	3.59E-05

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.77318.3.S1_at	Potri.004G182000	hypothetical protein	0.3158674	1.13E-05
PtpAffx.24863.2.A1_at	Potri.006G279000	Hypothetical protein	0.3164735	0.008226816
Ptp.5125.1.S1_s_at	Potri.009G118100	phenylcoumaran benzylic ether reductase 3	0.3178894	0.03158379
PtpAffx.129457.1.A1_at	---	---	0.3195885	0.006052306
PtpAffx.151956.1.A1_at	Potri.003G181600	hypothetical protein	0.3204664	0.002092016
PtpAffx.5087.1.A1_at	Potri.009G118100	phenylcoumaran benzylic ether reductase 3	0.3235475	0.01999068
Ptp.1093.1.S1_at	Potri.018G002900	---	0.324446	0.005669496
PtpAffx.202716.1.S1_at	Potri.002G255700	hypothetical protein	0.3267328	2.75E-05
PtpAffx.17234.1.S1_at	Potri.008G212500	Hypothetical protein	0.3289063	0.000569038
PtpAffx.24863.1.A1_at	Potri.018G002900	hypothetical protein	0.3315104	0.0112769
Ptp.4093.1.S1_at	Potri.019G043600	PtACT49[†]	0.3321203	0.000169779
PtpAffx.107976.1.A1_a_at	Potri.016G083500	hypothetical protein	0.3323183	0.01849764
Ptp.5125.1.S1_at	Potri.009G118000	phenylcoumaran benzylic ether reductase 3	0.3341359	0.03412937
PtpAffx.216729.1.S1_s_at	Potri.003G007400	cytochrome P450	0.3355203	0.001164319
PtpAffx.220470.1.S1_x_at	Potri.012G027400	hypothetical protein	0.3400855	0.000209259
PtpAffx.203072.1.S1_s_at	Potri.007G056900	oligopeptide transporter OPT family	0.3438312	0.02756819
PtpAffx.104683.1.S1_at	Potri.013G149200	MYB049	0.3439324	1.43E-06
Ptp.5321.1.S1_s_at	Potri.001G255100	hypothetical protein	0.3455952	2.95E-05
Ptp.5332.1.S1_at	Potri.014G017300	hypothetical protein	0.3475807	0.04604801
Ptp.4542.1.S1_at	Potri.014G092800	hypothetical protein	0.3478042	0.007509009
PtpAffx.24863.1.A1_s_at	Potri.018G002900	---	0.3493475	0.005934406
PtpAffx.160390.2.A1_at	Potri.015G002800	flavanone 3-hydroxylase-like protein (GO:0001666)	0.3542198	0.000641094
Ptp.2463.1.S1_a_at	Potri.013G041600	hypothetical protein	0.3546314	0.000572563
PtpAffx.2179.2.S1_at	Potri.005G069500	NAC domain trancription factor (NAC043)	0.357166	0.01988018
Ptp.6689.1.S1_at	Potri.006G029400	hypothetical protein	0.3573112	9.45E-07

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
Ptp.6689.1.S1_s_at	Potri.006G030000	---	0.3578904	3.47E-06
PtpAffx.151041.1.S1_s_at	Potri.004G055900	hypothetical protein	0.3625019	0.000176892
PtpAffx.107445.1.A1_at	Potri.003G088300	---	0.363026	7.18E-07
Ptp.5112.2.S1_a_at	Potri.013G153400	hypothetical protein	0.3677016	6.67E-07
PtpAffx.160390.3.A1_s_at	Potri.012G006300	oxidoreductase, 2OG-Fe(II) oxygenase family protein	0.369313	3.87E-05
PtpAffx.225113.1.S1_s_at	Potri.011G050100	hypothetical protein	0.3706698	0.000859113
PtpAffx.97010.1.S1_s_at	Potri.011G050100	hypothetical protein	0.3712084	0.000704086
PtpAffx.2179.1.S1_at	Potri.007G099400	NAC domain trancription factor (NAC045)	0.3718548	0.01120048
Ptp.3323.1.S1_s_at	Potri.003G150800	AP2/ERF domain-containing transcription factor	0.3728428	0.001773175
PtpAffx.6696.2.S1_s_at	Potri.T131600	hypothetical protein	0.3749542	2.40E-05
PtpAffx.208729.1.S1_at	Potri.010G074800	hypothetical protein	0.3788376	0.002737123
PtpAffx.71886.2.S1_at	Potri.012G091100	methionine sulfoxide reductase type	0.3794585	0.01161045
PtpAffx.18575.1.S1_s_at	Potri.014G179100	cytochrome P450 probable ent-kaurenoic acid oxidase	0.3794708	0.000314817
PtpAffx.201906.1.S1_at	Potri.002G092500	sulfate/bicarbonate/oxalate exchanger and transporter sat-1	0.380591	0.01924476
PtpAffx.119354.2.A1_s_at	Potri.015G147600	(MATE) family transporter-related (AtFFT like [68.0 % identity])	0.3806387	6.65E-06
Ptp.6647.1.S1_at	Potri.004G168000	Hypothetical protein	0.3834021	0.02051382
PtpAffx.5133.2.S1_s_at	Potri.008G218500	hypothetical protein	0.3837938	0.002434032
PtpAffx.225884.1.S1_at	Potri.005G013100	hypothetical protein	0.3842821	0.005671789
PtpAffx.17713.1.A1_at	Potri.005G170800	---	0.3842985	0.01944635
PtpAffx.454.12.S1_s_at	Potri.017G147200	hypothetical protein	0.385051	0.02535039
PtpAffx.208930.1.S1_s_at	Potri.010G121200	hypothetical protein	0.3861988	0.000514094
PtpAffx.221539.1.S1_at	---	---	0.3862303	0.00309052
PtpAffx.43573.1.A1_s_at	Potri.008G159000	hypothetical protein	0.3865614	0.02155405
PtpAffx.58683.1.S1_at	---	---	0.3878674	0.002538951
PtpAffx.203891.1.S1_at	Potri.004G043500	hypothetical protein	0.3882627	0.01386256

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.149231.3.S1_at	---	---	0.3909566	0.04309317
PtpAffx.225769.1.S1_s_at	Potri.005G056500	hypothetical protein	0.3910413	0.001078336
PtpAffx.205860.1.S1_at	Potri.005G248900	hypothetical protein	0.392676	0.000262425
Ptp.6743.1.A1_at	Potri.002G216400	acetyl-CoA C-acyltransferase	0.3929968	9.62E-05
Ptp.2730.1.S1_at	Potri.019G055200	hypothetical protein	0.3951778	0.004168419
PtpAffx.119354.1.A1_at	Potri.012G144900	(MATE) family transporter-related (AtFFT like [71.8 % idenitity])	0.3980224	8.72E-05
PtpAffx.60684.1.S1_at	---	---	0.3993356	0.004856131
PtpAffx.225884.1.S1_x_at	Potri.005G013100	hypothetical protein	0.4000252	0.004562735
PtpAffx.14494.1.A1_at	Potri.019G015300	hypothetical protein	0.4014969	0.002071218
PtpAffx.37785.1.A1_s_at	Potri.001G451900	hypothetical protein	0.4049553	0.00056388
Ptp.1239.1.S1_at	Potri.001G075600	---	0.4059875	0.000146423
Ptp.6123.1.S1_at	Potri.005G170200	---	0.4066081	0.000534929
Ptp.2176.1.S1_at	Potri.002G020700	hypothetical protein	0.4072804	0.01301041
PtpAffx.36779.2.A1_a_at	Potri.003G177600	hypothetical protein	0.4074203	0.02054916
PtpAffx.27218.1.A1_at	Potri.002G216400	acetyl-CoA C-acyltransferase	0.4087737	0.000637182
PtpAffx.120103.1.S1_at	Potri.001G134500	---	0.4094628	0.000296818
Ptp.3851.1.A1_at	Potri.017G118400	hypothetical protein	0.4094842	0.000442428
Ptp.1035.1.S1_at	Potri.009G067800	Hypothetical protein	0.4106612	0.001670309
Ptp.5925.1.S1_s_at	Potri.T131700	Hypothetical protein	0.4114265	2.10E-05
Ptp.4107.2.S1_a_at	Potri.007G133700	---	0.4116893	3.34E-05
PtpAffx.94172.1.A1_at	Potri.006G076500	Hypothetical protein	0.411754	0.002582849
PtpAffx.203834.1.S1_s_at	Potri.004G056800	hypothetical protein	0.4126723	0.005919432
PtpAffx.572.2.S1_at	Potri.010G129200	AP2 ERF and B3 domain-containing (RAV1)	0.4130817	0.006574912
PtpAffx.22767.1.S1_a_at	Potri.001G224000	Gamma carbonic anhydrase	0.4141914	0.000119384
Ptp.6743.1.A1_s_at	Potri.002G216400	acetyl-CoA C-acyltransferase	0.4148859	8.26E-05

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.203527.1.S1_at	Potri.003G204000	autoinhibited calcium ATPase	0.4160685	0.001980938
PtpAffx.75508.1.A1_at	Potri.016G036600	hypothetical protein	0.4172118	0.000750844
PtpAffx.35738.1.A1_at	---	---	0.4180244	0.02143712
Ptp.6936.1.S1_at	Potri.015G042000	---	0.4180358	0.000654561
PtpAffx.145341.1.A1_at	Potri.001G164700	---	0.4187307	0.000638821
PtpAffx.16590.1.A1_at	Potri.005G081000	---	0.4195494	1.75E-06
PtpAffx.3126.1.A1_at	Potri.007G089400	Hypothetical protein	0.4231772	9.68E-06
PtpAffx.25920.1.S1_a_at	Potri.005G100700	hypothetical protein	0.4242405	0.000299478
PtpAffx.5224.1.A1_at	Potri.006G263600	hypothetical protein	0.4269911	0.000173972
PtpAffx.99162.1.S1_at	Potri.001G468500	---	0.4280747	1.88E-05
Ptp.8062.1.S1_s_at	Potri.001G325800	glutaredoxin	0.4292798	0.003683732
PtpAffx.4462.1.S1_a_at	Potri.T167100	hypothetical protein	0.4296979	0.002152229
PtpAffx.40216.2.S1_at	---	---	0.4302493	0.000909939
PtpAffx.107622.1.A1_at	Potri.006G109100	hypothetical protein	0.4306704	0.001871152
PtpAffx.15679.1.S1_at	Potri.008G041400	---	0.4331198	0.004816731
PtpAffx.141636.1.S1_at	Potri.009G019200	NAC domain trancription factor (NAC113)	0.4336458	0.001565153
PtpAffx.37401.1.A1_s_at	Potri.001G077900	hypothetical protein	0.4339553	0.000188718
PtpAffx.217134.1.S1_s_at	Potri.017G141800	hypothetical protein protein	0.434343	0.000764614
PtpAffx.87145.1.A1_at	Potri.005G109200	aquaporin, MIP family, PIP subfamily	0.4347753	0.003851125
Ptp.7612.2.S1_at	Potri.012G012300	---	0.4363235	0.02391087
PtpAffx.1770.3.S1_s_at	Potri.T131500	hypothetical protein	0.4364153	0.04366146
PtpAffx.2286.3.S1_a_at	Potri.011G140700	---	0.4368428	4.87E-05
PtpAffx.210610.1.S1_at	Potri.012G090500	hypothetical protein	0.4388454	0.000354357
PtpAffx.225537.1.S1_at	Potri.005G056500	hypothetical protein	0.4414143	7.32E-05
Ptp.4443.1.S1_s_at	Potri.019G119400	hypothetical protein	0.4420668	5.39E-05

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.20905.1.S1_at	Potri.018G075000	aldehyde dehydrogenase-like protein (Panther:11699)	0.442972	0.000312754
PtpAffx.120439.1.A1_at	Potri.003G129400	hypothetical protein	0.4448315	0.000479854
PtpAffx.216685.1.S1_s_at	Potri.T129600	nbs-lrr resistance protein	0.4462517	2.22E-05
PtpAffx.43564.1.A1_at	Potri.015G123400	hypotheticalprotein	0.4470961	0.001165429
PtpAffx.5688.2.S1_at	Potri.008G223000	Hypothetical protein	0.4471263	1.70E-06
PtpAffx.1225.1.S1_at	Potri.010G041300	---	0.4479272	0.03344028
PtpAffx.59464.1.A1_at	Potri.019G110300	hypothetical protein	0.4491697	0.003803014
Ptp.3289.1.A1_at	Potri.004G025900	Hypothetical protein	0.4495509	0.000264025
PtpAffx.214417.1.S1_at	Potri.018G097700	hypothetical protein	0.4510844	0.009776894
PtpAffx.36845.1.S1_s_at	Potri.003G112600	hypothetical protein	0.4516383	0.000120895
PtpAffx.151306.1.S1_at	Potri.011G126400	Hypothetical protein	0.4517572	0.000445894
Ptp.9.1.S1_at	Potri.003G070600	hypothetical protein	0.4522786	0.02370581
Ptp.4110.2.S1_s_at	Potri.002G216400	acetyl-CoA C-acyltransferase	0.4531293	0.001771454
PtpAffx.5175.1.A1_s_at	Potri.013G153400	hypothetical protein	0.4532594	3.15E-05
PtpAffx.249.26.A1_s_at	Potri.005G195600	hypothetical protein	0.4537976	0.001606666
PtpAffx.27878.2.S1_a_at	Potri.017G017300	glutaredoxin	0.4545895	0.000455221
PtpAffx.107379.1.S1_s_at	Potri.013G153400	hypothetical protein	0.4554816	7.45E-06
Ptp.1478.1.S1_x_at	---	---	0.4556326	0.000197586
PtpAffx.214630.1.S1_s_at	Potri.004G233200	hypothetical protein	0.4560248	0.001550634
Ptp.1176.2.A1_at	Potri.016G080600	Hypothetical protein	0.4560491	0.001383922
PtpAffx.224972.1.S1_at	Potri.019G120000	hypothetical protein	0.4569212	3.39E-05
PtpAffx.2935.2.A1_a_at	Potri.004G021300	MYB166	0.45789	0.0296401
PtpAffx.225074.1.S1_at	Potri.018G088100	hypothetical protein	0.4582238	6.05E-05
PtpAffx.17830.1.A1_at	Potri.015G068200	hypothetical protein	0.4590012	0.000423058
Ptp.2064.1.S1_at	Potri.013G028300	histidine phosphotransfer protein	0.4629568	0.000107073

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.35127.1.A1_at	Potri.T071900	hypothetical protein	0.4656731	0.001804775
PtpAffx.36845.1.S1_at	Potri.003G112600	---	0.4666163	0.008280213
PtpAffx.68244.1.S1_at	Potri.002G119300	hypothetical protein	0.4667521	0.009831845
PtpAffx.32933.1.A1_at	Potri.T128700	hypothetical protein	0.4672783	0.000367233
PtpAffx.28535.1.S1_at	Potri.019G086200	hypothetical protein	0.4673024	0.00021187
Ptp.2730.1.S1_s_at	Potri.019G055200	hypothetical protein	0.4697964	0.000833273
PtpAffx.10247.1.A1_at	Potri.006G112500	---	0.4702148	0.001506998
PtpAffx.454.22.A1_at	Potri.002G234300	---	0.4705432	7.18E-05
PtpAffx.5078.6.S1_a_at	Potri.006G019600	hypothetical protein	0.4706861	0.003439891
Ptp.7679.2.S1_s_at	Potri.007G072300	hypothetical protein	0.4707613	0.002967462
Ptp.3507.2.S1_at	---	---	0.4715812	0.000265843
PtpAffx.215124.1.S1_at	Potri.003G088300	hypothetical protein	0.4723838	6.78E-05
Ptp.4226.1.S1_at	Potri.006G044600	F-box family protein	0.4725538	0.02459631
Ptp.5882.1.S1_at	Potri.002G250000	hypothetical protein	0.4726674	0.01366876
PtpAffx.152367.1.S1_s_at	Potri.006G263600	hypothetical protein	0.4730573	0.000166648
Ptp.4747.1.A1_s_at	Potri.004G153800	hypothetical protein	0.4738595	0.00046779
PtpAffx.225084.1.S1_s_at	Potri.018G117900	hypothetical protein	0.4740804	0.000351209
PtpAffx.10032.1.S1_a_at	Potri.005G015000	vacuolar H+-translocating inorganic pyrophosphatase	0.4741602	1.99E-05
PtpAffx.200308.1.S1_at	Potri.001G077900	hypothetical protein	0.4745117	5.56E-06
PtpAffx.225628.1.S1_s_at	Potri.017G141300	hypothetical protein	0.4747619	0.000195314
PtpAffx.54929.1.A1_at	Potri.005G053200	---	0.4759356	0.01300263

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.36779.1.A1_a_at	Potri.003G177600	hypothetical protein	0.4760002	0.029473
PtpAffx.27878.3.A1_a_at	Potri.017G017300	glutaredoxin	0.4760655	0.003332186
PtpAffx.114369.1.S1_at	Potri.010G094400	---	0.4763071	6.68E-05
PtpAffx.29379.1.A1_at	Potri.003G087400	hypothetical protein	0.4767615	0.000520478
Ptp.7175.2.S1_s_at	Potri.001G164800	hypothetical protein	0.4772867	0.005715396
PtpAffx.210610.1.S1_x_at	Potri.012G090500	hypothetical protein	0.4785605	0.000149726
PtpAffx.9204.1.S1_at	Potri.015G030700	hypothetical protein	0.4786495	0.000268305
PtpAffx.23842.1.S1_a_at	Potri.009G005700	hypothetical protein	0.4794379	0.03891936
PtpAffx.57539.1.A1_at	Potri.003G130100	---	0.4796353	4.77E-05
PtpAffx.163890.1.S1_at	Potri.003G151000	AP2 ERF domain-containing transcription factor (ERF2)	0.4798768	0.02583661
PtpAffx.225490.1.S1_s_at	Potri.001G063300	leucine rich repeat family protein with ABC domain	0.4807007	0.002295841
PtpAffx.249.100.A1_s_at	Potri.005G195600	hypothetical protein	0.4814786	0.006106844
PtpAffx.10911.1.S1_at	Potri.011G122100	hypothetical protein	0.4821086	0.02075135
Ptp.3816.1.S1_s_at	Potri.012G126900	Hypothetical protein	0.4837651	0.02707626
PtpAffx.11613.1.A1_a_at	Potri.008G218800	hypothetical protein	0.4839672	0.002150138
PtpAffx.36994.1.S1_s_at	Potri.001G077900	hypothetical protein	0.4842699	2.66E-05
PtpAffx.30130.1.S1_at	Potri.007G005300	GCN5-related N-acetyltransferase (GNAT) family protein / amino acid kinase family protein	0.4845617	0.01404792
PtpAffx.102143.1.A1_s_at	Potri.001G079800	AP2/ERF domain-containing transcription factor	0.4850044	0.01708024
PtpAffx.225083.1.S1_at	Potri.018G116600	hypothetical protein	0.4854374	0.02090737
PtpAffx.85479.1.S1_s_at	Potri.002G108600	Hypothetical protein	0.4877928	0.000233748
Ptp.4024.2.A1_at	Potri.008G159000	hypothetical protein	0.4880241	0.03048834
PtpAffx.59667.1.A1_at	Potri.010G231200	hypothetical protein	0.4881128	0.01120995
PtpAffx.110326.1.S1_at	Potri.005G062500	---	0.4884575	0.000891368
PtpAffx.221860.1.S1_s_at	Potri.008G205200	cytochrome P450	0.4888778	0.000609888
PtpAffx.133289.1.S1_at	Potri.015G018100	hypothetical protein	0.4893804	0.01116325

Table A-2 continued.

Probeset	Gene Model ID	Gene Title	Fold change	P-value
PtpAffx.79915.1.S1_s_at	Potri.006G175500	hypothetical protein	0.4913825	0.002994603
PtpAffx.18527.1.A1_at	Potri.007G135900	hypothetical protein	0.4914923	0.000115328
Ptp.1323.1.S1_at	Potri.010G176000	isocitrate dehydrogenase (NADP+)	0.4920794	0.000101148
PtpAffx.221475.1.S1_s_at	Potri.002G075900	hypothetical protein	0.494839	0.000161914
Ptp.5925.2.S1_a_at	Potri.002G234300	hypothetical protein	0.4949599	6.93E-06
PtpAffx.61115.1.S1_at	Potri.001G312800	---	0.4955034	0.00103788
PtpAffx.34463.1.S1_at	---	---	0.4967537	0.003511228
PtpAffx.46526.1.S1_at	Potri.004G181900	NAC domain trancription factor (NAC074)	0.4967758	0.003228318
PtpAffx.203973.1.S1_at	Potri.004G088700	hypothetical protein	0.4973264	0.000761425
PtpAffx.38981.2.S1_a_at	Potri.001G076600	---	0.4979646	0.005337432
PtpAffx.9204.2.A1_at	Potri.012G043200	---	0.4985766	0.001687679
PtpAffx.203598.1.S1_x_at	Potri.004G192500	hypothetical protein	0.4993183	0.000269793

Table A-3. Linear regression analysis between tannin concentration and concentrations of phenolic metabolites as well as relative gene expression level in WT and MYB115 overexpressing poplar.

Data shown include all positive transformants for both the 717 (*P. tremula* x *P. alba* [INRA 717-1-B4]) and 353 (*P. tremula* x *P. tremuloides* [INRA 353-38]) backgrounds. Condensed tannins were quantified using the butanol HCl assay. Other phenolic metabolites were quantified by HPLC analysis. Relative gene expression was measured by qPCR and normalized against elongation factor 1-β. Metabolites that were below the detection limit of HPLC are labelled "n.d.". R^2 values of > 0.5 are bolded. Asterisks indicate the level of significance of the R^2 value ($p \leq 0.05$, *; $p \leq 0.01$, **, $p \leq 0.001$, ***).

	353-38 (R^2) tannin	717-1-B4 (R^2) tannin
catechin	**0.947*** **	**0.8842*** **
dihydromyricetin	**0.8604*** **	n.d.
dihydroquercetin	**0.8964*** **	**0.8685*** **
tremulacin	**0.7925*** **	**0.596*** **
salicortin	**0.7897*** **	**0.5904*** **
salicin	0.01959	0.3912*
grandidentatin-1	**0.7674*** **	n.d.
grandidentatin-2	**0.8485*** **	0.3999*
chlorogenic acid-1	0.496**	**0.5241** **
chlorogenic acid-2	0.1784	**0.6577*** **
coumaroyl quinic acid-1	0.183	0.4277**
coumaroyl quinic acid-2	0.3111*	**0.6832*** **
HCA-1	n.d.	0.1363
HCA-2	n.d.	**0.5799*** **
HCA-3	**0.8137*** **	n.d.
MYB115	**0.5369*** **	0.4524**
MYB134	0.3267*	0.2021
MYB182	0.1497	0.03032
F3'5'H1	0.4957**	0.1431
DFR2	**0.7078*** **	**0.6373*** **
ANR1	0.3624*	0.000837

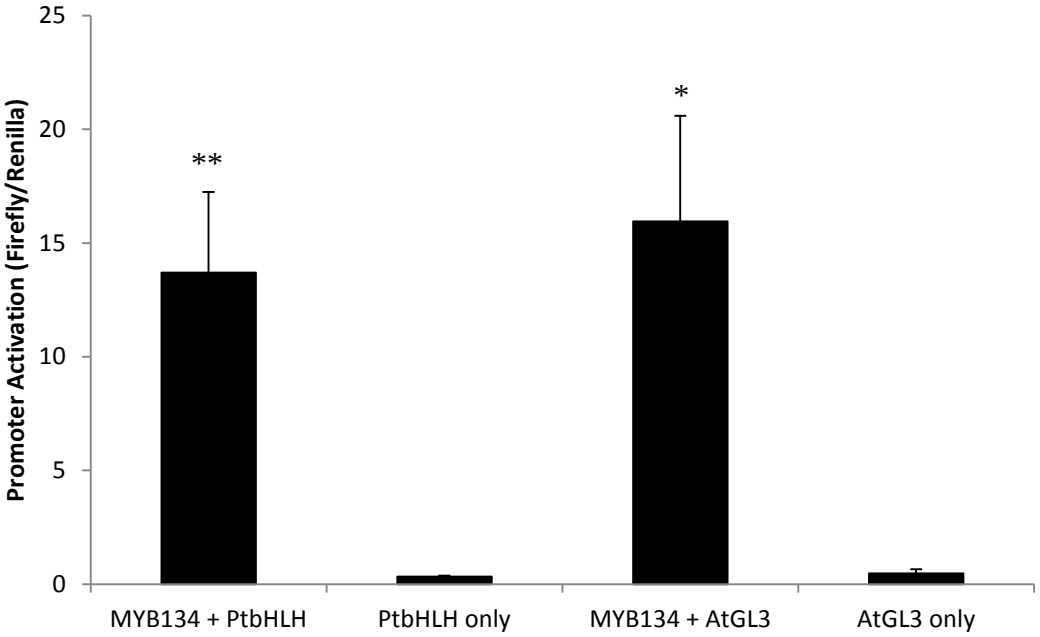

Figure A-3. Comparison of two bHLH factors as co-factors for MYB134.
Activation of the ANR promoter by MYB134 with PtbHLH131 (PtbHLH) Bombardment and luciferase assay performed by Dr. Kazuko Yoshida. Activators and co-factors are represented on the x-axis. Error bars indicate standard error (n = 4). Asterisks indicate results of a one-way ANOVA with date as a residual factor ($p \leq 0.05$, *; $p \leq 0.01$, **).

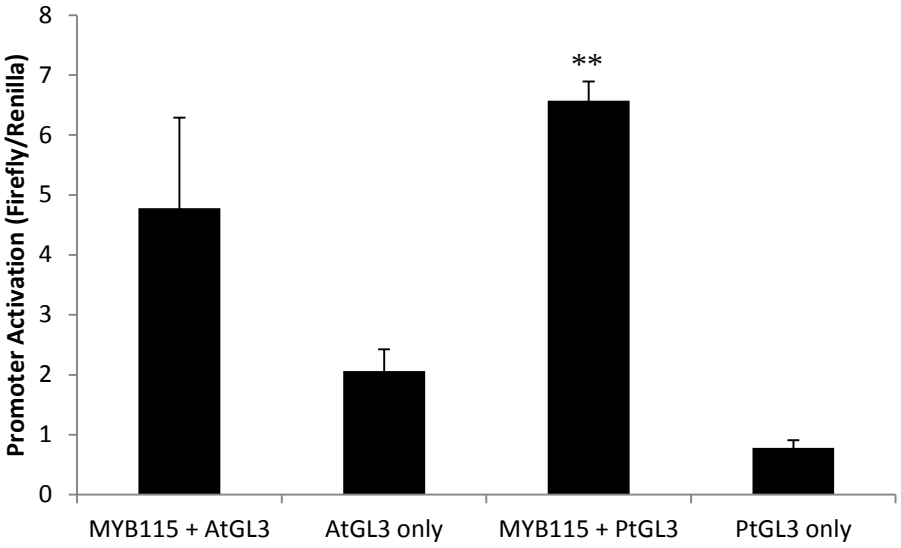

Figure A-4. Comparison of two bHLH factors as co-factors for MYB115.
Activation of the MYB115 promoter by the dual-luciferase promoter assay following bombardment of H11-11 poplar suspension cells. AtGL3 and PtGL3 (PtbHLH79) were co-bombarded with PtMYB115 or bombarded alone along with the reporter vector as a control and are represented on the x-axis. Error bars indicate standard error (n = 3). Asterisks indicate results of a one-way ANOVA with date as a residual factor ($p \leq 0.05$, *; $p \leq 0.01$, **).