# CRISPR/Cas9 Mutation of MYB134 and MYB115 To Study Regulation and Functions of Proanthocyanidins in Poplar Roots

by

Yalin Liu
Bachelor of Science (Honours), Beijing Forestry University, 2019

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

MASTER OF SCIENCE

in the Department of Biology

# Supervisory Committee

CRISPR/Cas9 Mutation of MYB134 and MYB115 To Study Regulation and Functions of Proanthocyanidins in Poplar Roots

by

Yalin Liu
Bachelor of Science (Honours), Beijing Forestry University, 2019

**Supervisory Committee**

Dr. C. Peter Constabel (Department of Biology)
**Supervisor**

Dr. Paul de la Bastide (Department of Biology)
**Departmental Member**

Dr. Barbara Hawkins (Department of Biology)
**Departmental Member**

# Abstract

Secondary metabolites play important roles in tree defense. Proanthocyanidins (PAs), one of the most common secondary metabolites, are widely distributed in trees and woody plants, and are abundant in poplar. In my research, molecular biology and biochemistry techniques were used to investigate the function of two important transcription factors, MYB115 and MYB134, in regulating the PA pathway in hybrid poplars. The importance of these transcription factors in regulating PA synthesis in leaves has recently emerged, but their roles in roots are not known. MYB134- and MYB115-overexpressing transgenic poplars showed a strong high-PA phenotype in leaves, but how these two regulators interact in vivo is still a mystery. This research aims to test the function of both MYBs in the regulation of PAs in poplar roots, and to explore the antimicrobial functions of root PAs. Both alleles of the MYB genes were sequenced in wild type poplars to design gRNAs for creating transgenic poplars with knocked-out (KO) MYB115 and MYB134 using the CRISPR Cas9 system. Both hairy root and whole plant transgenics with respective single- and double knock-outs were generated. Chemical and genetic characterization of both mutant types showed reduced PA content and down-regulated flavonoid genes in leaves. In poplar roots, only double-KOs showed a significant change in PA and salicinoid metabolism. These results indicated that the regulatory pathways for PA biosynthesis may differ in poplar leaves and roots. Significant PA concentrations remained in double-KO plants, suggesting other transcription factors for PA regulation are active. Because poplars accumulate large amounts of PAs in roots, potential functions of root tannins were also investigated. Antimicrobial activity of PAs was tested by disc inhibition assay *in vitro* and mycorrhizal co-culture sandwich assay *in vivo*. Pure PAs showed no inhibition towards the pathogenic fungi *Armillaria ostoyae* and *A. sinapina* but displayed slight inhibition to the mycorrhiza fungus *Laccaria bicolor*. These results provide preliminary insight into the functions of PAs in roots.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| AM | Arbuscular mycorrhizal fungi |
| ANR | Anthocyanidin reductase |
| ANS | Anthocyanidin synthase |
| BAN | BANYULS protein |
| bHLH | Basic helix–loop-helix proteins |
| CHI | Chalcone isomerase |
| CHS | Chalcone synthase |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| CTs | Condensed tannins |
| DFR | Dihydroflavonol reductase |
| DW | Dry weight |
| eGFP | Enhanced green fluorescent protein |
| ECM | Ectomycorrhizal fungi |
| F3H | Flavanone 3-hydroxylase |
| F3’H | Flavonoid 3’-hydroxylase |
| F3’5’H | Flavonoid 3’, 5’-hydroxylase |
| HTs | Hydrolysable tannins |
| LAR | Leucoanthocyanidin reductase |
| MATE | Multi-drug and toxic compound extrusion protein |
| MYB | Myeloblastosis transcription factors |
| PAs | Proanthocyanidins |
| ROS | Reactive oxygen species |
| RT-qPCR | Real-time reverse transcription polymerase chain reaction |
| UPLC-MS | Ultra High-performance liquid chromatography-Mass Spectrometry |

# Acknowledgments

This work was funded by NSERC as well as graduate scholarships from the Biology Department and Centre for Forestry Biology at the University of Victoria.

I would like to first thank my supervisor Dr. C. Peter Constabel for his support, guidance, and encouragement in all aspects. I am grateful for all your feedback throughout my project, and I have learned and gained so much over the last three years, such as plant biochemical knowledge, experimental skills, and TA experience. Thank you for being very patient as English is my second language, and with your patience and help on speaking and writing I improved a lot over the years. I would like to thank my committee members, Dr. Barbara Hawkins and Dr. Paul de la Bastide, for their suggestions, direction and encouragement during each meeting throughout this MSc project. I express my deepest gratitude to David Ma for his patience in teaching and training me on the molecular biology skills and for giving me insightful suggestions and helping on trouble shootings throughout my research. I am grateful to Eerik Piirtola for his patience and assistance in training me on how to use the UPLC system. I am grateful to Brad Binges for his teaching on how to transplant poplars and his assistance throughout my project with the greenhouse, my poplars, and the growth chamber. A special thanks to undergraduate student Simon Petley for assisting me with my poplars in the greenhouse and participating in mycorrhizal project. A warm thanks to Samantha Robbins for her patience in answering my questions on mycorrhiza. To Dr. Judith Felten lab from the Umea University, I am grateful to have had the chance to discuss my work with your lab and thank you for your suggestions for my project. Further thanks go to all the Constabel lab members and other graduate students in Forestry Biology. Thank you for your encouragement and support throughout these three years. I also acknowledge all the students, faculty, and staff in Biology department for their generous support and kindness during my master. Lastly, I express my thanks to my boy friend, Xiuwen, who accompanied me and helped me both academically and personally. Also, many thanks to my parents for their support on my abroad study and my friends here for making my time at UVic unforgettable.

# Chapter 1: Literature Review

## Introduction: Objective and Scope of the Review

Proanthocyanidins (PAs), also known as condensed tannins (CTs), are one of the most common secondary metabolites and are widely distributed in trees and woody plants. The biological functions of PAs in leaves and shoots include anti-herbivore and anti-pathogen effects as well as antioxidant functions. PAs have been observed in the roots of many tree species. However, little research attention has been given to the potential physiological and ecological roles of PAs in roots. These could include chelation of ions and nutrients or defence against soil pests and pathogens. Since PAs have been shown to play a role in defense against fungal and bacterial pathogens in leaves and stems, PAs may also function as anti-fungal compounds in roots. This review will outline the properties, distribution, regulation, and potential functions of PAs, with emphasis on their localization and potential functions in roots. My major focus is on the flavonoid-derived PAs.

## 1. Background on Tannins

### 1.1 Introduction to PAs and other tannins

#### *Basic structures of tannins*

Tannins are functionally defined as polyphenols that have the ability to precipitate proteins in solution. Tannins are derived from different biosynthetic pathways and are thus classified into two major groups: hydrolysable tannins (HTs) and condensed tannins (CTs, also known as proanthocyanidins, PAs). Hydrolysable tannins contain a central glucose, which is esterified with gallic acid and gallic acid derived moieties (Jourdes et al., 2013). PAs are oligomers or polymers built up by polymerization of flavan-3-ols, for example catechin, epicatechin, epigallocatechin, and fisetinidol (Hagerman and Butler, 1991). PAs from different species vary in the degree of polymerization, which can range from two to over 20, and can vary significantly even within the same genus (Scioneaux et al., 2011).

PA polymers from different species may differ due to the composition of the constituent monomers which vary in stereochemistry and hydroxylation pattern (Scioneaux et al., 2011). Based on the A ring type (Figure 1-1), PAs are divided into phloroglucinol and resorcinol categories (Aron and Kennedy, 2008). Procyanidins and prodelphinidins represent the most

common phloroglucinol type PAs found in nature, with different hydroxylation patterns on the B ring. Differences in inter-flavan linkages also contribute to the diversity of PA structures (Porter, 1992). Linkage between successive monomeric units in proanthocyanidins is usually between the C4 of the "upper'' extension unit and the C8 or C6 of the ''lower'' or ''starter'' unit. The B-type proanthocyanidins are oligomers or polymers of flavan-3-ols linked via C4β –C8/C6 bonds (Xie and Dixon, 2005). For example, procyanidins and the prodelphinidins are linear chains of flavan-3-ols linked via C4–C8 bonds. In A-type proanthocyanidins, a second ether linkage between an A-ring hydroxyl group of the lower unit and C-2 of the upper unit is present, namely one C–C and one C-O bond (C4–C8 and 2β → O → 7) (Xie and Dixon, 2005). The proportion of *cis* vs. *trans* monomers, namely the ratio of epicatechin to catechin can also determine the structures of PA polymers (Barbehenn and Constabel, 2011). Additionally, other phenolics may substitute groups of PAs, for example, gallic acid may be conjugated to oligomeric PAs to form 3-O-galloylated tannin (Li et al., 2010). Therefore, the diversity of PA structures depends on composition and proportion of the constituent monomers, subunit linkages, substitution by other phenolic groups, and degree of polymerization.

Hydrolysable tannins are synthesized from the shikimate-derived gallic acid (3,4,5-trihydroxybenzoic acid) and comprise two subclasses: galloyl glucoses and ellagitannins. Galloyl glucoses are derived from the esterification of sugar-type polyols with multiple galloyl units. Gallotannins are high-molecular-weight galloyl glucoses and figure importantly in protein precipitation, enzyme inhibition and metal chelation through forming various complexes in vitro (e.g., pentagalloyl glucose) (Okuda and Ito, 2011). Ellagitannins are synthesized by plants from galloyl glucoses by the oxidative coupling of adjacent galloyl groups, for example hexahydroxydiphenoyl (HDDP) unit, which is the structural characteristic that defines hydrolysable tannins as ellagitannins (Moilanen and Salminen, 2008). Examples of ellagitannins include peduncalagin and vescalagin (Okuda et al., 2009).

***Chemical properties of tannins***

Tannins exhibit a diverse range of *in vitro* biochemical properties. Tannins bind proteins at neutral or acidic to mildly basic pH principally by forming hydrogen bonds at their phenolic hydroxyls (Barbehenn and Constabel, 2011). Hydrophobic interactions between aromatic rings of tannins and proteins can also contribute to the formation of complexes, but are considered far

weaker than hydrogen bonding (Barbehenn and Constabel, 2011). Protein-binding efficacy is affected by multiple factors, including tannin structure, molecular weight and concentration, as well as protein amino acid composition or structure.

Tannins can also act as strong *in vitro* antioxidants due to their ability to scavenge free radicals or to reduce other oxidized compounds and form relatively stable semiquinone radicals. These effects are based on the capacity of the phenol functional group, such as PAs and flavonoids, to donate a hydrogen atom or a single electron to a free radical (Seyoum et al., 2006; Quideau et al., 2011). There are two physicochemical parameters that determine the potential efficacy of antioxidant action: the dissociation energy and the ionization potential of the phenol (Quideau et al., 2011), which can be affected by the number and position of additional hydroxyls and aromatic rings (Burda and Oleszek, 2001; Furuno et al., 2002). For example, the typical catecholic B-ring (phenolic units bearing adjacent hydroxy groups) of most PAs is the key factor determining their antioxidant capacity (Quideau et al., 2011). Furthermore, the number of galloyl subunits in PAs as well as the degree of polymerization and molecular weight may also affect their antioxidant property (Hagerman et al., 1998). This is because catecholic and pyrogallolic moieties can act as privileged radical-scavenging sites.

Another property of tannins *in vitro* is their ability to chelate metal ions, such as $Al^{3+}$, $Fe^{2+}$, $Zn^{2+}$ and $Cu^{2+}$. Tannins, like most plant polyphenols, contain two or more adjacent hydroxyl groups on aromatic rings, which enable tannins to chelate metal ions (Quideau et al., 2011). Zeng et al. (2019) found PAs in apple fruits interacted with metal ions mainly via catechol group-ion exchange, including $Fe^{2+}$, $Cu^{2+}$ and $Zn^{2+}$. Another study also revealed that tannin fractions obtained from three edible nuts were able to chelate $Fe^{2+}$, $Cu^{2+}$ and $Zn^{2+}$ (Karamać, 2009). A study by Chin et al. (2009) showed that immobilized tannic acid prepared from lateral roots of *Symphytum officinale* L., can chelate lead ($Pb^{2+}$). The hydrolysable tannins also exhibit the ability to chelate Fe ions and the additional galloyl group can enhance this ability (Moilanen et al., 2016). Complexation of heavy metals may play a role in many biological and ecological processes, for example, improve cationic nutrient cycling through plant-litter-soil interactions (Quideau et al., 2011). Additionally, binding of PAs to aluminum can reduce $Al^{3+}$ toxicity in soils (Osawa et al., 2011). The flavan-3-ol catechin as well as flavones have been shown to act as antioxidants by chelating iron and copper ions that are involved in ROS formation (Mira et al., 2002). The $Fe^{2+}$ ions binding ability of tannins may be important for their antimicrobial

properties as well (Scalbert, 1991).

## 1.2 PA biosynthesis

***Biosynthesis of PAs***

PAs (and anthocyanins) are major end products of the flavonoid pathway, downstream of the phenylpropanoid pathway (Dixon et al., 2005). The flavonoid biosynthetic pathway has been well characterized through biochemical and genetic analyses and the major enzymes and the corresponding genes in this pathway have been identified in many species, following early work in *Arabidopsis thaliana* (Xie et al., 2003). The general flavonoid biosynthesis pathways are displayed in Figure 1-1. The first committed step of the flavonoid pathway is the synthesis of chalcone by chalcone synthase (CHS), which is converted to flavanone by chalcone isomerase (CHI). Flavanones are then converted into dihydroflavonols, with different numbers of hydroxyl groups on the B ring catalyzed by three different flavanoid-hydroxylases. The dihydroflavonols are substrates for the NADPH-dependent reductase, dihydroflavonol reductase (DFR), which produces the leucoanthocyanidins (flavan-3,4-diols). Leucoanthocyanidins are converted into the colored, achiral anthocyanidins, pelargonidin, cyanidin and delphinidin through the action of anthocyanidin synthase (ANS) (Tanner et al., 2003). Anthocyanidin reductase (ANR) provides a biosynthetic route from anthocyanidin to the major mono-unit of PAs, 2,3-*cis*-epicatechin. Another key PA-specific enzyme is leucoanthocyanidin reductase (LAR), which can directly convert 2,3-trans-leucocyanidin to 2,3-*trans*-catechin. ANR and LAR together can provide *cis*- and *trans*-flavan-3-ol precursors needed for PA biosynthesis, and their relative activities can determine the types of PA in different species.

The mechanism of PAs oligomerization and polymerization from flavan-3-ol initiation units remains poorly understood, and both enzymatic and non-enzymatic mechanisms have been proposed. Early work by Haslam (1977) proposed the hypothesis that acid cleavage of a procyanidin dimer could generate an active intermediate (flav-3-en-3-ol) that could then couple with a starter unit, epicatechin, to form higher oligomers. Models for PA transport and polymerization, based on the phenotypes of *tt* mutants, were described in Zhao *et al*. (2010) and Dixon and Sarnala (2020). In species that lack the LAR gene, such as Arabidopsis, ANR produces epicatechin-only PAs, meanwhile ANR can transfer potential substrate flav-2-en-3,4-

diols into 2,3-cis-leucocyanidin, which may act as an epicatechin extension unit. In species such as *Medicago* that possess the LAR genes, LAR has the ability to determine PA chain length by converting 4 → 8 linked extension unit, Cys-EC [4β-(S-cysteinyl)-epicatechin], back to the starter unit epicatechin. This function of LAR was confirmed by the mutation of this gene in *M. truncatula*, which leads to accumulation of Cys-EC and increased levels of insoluble PAs (Liu *et al*., 2016). Another study by Yu *et al.* (2019) also found this dual-function of LAR in grapevine, where LAR can cleave the potential tannin extension unit precursors Cys-C [4β-(S-cysteinyl)-catechin] to generate catechin. This suggests that LAR can generate both catechin and epicatechin and may be involved in determining the degree of PA polymerization by regulating the relative proportions of starter and extension units. However, how PAs are polymerized is not yet clear.

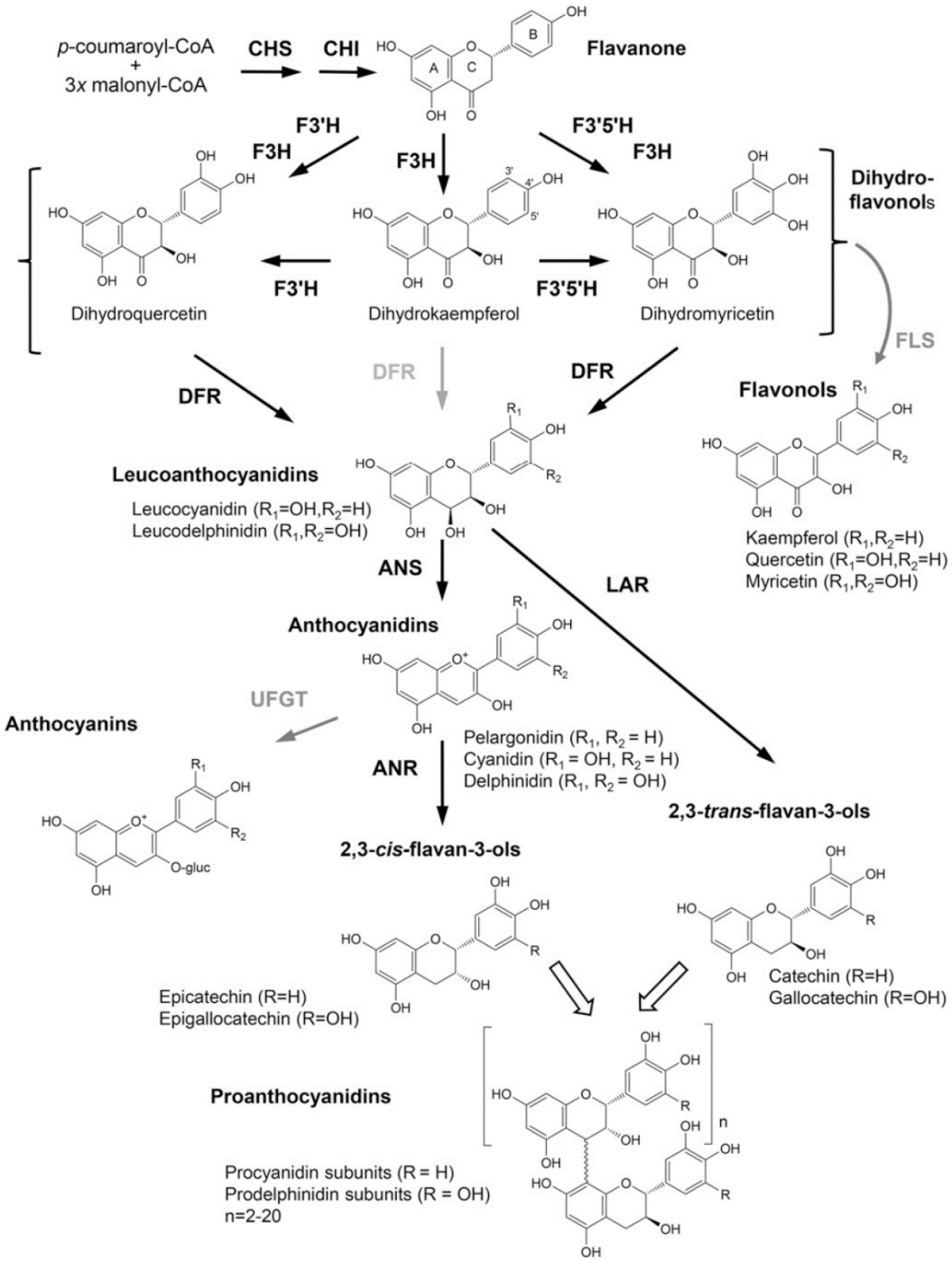



**Figure 1-1: General flavonoid pathway leading to the biosynthesis of PAs (from James et al., 2017).** Black arrows indicate the demonstrated metabolic routes to PAs in poplar leaves.

## 1.3 Ecological and physiological functions of tannins

### 1.3.1 Tannins as defensive phytochemicals

***Anti-digestive properties of tannins in vertebrates against herbivores***

The protein digestion-reducing activity of tannins (reviewed in Constabel et al., 2014), arising from their ability to bind enzymes and dietary protein in herbivores' digestive tracts, has been extensively studied. Tannins are proposed to inhibit the breakdown of plant material and thus reduce the assimilation of nitrogen, which may cause nutrient limitation and negative consequences for herbivore growth and fitness. At high pH such as found in lepidopteran guts, however, the phenolic groups are oxidized and thus unable to precipitate proteins. By contrast, in acidic environments such as vertebrate herbivore digestive tracts, tannin-protein binding occurs (Stevens and Hume, 2004). Tannins in forages can bind dietary protein and inhibit digestive enzymes, and therefore decrease protein utilization as well as reduce reproductive success of mammals (Wallis et al., 2012). Many vertebrate herbivores are known to select food plants based on levels of tannins and co-occurring nutrients (Perkovich and Ward, 2020), and PAs have been proven to be an effective feeding deterrent for mammal species including sheep, cattle and pigs (Hervás et al., 2003; Mueller-Harvey, 2006). Interestingly, some browsing animal species have evolved adaption to a high PA diet by producing proline-rich salivary proteins. These proteins have a strong affinity for tannins, and thus reduce the amounts of free tannins in the digestive tract, acting as defense against dietary tannins (Mueller-Harvey, 2006). Foley and coworkers (2003) provided evidence for an antiherbivore role of tannins by studying eucalypt-feeding mammals. They found ringtail and brushtail possums had differential susceptibility to co-occurring secondary metabolites, where brushtails avoided tannins while ringtails showed highly tolerance of foliar tannins (Marsh et al., 2003). In another study, tannins reduced the digestibility of nitrogen *in vitro* and strongly influenced the reproductive success of brushtail possum (DeGabriel et al., 2009).

***Feeding deterrent and anti-digestive effects of tannins against insects***

Karowe (1989) found that tannic acid decreases the efficiency of conversion of digested food, reduces growth, and increases mortality in *Malacosoma disstria* (forest tent caterpillar), but does not reduce digestibility. By contrast, adding PAs or HTs to the diet does not reduce protein utilization by *Lymantria dispar* (gypsy moth caterpillar) (Barbehenn et al., 2009). These

observations suggest that tannins may act as both toxin and feeding deterrent. The protein-binding and astringent properties of PAs also impart less palatability to high PA foliage (Barbehenn and Constabel, 2011). In cowpea seed, high tannin content correlates with resistance to bruchid beetles during storage, suggesting PAs role as biochemical defence (Lattanzio et al., 2005). In a different system, choice tests indicated that PAs in poplar may not act as an effective feeding deterrent to tree-feeding generalists. In feeding trials, two lepidopteran species, *Lymantria dispar* (gypsy moth caterpillar) and *Malacosoma disstria* (forest tent caterpillar) consistently showed a strong preference for high-PA transgenic poplars leaf disks and larvae growth on the transgenics was enhanced compared to that of wild-type control (Boeckler et al., 2014). However, these feeding choices were not entirely due to the preference for PAs, but may also be associated with the shifts in the balance of other phenolics in high PA transgenic lines (Boeckler et al., 2014). A recent study also revealed that nutritional context, specifically the protein: carbohydrate ratio of the diet, may affect the tannin deterrent effect on insects, since there may be a trade-off between development time and efficiency of food assimilation (Perkovich and Ward, 2020).

***Toxic and prooxidant effects of tannins***

The effects of tannins on herbivores are complicated by their toxic and prooxidant properties. At high pH, the phenolic groups of tannins are oxidized and thus tannins are unlikely to precipitate dietary protein as a defense against leaf-eating caterpillars with extremely alkaline guts. Foliar tannin composition may affect the level of phenolic oxidation in the midguts of caterpillars (Barbehenn et al., 2008), and tannins can become harmful prooxidants at pH>9 in herbivore digestive tracts which can lead to toxic oxidative damage. This was observed in non-adapted grasshoppers and lepidopterans feeding on leaves coated in tannic acid, where fatal lesions were produced in their midguts (Barbehenn and Constabel, 2011; Bernays et al., 1980). Instead of a direct toxic effect of tannins on the midgut epithelium, midgut lesions may also be caused indirectly by oxidative stress via oxidation of tannins. Tannin prooxidant activity and toxicity can thus cause detrimental post-ingestion effects on insects, as tannin oxidation in the gut lumen can generate high levels of reactive oxygen species (ROS) and damage cellular components. It is important to note that the type and structure of the tannin affects prooxidant activity. Barbehenn *et al*. (2006) found that hydrolyzable tannins (i.e., ellagitannins and galloyl

glycosides) are much more prone to oxidize than are condensed tannins at the high pH of the caterpillar midgut lumen (e.g., pH 10). Additionally, higher oxidative stress and lower protein utilization efficiency was observed in the midgut fluids of two caterpillar species fed on sugar maple compared with on red oak, mainly because of the higher levels of hydrolysable tannins in maple (Barbehenn et al., 2005). By contrast, many studies failed to demonstrate an effect of ingested tannins on protein utilization and caterpillar growth in herbivorous insects. For example, increasing tannins in hybrid poplar could not significantly improve the resistance against *L. dispar* (Barbehenn et al., 2009).

### 1.3.2 Antimicrobial and anti-pathogen roles and mechanisms

Tannins can also exert antimicrobial effects due to their ability to bind proteins and other macromolecules, including extracellular enzymes, secreted substrates and heavy metals. PAs may also inhibit fungal colonization by crosslinking and inactivating secreted enzymes and cell wall proteins (Scalbert, 1991). Activities of enzymes, including cellulases, pectinases and peroxidases, are all known to be negatively affected by tannins (Scalbert, 1991). This may explain the reduced breakdown of cellulose, pectins, or hemicelluloses in the presence of tannins that has been observed in many studies (Constabel et al., 2014).

Tannins can inhibit a wide array of microbes and filamentous fungi *in vitro*, including human pathogens, plant pathogens, food-spoiling saprophytes, and rumen symbionts (Mila et al., 1996; Puupponen-Pimiä et al., 2005; Engels et al., 2009; Anderson et al., 2012; Yoshihara et al., 2013; Ullah et al., 2017). For example, PAs may inhibit gastrointestinal bacteria via binding to membrane or cell-wall components, which could underlie their benefit in human health (Smith et al., 2005). Inhibition of microbes is also important in livestock health, where tannins protect ruminants from antinutritional effects via multiple mechanisms (McSweeney et al., 2001). In some cases, higher-molecular-weight tannins exhibited greater inhibition of microbes, as larger molecular weight tannins are better protein precipitants than smaller tannins (Quideau et al., 2011). However, other experiments also showed smaller phenolics without protein-precipitation activity were as effective as larger polymers, suggesting that not only protein-binding but other strategies are involved in antimicrobial mechanisms (Scalbert, 1991). In addition, metal-binding activity of tannins may contribute to their interaction with microbes (Scalbert, 1991). For example, the bacterial pathogen *Erwinia chrysanthemi* was inhibited significantly by tannins *in*

*vitro* as a result of iron chelation, which is supported by the observation that strains deficient in iron transport were more sensitive to the tannin (Mila et al., 1996). In addition, a recent study by McGivern et al. (2021) revealed that some soil microbiota can degrade polyphenols including PAs under anoxia. In summary, antimicrobial effects of tannins could involve multiple potential mechanisms in different organisms.

The *in vitro* chemical properties of tannins described above may give rise to *in vivo* biological effects of tannins in plants, in particular anti-pathogen functions. As metal ions may be essential for fungal growth, chelating metals may provide root PAs with anti-pathogen properties. The hypothesis that PAs are potential antifungal compounds for plant defense is supported by several studies, which demonstrate a stimulation of PA synthesis by pathogen attack (Miranda et al., 2007). This was also observed in *Populus tomentosa Carr.* leaves following infection by the fungal pathogen *Marssonina brunnea* (Yuan et al., 2012). The overexpression of a PA biosynthetic gene led to improved resistance against this pathogen *in vitro*. Similar results were also observed in black poplar (*Populus nigra*), where catechin and PAs accumulated in leaves after rust fungus *Melampsora. larici-populina* infection (Ullah et al., 2017). As expected, poplar genotypes with higher levels of PAs showed reduced rust susceptibility, while poplar with down-regulated PA biosynthesis had increased rust susceptibility. Ullah et al. (2019) also found that another fungal pathogen (*Plectosphaerella populi*) induced flavan-3-ol accumulation in *P. nigra* stems after infection. A recent study reported that the overexpression of two transcription factors enhanced the pathogen resistance against *Botrytis cinerea* and *Dothiorella gregaria* infection in *Populus alba* leaves by promoting accumulation of flavonoids, including not only anthocyanins and PAs, but also the flavonols quercetin and kaempferol (Bai et al., 2020). Plant endophytic fungi appear to be affected by PAs as well, as a negative relationship between PAs concentration and success of fungal endophyte infection was recorded in twig bark of Fremont cottonwood (Bailey et al., 2005). These studies all indicate that PAs have the potential to influence plant-associated fungi.

### 1.3.3 In planta protective effects of PAs against oxidative stress

The antioxidant capacity of PAs *in vitro* is well established, but only recently has their *in vivo* role against oxidative stress in woody plant leaves been demonstrated (Gourlay and Constabel, 2019; Gourlay et al., 2020). Reactive oxygen species (ROS) accumulation is a common feature under exposure to stressful conditions such as high light, UV-B irradiation,

drought stress and pathogen infection. As antioxidants, PAs may therefore contribute to plant tolerance against a range of abiotic stresses. PA accumulation could be an adaptive response to different oxidative stress conditions as a result of enhanced ROS in plant tissues, especially in tree leaves with abundant PAs (Gourlay and Constabel, 2019). Some studies have found an association between oxidative stress and increased foliar PA concentration in trees as indirect evidence for PAs antioxidant effects in plants. For example, PA accumulation after elevated UV-B irradiation exposure was reported in birch (B. *pendula*) and hybrid poplar (Lavola, 1998; Mellway et al., 2009). Ward and Young (2002) found that the canopy of acacia trees with exposure to solar irradiation, but not herbivory, produced the highest levels of tannin (Close and McArthur, 2002). Importantly, Gourlay and Constabel (2019) demonstrated that transgenic poplar engineered for very high PA content in leaves were more resistant to the oxidative damage caused by methyl viologen (MV), an oxidative stress-inducing herbicide as well as UV-B radiation and drought stress, by reducing ROS accumulation in poplar leaves (Gourlay et al., 2022). In contrast, poplar plants with reduced tannin synthesis in leaves increased susceptibility to these oxidative stress (Gourlay et al., 2020; Gourlay et al., 2022). These studies provide direct evidence for PA's antioxidant function in poplar leaves.

## 2. Distribution and localization of PAs

### 2.1 Overview of PA distribution in plant tissues

A number of histochemical methods have been developed to analyze the tissue-specific accumulation patterns of PAs in fresh plant material. These include 4-dimethylaminocinnamaldehyde (DMACA), which is a PA-specific stain but does not stain hydrolyzable tannins (Feucht et al., 1986; Constabel et al., 2014), and vanillin-HCl, which detects not only PAs but also their immediate precursor molecules monomeric flavan-3-ols and other phenolics (Li et al., 1996). There are also some other less commonly used staining tests such as the Hoepfner-Vorsatz (H-V) test, which detects phenolic groups; dimethoxybenzaldehyde (DMB), which stains PAs red; and ferrous sulfate in formalin (McKenzie and Peterson, 1995a). These stains have been used for detecting both cell- and tissue-specific PA accumulation.

PAs are typically found at high concentrations in vegetative woody plants, including both conifers and broad-leaved poplars, where they occur in most vegetative tissues, including bark,

leaves, and roots (Xie and Dixon, 2005; Hoffmann et al., 2012). PAs are also prevalent in reproductive organs, especially in seeds and fruits. In herbaceous species, PAs are more restricted in distribution, for example, in leaves of some legume family plants and seed coat (McMahon et al., 2000; Lepiniec et al., 2006). Thus, the root accumulation of tannins appears to be primarily associated with the woody habit.

Tannins are commonly found in leaves of forest trees, such as *Populus tremuloides* (Osier and Lindroth, 2006), *Quercus petraea* (Peng et al., 1991), *Betula papyrifera* (Laitinen et al., 2002), *Liquidambar formosana* (Hatano et al., 1986), *Acer saccharum* (Inoue and Hagerman, 1988). PA localization was studied in newly emerged and fully expanded leaves by both vanillin-HCl and DMACA staining in *Populus tremuloides* (Kao et al., 2002). In newly emerged leaves, fewer PAs were found in palisade and the lower tier of spongy mesophyll cells, and more PAs were detected in the abaxial epidermis. A similar pattern was observed in fully expanded leaves, and a shift of PAs distribution from the abaxial epidermis to the subepidermal layer of spongy mesophyll. The shift between epidermal and hypodermal cells was also observed in developing stems, where PAs were found in epidermis, hypodermis, phloem idioblasts, and developing ray parenchyma (Kao et al., 2002). By contrast in herbaceous plant vegetative tissues, tannins are generally absent, although they may be found in leaves of some legume species. In leaves of forage legumes *Onobrychis viciifolia*, flavan-3-ols were detected in subepidermal cells of both abaxial and adaxial surfaces, but tannin accumulation was also observed in these cells (Lees et al., 1993). Lees et al. (1995a) also observed a shift of PAs distribution in sainfoin (*O. viciifolia*) seedlings from the abaxial to the adaxial side as the leaflets unfold and mature, and PAs continue to increase in specialized cells called idioblasts in the epidermis. The epidermal localization of PA accumulation is suggestive of a protective function against pathogens or a function in protection against UV stress and associated ROS activity (Close and McArthur, 2002).

Seeds and fruits are rich in PAs and their distribution have been investigated in these structures as well. For example, in *Arabidopsis* seed coats, PAs undergo oxidative polymerization and form brown pigments, before being deposited on to the cell walls, where the deposition was limited to the endothelium layer, which is the innermost layer of the seed coat (testa) (Debeaujon et al., 2001; Pourcel et al., 2005). These studies have demonstrated the importance of tannins in seed coat physiology. In blueberry, PAs are detected throughout the entire fruit early in development, including ovary tissues, the developing seed coat, and the fruit

flesh (mesocarp). As the fruit develops, the PA staining becomes minimal in the mesocarp but remained concentrated in the seed coats, placentae, and the exocarp. At fruit maturity, both PAs and anthocyanins were detected at high levels in the epidermis (Zifkin et al., 2012). Likewise, apple PAs deposit in the seed coat and in the hypodermis, but not the cortex (Lees et al., 1995b). In Tunisian Dates (*Phoenix dactylifera L.*), PAs were localized to both edible tissues called pericarp and hard tissues called stone (Hammouda et al., 2014). This study showed that tannins are not only located in the epidermis of dates but also more deeply in the mesocarp. Stone tannins are found in a specific cell layer near the sclereid cells of the seed coat of *Phoenix dactylifera L.* (Hammouda et al., 2014). In *Silybum marianum* (L.) Gaertn, accumulation of PAs was observed in the pericarp subepidermal cell layer of fully-ripened fruits (Giuliani et al., 2018). In mangosteen fruit, the PAs are also found in the pericarp (Fu et al., 2007; Zhou et al., 2011). A recent study by Brahem et al. (2020) characterized procyanidin localization at ripe and overripe stages in perry pear flesh, where tannins were found in parenchyma cells of the mesocarp. These studies indicated the diverse localization of PAs in the seed coat and mesocarp of fruits. This localization at the outer layer of seeds and fruits may suggest their defensive functions against pests,  pathogens or frugivores.

### 2.2 PA distribution and localization in roots

Compared to the large number of studies focused on tannins in leaves and other above-ground organs of plants, the literature on root tannins is very limited. However, it has not previously been summarized and thus this review emphasizes PA distribution in roots. The distribution of PAs in roots has been studied in only a handful of species, mostly in woody plants. Root tannin distribution varies with plant species, for example, root cap localization in *Populus* (Kao et al., 2002) and *Rosaceae* (Hoffmann et al., 2012), older brown root localization in *Pinus* and *Eucalyptus* (McKenzie and Peterson, 1995a), or complete absence of root PAs (despite being present in stem and leaf tissues) in herbaceous plant such as *Onobrychis viciifolia* (sainfoin) (Lees et al., 1993).

Based on previous anatomical research with jack pine (*Pinus banksiana*) and *Eucalyptus pilularis*, typical woody roots are defined by three anatomically different zones: the white zone, condensed tannin (PA) zone, and cork zone (McKenzie and Peterson, 1995a, b). The white zone

roots contain living cortex, non-suberized endodermis and a central stele containing xylem. The brown zone is characterized by a dying cortex, occurring >2 cm from root tip and prior to the cork zone. The cork zone is oldest part of the root and terminates at the root-stem interface. It contains secondary xylem and phloem and a ring of mature cork cells (Peterson et al., 1999). The white zone and brown zones lack secondary growth, whereas the cork zone undergoes secondary growth.

McKenzie and Peterson (1995a) found that PAs were primarily present in visibly brown zones, using four tests: vanillin-HCl, ferrous sulphate, DMB and H-V (Table 1-1). Staining intensity in the brown zone increased with distance from root tip. By contrast, the white zone only stained very faintly for PAs. The brown pigmentation in the brown zone was attributed to PAs that were visualized in the younger region of brown zone, and to the deposition and oxidation of PAs into cortical cells walls as these undergo cell death (McKenzie and Peterson, 1995a). Brown pigmentation in the brown zone was not correlated with lignin or suberin, which were also histochemically visualized (McKenzie and Peterson, 1995a). In the cork zone, however, the brown colour observed was attributed in part to PAs but also to lignin and suberin deposition in the form of Casparian bands (McKenzie and Peterson, 1995b). However, in other species the distribution of PAs in roots can be quite different.

In woody plants, PAs are detected in different zones or cell layers summarized in Table 1-1. In roots of Pinaceae *Pseudotsuga menziesii*, PAs were detected in thickened cell walls of inner cortical tissue as seen by DMACA staining (Weiss et al., 1999). In Rosaceae *Malus* x *domestica,* PA-specific histochemical staining detected PA deposits in the white zone; specifically in the root cap as well as cortical and epidermal cells of the distal root tip (Hoffmann et al., 2012). In Fagaceae trees, including *Quercus, Castanopsis and Lithocarpu*s, PAs were found mainly in the root cap, the epidermal layer and the endodermis (Tam and Griffiths, 1993). In the camphor tree (*Cinnamomum camphora*), PAs were observed by DMACA staining mostly in the outer epidermal cells of the root tip (Osawa et al., 2011; see Section 3.1). In grapevine (*Vitis vinifera*), uneven localization of PAs in hairy root tissues was observed. PAs were detected and quite abundant in the youngest part of the hairy roots, namely the cell division zone below the apex (Terrier et al., 2009). The epidermis, endoderm, and vascular bundles were found to be the richest area accumulating high level of PAs in the division zone of the root tip. Almost no DMACA staining could be observed in the cortical parenchyma cells of hairy roots (Terrier et al.,

2009). In the non-woody plant *Fragaria* x *ananassa*, PAs were detected in white zone as well as in cortical and epidermal cells of the root tip (Hoffmann et al., 2012).

In *Populus tremula* x *alba*, PA localization and PA concentrations and the relationship between PA distribution and root pigmentation was studied using histochemical staining (DMACA) and quantitative assays (Westley, 2015). PA was found to accumulate in the white zones, thus its concentrations was not correlated with brown pigmentation as in previous studies. Additionally, PAs accumulated predominantly in cells bordering the plant-soil interface and tissues of the young, active white zones. They were also sporadically localized in the cortex (Westley, 2015). Similar findings were observed in *P. tremuloides* (Kao et al., 2002), where PAs localization was determined by DMACA and vanillin-HCl staining, detecting a range of phenolics. Positive staining was found in the root cap as well as in epidermal cells of the root tip, and was sporadically present in cortical cells as well as the meristematic and early elongation zones in root tips (Kao et al., 2002).

The precursors of PAs, catechin, can also be detected in roots of different species. Mace and Howell (1974) reported the histochemical localization of catechin and gallocatechin in fresh root sections of Acala cotton (*Gossypium hirsutum*) seedlings, a herbaceous plant. PA precursors were heavily concentrated in the root endodermis and hypodermis by DMB staining. Although catechins were not detected in the main body of the root cap, they were present in a few scattered epidermal cells in the maturation zone, and also existed in a continuous ring of endodermal cells in the taproot endodermis. They were also found in randomly scattered parenchyma cells of the root stele as well as a discontinuous ring of one or two cell layers in hypodermal cells. Catechins, though not PAs, were present at different levels in different locations in *Camellia sinensis* (a woody plant), around the pericycle of the root, but no catechin was observed in the cortex or xylem vessels (Liu et al., 2009).

**Table 1-1: The localization of PAs or precursors in roots of different species and the corresponding staining methods.**

| species | localization | Detected compound | Staining methods | reference |
|---|---|---|---|---|
| Pinaceae (*Pinus*) & Myrtaceae (*Eucalyptus*) | brown zones (>2 cm from root tip and prior to the cork zone) | PAs | vanillin-HCl, ferrous sulphate, DMB[a] and H-V[b] | McKenzie and Peterson (1995a,b) |
| Pinaceae (*Pseudotsuga menziesii*) | inner cortical cell walls | PAs | DMACA[c] | Weiss et al. (1999) |
| Salicaceae (*Populus tremula* x *alba*) | root cap, white zone, cortex, border cells | PAs | DMACA | Westley (2015) |
| Salicaceae (*Populus tremuloides*) | root cap, epidermal cells of root tip, cortical cells, early elongation zones in root tips | PAs | DMACA and vanillin-HCl staining. | Kao et al. (2002) |
| Rosaceae (*Fragaria* x *ananassa* and *Malus* x *domestica*) | white zone: specifically in the root cap as well as cortical and epidermal cells of the distal root tip | PAs | DMACA | Hoffmann et al. (2012) |
| Fagaceae (*Quercus, Castanopsis and Lithocarpus*) | root cap, the epidermal layer and the endodermis | PAs | vanillin-HCl and DMB | Tam and Griffiths (1993) |
| Lauraceae (*Cinnamomum camphora*) | outer epidermal cells of the root tip | PAs | DMACA | Osawa et al. (2011) |
| Vitaceae (*Vitis vinifera*) hairy root | cell division zone: epidermis, endoderm, and vascular bundles | PAs | DMACA | Terrier et al. (2009) |
| Malvaceae (*Gossypium hirsutum*) | root endodermis and hypodermis, maturation zone, endodermis | catechins | DMB | Mace and Howell (1974) |
| Theaceae (*Camellia sinensis*) | around the pericycle | catechins | vanillin-HCl | Liu et al. (2009) |

[a] DMB is dimethoxybenzaldehyde.

[b] H-V test is Hoepfner-Vorsatz test, containing sodium nitrate, acetic acid, urea and NaOH.

[c] DMACA is 4-dimethylaminocinnamaldehyde.

Although pigmentation has been a key factor in describing the cork zone, the cause of the brown coloration is not well understood or investigated. It has been potentially correlated with root dormancy, cortical breakdown, periderm formation and PA deposition (McKenzie and Peterson, 1995a). Brown pigmentation in mature *Arabidopsis* seed coats has been attributed to the oxidation of PAs and other flavonoids (Lepiniec et al., 2006). Many reports use vanillin-HCl histochemical staining for PA localization. However, this method is limited in its ability to detect PA specifically, since other flavan derivatives can also be detected (Li et al., 1996). Therefore, the brown pigmented cells in the root brown zone that stained positively for PAs by vanillin-HCl may have been due to other flavonoids that are complexed and deposited in cell walls. Meanwhile, it was shown that flavonoids (including PAs) are not present in the white zone of *Pinus* and *Eucalyptus* (McKenzie and Peterson, 1995a). Therefore, presence of PAs is not always correlated with brown pigmentation in roots. In addition, the occurrence of PAs is very clear in root tips, which are typically white. PA polymers and flavan-3-ols monomers were found in the meristematic region of root tips in *Fragaria* x *ananassa* (Hoffmann et al., 2012). The difference in PA localization and correlation with pigmentation provides evidence to suggest that PA containing tissues are not always brown, despite this being a common assumption.

As summarized in Table 1-1, the localization of PAs in roots has been reported in mostly woody species. In *Pinus* and *Eucalyptus*, PAs accumulate in the visibly brown zone at some distance from the root tip. By contrast, *Populus tremula x alba*, *P. tremuloides* and some species of Rosaceae, localize PAs around the root cap as well as in epidermal cells near the root tip and the white zone. This is consistent with PA accumulation in the epidermal cells of stems and leaves in young poplar (Kao et al., 2002). Fagaceae trees also produce PAs in the root cap, epidermal layer, and the endodermis. Other plants such as cotton and tea contain PA precursors such as catechins in the root endodermis and pericycle respectively, but not in the root cap. It is important to note that different staining methods have varying specificities for PAs which may cause ambiguity in determining the nature of the positively stained compounds. To describe the general distribution patterns of PAs in roots, many more species still need to be investigated.

## 3. The potential functions of tannins in roots

The previously described studies have shown that PAs accumulate predominantly in the outer cortical cells and epidermal cells in both active white zones and brown zones of roots. These findings are important when considering the potential roles of poplar root PAs, which include defense, aluminum resistance via chelation of metal cations, interactions with mycorrhizal symbionts, and soil nutrient cycling.

### 3.1 Interactions with herbivores and pathogenic soil organisms

Root-feeding herbivores can reduce plant productivity and increase the chance of fungal and bacterial infection in roots (Hunter, 2001). Even though the diversity of below-ground herbivores is lower than that of above-ground herbivores, below-ground chemical defenses appear to influence above-around tissues (Rasmann and Agrawal, 2008). This was shown by Kaplan *et al.* (2008), who found foliar herbivory induced significant increases in secondary metabolites in *Nicotiana tabacum* leaves, but not in roots. In contrast, root herbivory by nematodes elicited a whole plant response, with elevated secondary metabolite production in both roots and leaves. Anti-herbivore defenses are widely considered roles of PAs due to the protein binding properties described above. Root anti-herbivore effects of PAs were investigated using *Melolontha* larvae, but PAs were found to have no negative effect on growth, development or feeding selection (Sukovata et al., 2015). This is consistent with a lack of effect on the survival rates of *Costelytra zealandica* larvae by PAs (Sutherland et al., 1982).

The localization of PAs in specific cells in the cortex and root epidermis could suggest a protective function against root pathogens. Aggarwal et al. (2018) found decreased root colonization by fungal pathogen *Fusarium oxysporum* and significantly reduced fungal biomass in high-PA chickpea hairy roots which overexpress *Arabidopsis* AtTT2 gene, as compared to wildtypes. So far, there is little evidence for root PAs as anti-pathogen compounds below ground. Further studies on root-rhizosphere interaction may give evidence to the relationship between root PAs and fungal resistance.

### 3.2 Metal chelation and aluminum tolerance

The presence of two or more hydroxyl groups on the phenyl (B) ring of PAs enables

chelation (Quideau et al., 2011), and this binding potential is greater under acidic conditions (Waterman and Mole, 1994). Ionic aluminum (Al) is one of the harmful metals that inhibit root elongation in acidic soils, thus reducing the plant's capacity for water and nutrient absorption and thereby reducing aboveground productivity (Kochian et al., 2004). The secretion of organic anions such as citrate, malate, or oxalate are an Al resistance in annual crops. However, these root exudates can be rapidly decomposed by soil microbes (Jones et al., 1996). In camphor, there is evidence to suggest that PA chelation of $Al^{3+}$ is more effective in detoxifying $Al^{3+}$ rich soils than the release of organic anions (Osawa et al., 2011).

Osawa et al. (2011) discovered a novel Al resistance mechanism based on root PAs to minimize the negative effects of $Al^{3+}$ on root growth. This study compared properties at the surface layer of the root apex in a PA producing species *Cinnamomum camphora* (camphor tree) with non-PA producing species *Glycine max* (soybean) on differing levels of Al in soils. The camphor tree was highly resistant to high levels of $Al^{3+}$, resisting 20-fold higher concentration (500μM) than soybean. Exposure to higher $Al^{3+}$ levels only led to slight induction of citrate release, suggesting that Al resistance in the camphor tree was not dependent on citrate release, and there must be other mechanisms involved. Interestingly, Osawa et al. (2011) detected PA-accumulating cells outside the epidermal cells of camphor roots after Al exposure. This revealed that outer PA-accumulating cells constitutively proliferate under high Al conditions and shield the inner epidermis cells from Al toxicity. The proliferation is followed by the abrupt collapse and detachment of Al-PA rich cells. The Al-damaged cell surfaces were then replaced with new extensive propagation of epidermis cells behind the tip. The constant turnover of PA-accumulating cells may facilitate the expansion of epidermis cells away from toxic effects of $Al^{3+}$ during root elongation in the camphor tree.

The detoxifying role of PAs on other heavy metal ions has also been studied. For example, there is evidence to suggest that root tannins can bind to $Fe^{3+}$ in mangrove roots and form black ferric pigment, providing mangroves with an effective $Fe^{3+}$ tolerance mechanism (Kimura and Wada, 1989). Blackening of living roots was observed in the mangroves with increased biomass, compared to those trees lacking blackened roots.

Hydrolysable tannins, for example oenothein B in roots of *Eucalyptus camaladulensis,* provides a similar mechanism for Al tolerance (Tahara et al., 2014). *Eucalyptus camaladulensis* has the capability to grow in acid sulfate soils with Al concentration reaching mM level and can

accumulate up to 11 mg Al $g^{-1}$ dry weight in root tissue with no symptoms of toxicity. Oenothein B localized in the root symplast can chelate a minimum of four Al ions per molecule by the free hydroxyl groups on the dimer (Tahara et al., 2014). However, the subcellular localization of oenothein B in root requires further investigation. Oenothein B is not an Al specific binding agent, but does preferentially bind Al when multiple metals are present (Tahara et al., 2014).

In summary, there is good evidence that both HTs and PAs can reduce toxic effects of $Al^{3+}$ in some plants, likely through chelation.

### 3.3 Possible roles of PAs in modulating mycorrhizal colonization

Mycorrhizas are defined as associations between symbiotic soil fungi and the absorbing organs of root tissues. They are classified into arbuscular mycorrhizal fungi (AM) which colonize inside root cell walls, and ectomycorrhizal fungi (ECM) which enclose host roots with a mantle and form the Hartig net between the epidermal and cortical cells (Smith and Read, 2008).

Several reports have focused on the role of catechins and PAs as regulators of ectomycorrhizae (ECM) colonization but have shown contradictory results. Weiss *et al.* (1999) looked into the tissue-specific distribution of metabolites in mycorrhizae of *Abies alba* and *Pseudotsuga menziesii*. In these two species, the cell walls of the inner cortex showed a netlike thickening, which stained positively with DMACA. In the apical part of *P. menziesii* roots, soluble flavanols were found in the outer cortical cells. By contrast, only a weak accumulation of flavanols in root hairs and rhizodermis cells was found in non-mycorrhizal roots grown in the greenhouse. The opposite observations were reported by other studies. For example, Weiss *et al.* (1997) performed histochemical studies with larch mycorrhiza (*L. decidua*, Pinaceae - ECM *Boletinus cavipes*) using DMACA and HPLC. Mycorrhization led to rapidly increased concentrations of catechin and epicatechin in the endodermis and the central part of cortex in apical root tissue, as well as in the parenchyma tissue enclosing the apical meristem of the root (Weiss et al., 1997). However, no significant difference in procyanidin levels between noninoculated control roots and mycorrhizas were observed. By contrast, in beech (*Fagus sylvatica*), levels of the soluble phenolics catechin and epicatechin in roots tips are reduced by mycorrhizal colonization, suggesting that this reduction in catechin and epicatechin may be an important step for rapid mycorrhization (Beyeler and Heyser, 1997). Similar results have also been observed following mycorrhiza establishment in *Picea abies-Lactarius deterrimus* and

*Picea abies-Laccaria amethystea* (Munzenberger et al., 1990). Beyler and Heyser (1997) suggested that higher levels of catechin and epicatechin found in nonmycorrhizal roots may function as inhibitors of enzymes needed for ectomycorrhizal formation, and the reduction of their levels in the root tissue leads to the activation of such enzymes. Additionally, procyanidin contents remain constant in this study, so the reduced levels of corresponding monomers are not used to enhance procyanidin polymer content of the roots. Schützendubel and Polle (2002) found that the inoculation of *Paxillus involutus* on Scots pine (*Pinus sylvestris*) root tips didn't affect catechin content compared to non-mycorrhizal root tips. Niemi *et al*., (2007) also found that during the establishment of an ECM symbiosis with *Suillus variegatus* on Scots pine seedling, roots contained remarkably more catechins and PAs at the beginning of the experiment than shoots and showed a tendency to decrease towards the end of the experiment. However, in contrast to shoots, no significant difference of catechins and PAs between non-inoculated and inoculated seedlings was observed in roots. Despite the positive relation between PAs or catechins and mycorrhiza colonization found in study by Weiss *et al*. (1999), there is no direct evidence that catechin or PAs may serve as a defense against fungus infection.

The effect of arbuscular mycorrhizal (AM) fungal colonization on PA synthesis in birdsfoot trefoil (*Lotus japonicus L.*) was investigated by Solaiman and Senoo (2018). Soluble and insoluble PAs concentrations in roots were lower in plants inoculated with AM fungi compared to uninoculated plants, but AM fungus inoculation increased the concentration of flavonols and soluble tannins in shoots (stems and leaves). A potential explanation is that the high tannin concentration of non-mycorrhizal roots may have been induced by deficiency of nutrients in the plant tissue, which is countered by the supply of nutrients from AM mycorrhizal pathway. AM fungi inoculation can be more efficient in supporting the production of primary and secondary metabolites compared with the uptake from soil (Solaiman and Senoo, 2018). Overall, this study lacks direct evidence that PAs impact AM symbiotic systems.

### 3.4 Possible roles of tannins in root nodule symbioses

Host plants that form root nodule symbioses with nitrogen-fixing bacteria belong to a single clade of angiosperms called the Nitrogen-fixing Clade (Soltis et al., 1995). Legumes nodulated by proteobacterial rhizobia are most widely studied in root nodule symbioses and

actinorhizal plants such as *Parasponia sp.* (*Cannabaceae*) can also develop nodules in symbiosis with the actinobacterium *Frankia* (reviewed in Svistoonoff et al., 2014).

So far, all studies on root nodule symbioses focus on isoflavone content and only one study on tannins has been found. Laplaze *et al*. (1999) investigated the involvement of polyphenols in the *Casuarina glauca-Frankia* symbiosis. Histochemical analyses of *C. glauca* nodules revealed that *Frankia* infected cells were surrounded by tannin containing cell layers which located below the periderm, in the endodermis, and in the cortex. Flavans (flavan-3-ols, flavan-4-ols and flavan-3,4-diols) are the major components of the deposits of phenolic compounds found in uninfected nodule cells, while gallic tannins accumulated specifically at the boundaries of the infected cortex. Biochemical studies also showed that nodules accumulated the same compounds as roots but there was a strong increase in flavan content in nodules compared with uninfected roots, suggesting the potential role of flavans in plant-microbe interactions.

### 3.5 Effects of tannins on decomposition and nutrient cycling

Shoot exudates and the *in vivo* biological activities of tannins in shoot exudates has been studied in very few cases. Eucalyptus can secrete red-coloured wood exudates called kinos which are tannin-rich and there is a positive correlation between total tannin content and antibacterial effects (Martius et al., 2012). Exudated root tannins could thus play roles in facilitating belowground interactions with a diverse community of microorganisms and improving the breakdown of soil organic matter. However, roots may not secrete PAs as exudates like the tannin-rich kinos found in wood. And root PAs may not be secreted like other flavonoids as a result of their size and are more likely to enter the soil from storage of epidermal cells. By contrast, the PA precursor, catechin, can be exudated by roots and exert different functions in the root rhizosphere.

Soil tannins, either from root turnover or leaf litter, have the ability to alter soil nutrient cycling and microbial communities and consequently impact plant nutrient availability (Schweitzer et al., 2008). The binding capacity of PAs to positively charged ions like $Ca^{2+}$, $Al^{3+}$ and $Fe^{3+}$ leads to the speculation that PAs may bind and store nutrient cations in the root cells. However, no correlation between $NH_4^+$, $Ca_2^+$ or $NO_3^-$ fluxes and PA concentrations along the roots of *Populus tremula x alba* could be detected (Westley, 2015). PAs can also bind to $NH_2$

groups of organic nitrogen compounds and form large inorganic complexes of low solubility. These are resistant to decomposition, and therefore reduce nitrogen losses through leaching. They may also enhance organic nitrogen availability to mycorrhizal plants (Schweitzer et al., 2008). PA-$NH_2$ complexes reduce the conversion of organic nitrogen to inorganic nitrogen forms ($NH_4^+$ and $NO_3^-$) that plants can typically utilize, which suggests that soils with high PAs levels could have negative impacts on plant growth. In addition, high PA leaf litter produced from *Populus balsamifera* (balsam poplar) has the potential to inhibit the ability of *Alnus rubra* (red alder) to fix atmospheric $N_2$ in nodules and limit the availability of soil inorganic nitrogen that alder could use as a substitute for fixed $N_2$ (Schimel et al., 1998). PAs produced by poplar leaf litter may inhibit $N_2$-fixation in alder nodules by altering microbial activity, and root PA may potentially exert the same function.

Denitrification is a microbially facilitated process where $NO_3^-$ is reduced and ultimately produces $N_2$. This is known as an alternative anaerobic respiration route. This process can be responsible for nitrogen losses from soils and decreased productivity in plants. A recent study by Bardon et al. (2016) purified PAs from roots and rhizomes of *Fallopia spp.,* and this drastically inhibited the denitrification and aerobic respiration of two denitrifying bacteria. This indicated that *Fallopia spp.* may reduce denitrification through the release of PAs and directly targeting the respiratory process in these bacteria. By contrast, monomeric flavan-3-ol forms such as catechins had no effect on denitrification. Furthermore, the effects of *Fallopia* extracts on bacterial respiration were found to be correlated with their protein precipitation capacity rather than their antioxidant properties, again suggesting the involvement of PAs rather than catechin and epicatechin. It was also supposed that lower polymerized procyanidins (≤ tetramers) were more likely to be involved in the biological denitrification inhibition of *Fallopia spp.* than high polymerized forms, due to their higher solubility in water, which may help them target the biological membranes where denitrification occurs (Bardon et al., 2016). Apart from denitrification, PAs could be toxic to microbes through oxidative damage or by directly binding to microbial enzymes and inactivating them. It should be noted that some microbes have adapted to high tannin soils via secretion of tannin-binding polymers and the production of tannin resistant enzymes (Kraus et al., 2003). A recent study has provided evidence that anoxic soil microbiota can depolymerize PA polymers and degrade the subsequent monomers, and the interaction between root PAs and the rhizosphere microbiome suggests root PAs may serve as a

defense barrier but could be degraded and utilized by microbes (McGivern et al., 2021).

In all, PAs might be involved in plant nutrient availability either through binding to soil components or by impacting plant-microbe interactions, leading to the speculation that root PAs could have a physiological role in modulating nutrient uptake. However, whether soil PAs originate from both root and foliar inputs, or PAs in root tip and epidermal cells is an important question. The impact and functions of PAs specifically in roots still needs further investigation.

## 4. Transcriptional regulation of PA synthesis

The PA and flavonoid pathways are regulated transcriptionally, and a host of transcription factors have been characterized in a number of plants. Most important are the R2R3 MYB factors (Liu et al., 2015). Transcriptional activation of the genes for the PA pathway requires not only a MYB factor, which determines the specificity of binding sites, but also a basic helix-loop-helix (bHLH) protein and a WD-repeat (WDR) protein cofactor, which both interact with the MYB. They form the so-called "MBW" complex and bind to the cis-elements in promoters of PA biosynthetic pathway genes to activate transcription (Xu et al., 2015).

Many MYB PA activators are expressed in seeds and fruits and have also been found in vegetative organs such as leaves and roots. There are two major types of MYB PA activators: the TT2-type and PA1-type. TT2/MYB123 was first characterized in the Arabidopsis seed coat (Nesi et al., 2001). PtMYB134 is a homologue of *Arabidopsis* TT2, and is mainly expressed in *Populus* leaves and is also found in roots (Mellway et al., 2009). TT2-type MYB factors LjTT2 family has also been discovered in birdsfoot trefoil (*Lotus japonicus*) (Yoshida et al., 2008). The three LjTT2s showed distinct transcription patterns in different tissues, including leaves, stems and roots. Among three LjTT2s, LiTT2b is most related to TT2 expression and its transcripts appeared to accumulate specifically in roots (Yoshida et al., 2008). VvMYBPA2, the orthologue of Arabidopsis TT2, is mainly expressed in the exocarp of young berries and in the leaves of *Vitis vinifera* (Terrier et al., 2009). In *Trifolium repens* PA accumulation is regulated by TT2-type MYB factors MYB14 in leaf and floral tissues, and leaves in *Trifolium arvense* (Hancock et al., 2012). The other type of MYB PA activator, PA1-type, also shows diverse patterns in different species. In *Populus*, PtMYB115, which belongs to the PA1 type of PA regulator clade, is mainly expressed in leaves and roots (James et al., 2017). VvMYBPA1 is involved in the regulation of the PA pathway during seed development in grapevine (Bogs et al., 2007).

In addition, there are MYBs from other clades that are also involved in PA regulation. VvMYB5b, expressed in both berries and vegetative tissues of grapevine, is a potential regulator for both PAs and anthocyanins, though its specificity still needs to be confirmed (Deluc et al., 2008). In *Medicago trunculata*, MtMYB5 belongs to the general phenylpropanoid clade. It mainly contributes to the regulation of seed PAs and is also expressed in flowers and roots, with less activity in leaves (Liu et al., 2014).

In addition to MYB activators, the involvement of MYB factors with broad repressor functions in phenylpropanoid metabolism have been uncovered (Ma and Constabel, 2019). Other transcription factors involved in PAs regulation are also reported. For example, B-box (BBX) transcription factors (TFs) play crucial roles in the transcriptional regulation of flavonoid biosynthesis (Bai et al., 2019). Recently, BBX23 was identified as a novel regulator of flavonoid biosynthesis in poplar, showing a distinct response to high light. Overexpression of BBX23 in poplar activated expression of MYB TFs, such as MYB115 and MYB119, as well as structural genes in both PA and anthocyanins pathways (Li et al., 2021). Furthermore, Li et al. (2022) characterized a CCCH type transcription factor, PuC3H35, which positively regulates the expression of genes involved in PA and lignin biosynthesis as a response to drought stress in *Populus ussuriensis* root. These recent studies reveal that there are more transcription factors in addition to MYBs that may play a role in regulating PA and other flavonoids. In particular, our understanding of the transcriptional regulatory mechanisms for PA biosynthesis in plant roots is still rudimentary and needs further research.

In previous work we characterized two MYB transcription factors, MYB134 and MYB115, that regulate PA synthesis in poplar leaves. PA biosynthesis can be induced by stresses such as wounding, UV light, and nitrogen deficiency (Osier and Lindroth, 2001; Peters and Constabel, 2002; Mellway et al., 2009). The TT2-type PA-regulator MYB134 was first identified in the Constabel Lab (Mellway *et al.,* 2009) and found to be a key stress-inducible regulator of PAs in leaves and stems of *Populus*. High PA lines have been engineered by overexpressing MYB134 (Mellway *et al.,* 2009). When overexpressed, the MYB134 gene causes the hyperaccumulation of PAs especially in leaves, together with a strong increase in the abundance of transcripts encoding the major flavonoid and PA enzymes. MYB134 is a specific regulator of the PA pathway, since it induces synthesis of PAs, but not anthocyanins, flavonols, and other flavonoids (Mellway *et al.,* 2009; Gesell *et al.,* 2014). High PA lines have also been

engineered by overexpressing MYB115, which likewise causes enhanced accumulation of PAs in leaves (James et al., 2017). Similar to MYB134, MYB115 specifically regulates the PA branch of flavonoid metabolism. Overexpression of MYB115 and MYB134 both lead to reduced salicinoid phenolic glycosides levels, suggesting a trade-off in allocation to PA or salicinoid pathways. In addition, both MYB115 and MYB134 appear to activate their own and each other's promoters, which could provide potential positive feedback loops in PA induction (James et al., 2017). All these studies on MYBs provide new insight into the complexity of PA and flavonoid regulatory network in poplar, but also raise many questions.

In *Populus*, MYB134 and MYB115 are expressed differentially in leaves and roots, with MYB134 dominating in leaves (James et al., 2017), while MYB115 is more highly expressed in roots (James et al., 2017; Wang et al., 2017). The MYB115-overexpressing transgenic *Populus* had slightly higher root PAs levels compared to the wild-type, while the MYB134-overexpressing genotype showed no obvious difference in roots (Westley, 2015). RNAi knockdown lines have also been engineered recently to eliminate MYB134 and create low PA poplar lines (Gourlay et al., 2020). However, there was no reduction in root PA concentrations when MYB134 was suppressed using RNAi, indicating that MYB134 is likely not required for PA synthesis in roots (Gourlay et al., 2020). An independent study on MYB115 suppression found only a small reduction in PAs in leaves of *P. tomentosa* (Wang *et al.,* 2017*),* suggesting this regulator is less important in leaves compared to other tissues. The evidence above suggests that MYB134 is the major PA regulator in leaves. However, in roots MYB115 might be the major regulator of PA synthesis.

## 5. Objectives of the research

This research aims to generate MYB115 and MYB134 single CRISPR knockout as well as double knockout transgenic poplars in order to obtain low PA lines with reduced or eliminated PA levels. These transgenic CRISPR lines can be used to check the effect of knocking-out MYB115 and MYB134 on PA accumulation as well as expression of key flavonoid genes. This can help investigate the function of both MYBs in the regulation of PAs in poplar roots and shoots and help interpret the relationship between MYB115 and MYB134. More specifically, my study used hybrid poplar (two aspen hybrids: *P. tremula x alba* and *P. tremula x tremuloides*) as a model system to investigate the function, especially priority for tissue specificity, of these two

MYBs in regulating PA biosynthesis in poplar. Poplars accumulate large amounts of PAs in roots and poplar leaves, and can synthesize lots of PAs by stress induction. I hypothesized that both MYB134 and MYB115 are key transcription factors for PA biosynthesis in poplar, but that their interactions and functions in roots may differ from those in shoots. A long term objective is to explore the biological effects of PAs in roots. The availability of transgenic *Populus* lines with low or high PA levels will be also very useful for exploring poplar root – microbe interactions.

# Chapter 2: Methods and Materials

## 2.1 Plant growth conditions

*Populus tremula* × *P. tremuloides* (clone INRA 353-38) wild-type (herein referred to as '353') and transgenic MYB134- and MYB115-overexpressing plantlets as well as *P. tremula* × *P. alba* (clone INRA 717-1B4) wild-type plantlets (herein referred to as '717') (Mellway et al., 2009; James et al., 2017) were available in the Constabel lab. All clones were maintained and micropropagated *in vitro* on Lloyd and McCown's Woody Plant Media (WPM; Caisson, North Logan, Utah, United States) supplemented with 1 µM final concentration indole-3-butyric acid (IBA).

## 2.2 Sequencing of MYB115 and MYB134 genes

Approximately 100 mg of fresh frozen leaf tissue from 353 and 717 wild-type tissue cultures were finely ground in liquid nitrogen using mortar and pestle. DNA was extracted using DNeasy Plant Kit (Qiagen, Mississauga, ON, Canada). For PCR amplification, 1–2 µl of the genomic DNA preparation was used as a template with primers flanking the target site for MYB115 and MYB134 respectively (Supplemental Table 1.1). The primers were designed depending on the MYB115 and MYB134 sequences of *P. tremula*, *P. tremuloides* and *P. alba* from online database Popgenie (https://popgenie.org/) and AspenDB (http://aspendb.uga.edu/s717). Amplicons were further purified using EZ-10 Spin Column PCR Product Purification Kit (BioBasic, Markham, Canada) and then Sanger-sequenced (Sequetech, CA, USA). Resulting sequences of 353 wild-type were aligned with *P. tremula* and *P. tremuloides* sequences from UPSC database (v1.1) in Popgenie, while that of 717 wild-type were aligned with *P. tremula* and *P. alba* sequences from sPta717 Variant Database (v2) in AspenDB, using ClustalW in BioEdit. Specific information such as single nucleotide polymorphisms (SNPs) were checked manually. Double peaks and deletions indicated SNPs. In order to separate the two alleles, PCR products were cloned using pUCM-T Cloning Vector Kit (BioBasic) and moved into *E. coli* XL1-Blue. Up to 10 colonies were sequenced to get all potential SNPs in the exon region and checked against the corresponding sequences in databases.

## 2.3 Generation of transgenic poplars with knocked-out MYB115 and MYB134

### 2.3.1 Guide RNA (gRNA) design and vector construction

Potential target gRNA sequences for MYB115 and MYB134 in the *P. tremula* × *P. tremuloides* hybrid (INRA 353-38) and *P. tremula* × *P. alba* hybrid (INRA 717-1B4) were obtained from AspenDB (http:// aspendb.uga.edu/s717). Target gRNA regions without SNPs with the lowest off-target potential and near the 5'-end of the coding sequences in exon regions were selected (Figure 3-3). Potential off-target sites were identified by BLAST searching the gRNA sequences against the *P. tremula x P. alba* clone 7171-B4 genome using the Probe Search function in the sPta 717 Variant Database (http:// aspendb.uga.edu/s717). If the hits were localized within the coding region of the targeted gene, contained the NGG PAM region, and had less than three mismatches within the core region of the gRNA, they were considered as potential off-target sites and were not selected.

CRISPR constructs were created under the help of David Ma. The constructs were prepared using the vectors previously developed by Jacobs *et al*. (2015) and obtained from Addgene (Addgene.org), with minor modifications as described in Fellenberg et al. (2020). Two types of plasmids were used for transformation. P201H and p201N plasmids with antibiotic markers were used for whole plant transformation, while p201G containing an eGFP marker was used for hairy root transformation. P201H and p201N were double digested with *Swa*I and *Spe*I and p201G with *Spe*I and *Apa*I, before being purified by gel extraction. MtU6 and scaffold DNA fragments were PCR amplified from the pUC gRNA shuttle vector using the primers 35S*Spe*I_MtU6F/MtU6R and ScaffoldF/*Apa*I_ScaffoldR for p201G (for p201N or H, I used *Swa*I_MtU6F and *Spe*I_ScaffoldR instead), creating 20 nt overhangs with backbone to enable DNA assembly (Yoshida et al., 2015). Linearized plasmid of above vectors, purified MtU6, scaffold and the single-stranded sgRNA of MYB115 or MYB134 with 20 nt overhangs on both ends were combined and incubated for 1 h at 50°C using the NEBuilder DNA Assembly Kit (New England BioLabs). The reaction products (1 µl) were transformed into *E. coli* XL1-Blue via electroporation and colonies were grown on LB Kan50. Purified plasmid was sequenced in order to confirm correct assembly.

To assemble two gRNAs into one vector for each construct (all constructs contain two gRNAs) the first gRNA fragment was amplified with primers 35S*Spe*I_MtU6F and UNS1_ScaffoldR for p201G plasmid, while the first gRNA fragment for p201N or H plasmids

were amplified with *Swa*I_MtU6F and UNS1_ScaffoldR. The second p201G gRNA fragment was amplified with UNS1_MtU6F and *Apa*I_ScaffoldR, while p201N or H was used with *Spe*I_Scaffold R. Linearized plasmid, two purified gRNA fragments with 20 nt overhangs on both ends were combined and transformed into *E. coli* XL1-Blue as described above. Both p201G or p201H-MYB115 and p201G or p201N-MYB134 contained two gRNAs targeting two exon regions respectively, while p201G- or p201N-double-knockout contained one gRNA from MYB115 exon1 and another gRNA from MYB134 exon1. Purified plasmids were sequenced again and confirmed constructs were further electroporated into *Agrobacterium tumefaciens* strain GV3101 (MP90) for p201N or H plasmids and *Agrobacterium rhizogenes* strain ARqua1 for p201G plasmids. Positive transformants were identified by colony PCR with 35S*Spe*I_MtU6_F/*Apa*I_Scaffold_R for p201G plasmids and *Swa*I_MtU6F/*Spe*I_ScaffoldR for p201N or H plasmids. The primers for amplifying each fragment are shown in Supplemental Table 1.2.

### 2.3.2 Plant transformation to obtain transgenic hairy roots and whole plants

Hairy roots were generated on *in vitro* cultured *P. tremula* × *P. alba* clone 717-1B4 as described previously (Yoshida et al., 2015; Fellenberg et al., 2020). Medium-sized leaves were excised and placed on preculture solid medium (0.25 g of MES, 0.1 g of myo-inositol, 30 g of Suc, and 4.33 g of MS [Caisson] salts in 1 L, pH 5.7) for 16 h. These leaves were wounded by cutting the leaf veins gently before transformation with *Agrobacterium rhizogenes* strain ARqua1 (Yoshida et al., 2015). Transformed colonies, containing P201G-MYB115, p201G-MYB134, p201G-double-knockout and p201G (empty vector), were grown on MG/L (5 g $L^{-1}$ tryptone, 2.5 g $L^{-1}$ yeast extract, 5.2 g $L^{-1}$ NaCl, 2.0 g $L^{-1}$ glutamic acid, 10 g $L^{-1}$ mannitol, 0.2 g $L^{-1}$ $MgSO_4 \cdot 7H_2O$, and 0.5 g $L^{-1}$ $K_2HPO_4$, 0.8 µg/L biotin)-agar medium at 28°C with 50 mg $L^{-1}$ spectinomycin and 50 mg $L^{-1}$ kanamycin for vector selection. Liquid cultures were grown in MG/L medium for 16 h before collecting cells by centrifugation at 3500rpm for 20 min. The bacterial pellet was resuspended in induction broth (0.25 g of MES, 0.1 g of myo-inositol, 30 g of Suc, and 4.33 g of MS salts (Caisson Labs, Smithfield, UT) in 1 L, pH 5.7-5.9) with 200 mM acetosyringone. A suspension with optical density of 0.6 to 0.8 was used to inoculate excised leaves. Wounded explants were shaken in the *Agrobacterium* suspension at 100 rpm for 90-100

min at 28ºC. Transformed leaves were maintained on solid antibiotic-containing medium (preculture medium with 500 mg $L^{-1}$ cefotaxime and 300 mg $L^{-1}$ timentin) for at least 2 weeks. After hairy roots appeared on transformed leaves, positive transformants were selected via GFP fluorescence under UV light (395 nm) using an Olympus SZX7 Zoom Stereomicroscope. Positive roots were excised and sub-cultured on hairy root medium (WPM supplemented with 0.25 mg $L^{-1}$ indole-3-butyric acid and 0.25 mg $L^{-1}$ 1-naphthaleneacetic acid) in the dark at 25°C until sufficient mass for analysis was obtained. For analysis, approximate 100 mg of hairy roots were harvested, and flash frozen in liquid nitrogen for DNA extraction and chemical analysis.

For generating transgenic whole plants, *P. tremula* × *P. tremuloides* clone (INRA 353-38) leaf explants were transformed as described previously (Mellway et al., 2009; James et al., 2017). Medium-sized leaves were excised and placed on preculture solid medium for two days. These leaves were wounded by cutting the leaf veins gently before transformation with *Agrobacterium tumefaciens* strain GV3101, containing P201H-MYB115, p201N-MYB134, p201N-double-knockout and empty vector, respectively. Positive transformed lines were selected by growth on hygromycin or kanamycin-containing selecting medium (callus, shoot and root induction medium) (Yoshida et al., 2015; Fellenberg et al., 2020). After transgenic plantlets rooted on antibiotic-containing medium, the leaves were harvested for DNA extraction and positive transformants were confirmed by PCR using primers that amplify each exon of MYB115 and MYB134 genes, respectively (Supplemental Table 1.1) and sent for sequencing. Transgenics were maintained and micropropagated on solid WPM as described above prior to acclimation to the greenhouse for further analysis.

### 2.3.3 Plant growth conditions and stress treatments

Prior to characterizing the phenotypes of whole plant transgenics, tissue culture plantlets were transferred to peat moss-based soil-less mix (Sunshine Mix 4, Victoria, BC) and acclimated in a mist chamber for one month. The plants were then transferred to gallon pots with additional slow-release fertilizer (Major and Constabel, 2006). Plants were kept in the greenhouse under 16 h daylight, with temperatures between 18 °C and 26 °C. Plants were grown in the greenhouse for another two weeks before being moved into a growth chamber specially equipped with LEDs for high-light treatments. Potted plants were placed in saucers and watered daily. The average daily

light intensity was set to 500 - 600 μmol $m^{-2}$ $s^{-1}$ for the first 5 days, and then increased to 1000 μmol $m^{-2}$ $s^{-1}$ for 15 days. Plants were rotated every other day to minimize light differences in the growth chamber. After 20 days exposure to high-light, samples including apex (three small leaves as well as the stem were included), medium leaves (leaf plastochron index (LPI) 10-12), young stems (around 10cm under the apex) and young roots (around 5cm in white zone containing root tips) were harvested, flash frozen in liquid nitrogen and stored at -80°C for further experiments. The plant tissues were then ground into fine powder and freeze-dried prior to phenolic extraction.

### 2.4 DNA extraction and genotyping

DNA extraction of positive hairy roots and transgenic plants, target gene amplification and sequencing were performed as described in 2.2. To check the CRISPR/Cas9 system knockout efficiency and characterize the mutation types, genomic DNA was extracted, and amplicons were PCR-amplified, with primers targeting each exon of MYB115 and MYB134 genes respectively (Supplemental Table 1.1), from each transgenic line to include SNPs to identify both parental alleles. The PCR products containing two gRNA sites were sequenced directly and analyzed using the web-based tool TIDE (https://tide.nki.nl/) to preliminarily identify mutations and distinguish bi-allelic from monoallelic mutants. Distinct allele-specific SNPs around the gRNA target region were used to confirm that both parental alleles had been amplified and mutated in these sites. For ambiguous sequence results that could not be unambiguously identified by TIDE, PCR products were then cloned into T-vector (pGEM®-T Easy Vector, Promega, USA) for allele-specific sequencing to determine the mutation obtained on each allele. The rates of edited hairy roots or whole plants were calculated based on the PCR subcloning sequencing results.

### 2.5 Histochemical staining

To detect PAs and flavan-3-ols, fresh plant tissues (including apex, stem, young leaf and root) were harvested after two months of growth in the mist chamber and stained for 5 min with DMACA (1% 4-dimethylaminocinnamaldehyde w/v in ethanol:6 N HCl, 1:1, v/v), following de-

staining in 95% ethanol (Mellway et al., 2009). Images were recorded using a SeBaCam™ digital camera (Laxco) mounted on a stereo microscope (Laxco, Washington, U.S.).

### 2.6 Phytochemical extraction, butanol-HCl tannin assay and Ultra-Performance Liquid Chromatography analysis

The butanol-HCl method (Porter et al., 1986) was performed to quantify condensed tannins in different poplar tissues as described previously (Yoshida et al., 2015; Gourlay and Constabel, 2019). Freeze-dried tissue was ground in liquid nitrogen and around 50 mg powder was weighed in a 2-ml CryoTube. Ground tissues were extracted in 1.5 ml of 100% MeOH and four steel beads were added to the mixture, which was homogenized using a Precellys tissue homogenizer (Bertin Technologies, Rockville, MD, USA) for two cycles (each cycle for 45 s at 5000 Hertz per min). This was followed by sonication for 10min in a sonicating water bath and further centrifugation for 10 min at 15,000g, after which the supernatant was collected. The extractions were repeated twice more with an additional 1 ml MeOH, resulting 3.5 ml of extract in total. For tannin assays, reaction mix containing 66.7 µl iron reagent (2% w/v $FeNH_4(SO_4)_2$ in 2 N HCl) and 2ml butanol–HCl (95:5 v/v) was added to 500 µl of MeOH extract and mixed well. 200 µl of the mixture was removed to 96-well plates for use as unheated blanks. The remaining reaction mixture was incubated in a 95 °C water bath for 45 mins then allowed to cool to room temperature. Again, 200 µl of the assay mixture was moved to 96-well plates and read at 550nm using a PerkinElmer Victor X5 2030 Multilabel Plate Reader (Turku, Finland). The absorbance readings of unheated blanks were subtracted from that of the heated mixture for background correction. Tannin concentration was calculated using standard curves using with purified poplar CTs (Mellway et al., 2009).

The remaining 3 ml of methanol extract mentioned above was dried in a SpeedVac Plus concentrator (ThermoSavant SC110A) for several hours. The dried extracts were sonicated in 50% HPLC-grade MeOH to resuspend to a final concentration of 80 mg tissue DW/ml methanol. UPLC-MS and analysis was performed as described in Fellenberg et al. (2020). The suspension was filtered through a syringe filter (4 mm, 0.2 µm PTFE; Phenomenex), and 1.0 µl of each sample was injected into a Waters Acquity UPLC System coupled to an Acquity PDA eLambda Detector and an Acquity QDa single quadrupole mass spectrometer (Waters, Milford, MA, USA). To separate phenolic compounds, two solvents were used: solvent A (ddH2O with 0.1%

formic acid) and solvent B (acetonitrile with 0.1% formic acid). The flow rate of the gradient was set at 0.5 ml $min^{-1}$. The detailed gradient profile, UV wavelength range and MS conditions were described by Fellenberg et al. (2020). Optimized SIR methods and negative ionization mode were used for collecting MS data and MS analysis. Catechin, epicatechin and salicin (Sigma-Aldrich) as well as tremulacin, tremuloidin and salicortin (gifts from Dr. Richard Lindroth, University of Wisconsin) were used as standards to create calibration curves using serial dilutions (1000–5 ng $\mu l^{-1}$). The SIR method targeted masses, and optimized cone voltages for salicinoids were recorded in Fellenberg et al. (2020). The targeted mass and optimized cone voltage for catechin and epicatechin is 289 m/z, 16 V. MS data were processed in TargetLynx (version 4.1) software, and each compound from each sample was quantified by measuring the peak area.

## 2.7 RNA extraction, cDNA synthesis, and RT-qPCR analysis

Total RNA was extracted from both poplar roots and leaves using the method modified from Muoki et al. (2012) as described in Yoshida et al. (2015) and James et al. (2017). Approximately 50 to 75 mg of frozen ground leaf or root tissue was used for RNA extraction. To remove genomic DNA, RNA preparations were treated with RQ1 DNase (Promega, Madison, WI, USA). First strand cDNA was synthesized using Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA).

RT-qPCR was performed in 20µl reactions with Luna® Universal qPCR Master Mix (NEB, Ipswich, MA, USA). Each reaction contained 100ng of 1:10 diluted cDNA template, 0.5µl of 10 mM forward and reverse primers, 5µl master mix, and DEPC water. QPCR was performed on a CFX96 Real Time system and C1000 Thermocycler (Bio-Rad), following the thermocycling protocol provided by NEB. For each sample, no-reverse transcriptase reactions were included as negative controls. All transcript expression data were normalized against the geometric mean of poplar elongation factor1b (EF1b; Potri.001G224700) and actin (Potri.001G309500) expression, and were quantified by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). Primers are shown in Supplemental Table 1.3. Annealing temperature for each pair of primers was optimized for highest primer efficiency, which was estimated based on the slope of the standard curve created with a series dilution of templates (Yoshida et al., 2015).

## 2.8 Fungal cultures and disc inhibition assay

Six different isolates of *Armillaria* (*A. ostoyae* and *A. sinapina*) were obtained from Dr. Renate Heinzelmann (Forest Sciences, University of British Columbia) and were maintained on 3% malt extract medium at 25ºC in the dark. For propagation, a plate of fungus mycelium was blended with liquid potato dextrose broth (PD broth) in a sterilized cup and the mixture was cultured for 14 days at 100 rpm on a shaker. The culture in PD broth was homogenized again for 30s and was spread on 1% malt extract agar. The plates were incubated at 25ºC for 7days in the dark, before being used for propagation or disc assay.

The dikaryotic vegetative mycelium of the ectomycorrhizal fungus *Laccaria bicolor* strain S238N (obtained from Dr. Judith Felten, Umea University, Sweden) was maintained on modified Pachlewski medium P5 at 25ºC in the dark (Deveau et al., 2007; Felten et al., 2009).

Both *Armillaria* and *Laccaria bicolor* hyphal inhibition by PAs and catechin was tested using disc diffusion assays according to the method described in Foss *et al.* (2014). Agar plates were centrally inoculated with a mycelial plug from each culture and incubated at room temperature in the dark until actively growing (approximately 3-7days). Discs (0.5 $cm^2$) were cut from Whatman™ grade 1 qualitative filter paper (Fisher Scientific, ON, Canada) and autoclaved. Purified poplar PAs available from the Constabel lab and commercial catechin were dissolved in methanol. PA or catechin solutions (50 mg/mL and 100 mg/mL) were applied to discs to give a final dose of 500, 800 and 1000 µg PAs or catechin/disc, and the methanol was allowed to evaporate. Discs treated with same amount of methanol were used as negative solvent controls. Positive controls were treated with 10 µg hygromycin/disc, a known antifungal agent applied in MeOH. Treatment discs were placed at an equal distance from the edge of actively growing culture and incubated at room temperature in the dark. Hyphal inhibition was assessed by measuring the zone of inhibition surrounding each disc after an incubation of 48h.

## 2.9 Mycorrhizal co-culture sandwich assay

For ectomycorrhiza formation on in vitro poplar by *L. bicolor*, we followed the modified sandwich coculture method developed by Felten et al. (2009). Assays were carried out by Simon Petley. Poplar stem cuttings of wild-types and double-knockouts were propagated and precultured on woody plant medium (WPM) for a week to get rooted plantlets, which were

transferred to square petri dishes (100 x 100 x 20 mm, Sarstedt, Germany) with WPM half covered with a cellophane membrane and cultured vertically for another 3 weeks. In parallel, mycelium plugs of *L. bicolor* S238N were placed on cellophane-covered P20 (low-carbon/ sugar-reduced Pachlewski medium, pH 5.8, buffered with 1 g $L^{-1}$ MES sodium salt) in similar square petri dishes and grown for 10 days. For cocultures, plants were first transferred to square petri dishes containing P20 half covered by cellophane. Then the half piece of mycelium-covered cellophane membrane was placed fungus side down on the poplar roots for direct interaction. The plates were sealed with microtapes or parafilms and the lower part of the square petri dish was covered with a black plastic bag to avoid light. The inoculated cultures were arranged vertically in the growth chamber, which was maintained under the same conditions as the poplar plant cultures. Poplar plants covered with a cellophane membrane without fungal mycelium were setup at the same stage as a control.

To determine the extent of poplar root colonization by *L. bicolor*, around 10 individual poplar plants (three or four per square petri dish) were observed every week after inoculation using a stereomicroscope (Laxco, Washington, U.S.). All root tips were counted, infected root tips were quantified, and root morphology was recorded by images. The observation and counting of root tips were done by Honours student Simon Petley.

## 2.10 Statistical analyses

Data were analyzed by t-tests or one-way analysis of variance (one-way ANOVA) with factors for genotype. Tukey's honest significant difference (Tukey's HSD) post-hoc tests were used to determine which specific groups' means (compared with each other) were significantly different ($P = 0.05$). All statistical analysis was done in R (version 4.1.1).

# Chapter 3: Results

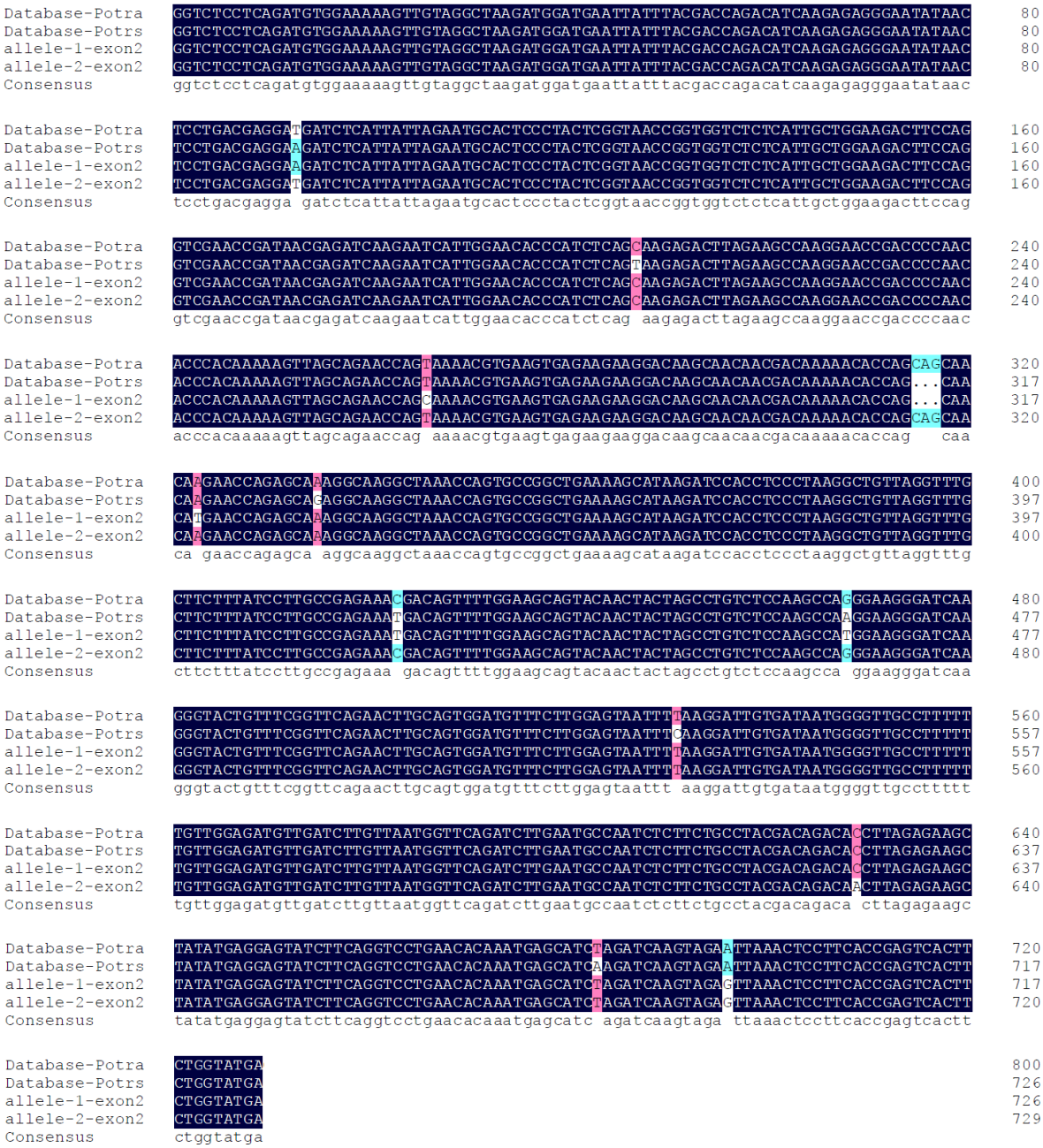
## 3.1 Validation of MYB115 and MYB134 target genes

MYB115 and MYB134 genes were cloned with primers based on the genome sequences of *P. tremula*, *P. tremuloides* and *P. alba* from online databases Popgenie and AspenDB and sequenced (Sequetech, CA, USA). Sequences were aligned to the database and single nucleotides polymorphisms (SNPs) in both alleles were identified. For *P. tremula* × *P. tremuloides* hybrid INRA 353-38 (WT353), SNPs in MYB115 and MYB134 genes were not matched with those from Popgenie and were updated as highlighted in red in Figure 3-1(A)(B) and Figure 3-2(A)(B). For *P. tremula* × *P. alba* hybrid INRA 717-1B4 (WT717), SNPs were confirmed and matched with those from AspenDB as highlighted in blue in Figure 3-1(C)(D) and Figure 3-2(C)(D).

A


Database-Potra ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCGGAGAAGATGCATTGCTTACCGA 80
Database-Potrs ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCAGAGAAGATGCATTGCTTACCGA 80
allele-1-exon1 ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCAGAGAAGATGCATTGCTTACCGA 80
allele-2-exon1 ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCAGGGAAGATGCATTGCTTACCGA 80
Consensus atgggaagggctccttgttgctctaaggtgggtttgcgtagaggaccatggactccc g gaagatgcattgcttaccga
Database-Potra GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCAAAAAAATCT 160
Database-Potrs GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCTAAAAAATCT 160
allele-1-exon1 GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCTAAAAAATCT 132
allele-2-exon1 GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCTAAAAAATCT 132
Consensus gtatatccaggctcacggcgagggccattggaggtctttgcc aaaaaatct

B


Database-Potra GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC 80
Database-Potrs GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC 80
allele-1-exon2 GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC 80
allele-2-exon2 GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC 80
Consensus ggtctcctcagatgtggaaaaagttgtaggctaagatggatgaattatttacgaccagacatcaagagagggaatataac
Database-Potra TCCTGACGAGGATGATCTCATTATTAGAATGCACTCCCTACTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG 160
Database-Potrs TCCTGACGAGGAAGATCTCATTATTAGAATGCACTCCCTACTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG 160
allele-1-exon2 TCCTGACGAGGAAGATCTCATTATTAGAATGCACTCCCTACTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG 160
allele-2-exon2 TCCTGACGAGGATGATCTCATTATTAGAATGCACTCCCTACTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG 160
Consensus tcctgacgagga gatctcattattagaatgcactccctactcggtaaccggtggtctctcattgctggaagacttccag
Database-Potra GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGCAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC 240
Database-Potrs GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGTAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC 240
allele-1-exon2 GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGCAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC 240
allele-2-exon2 GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGCAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC 240
Consensus gtcgaaccgataacgagatcaagaatcattggaacacccatctcag aagagacttagaagccaaggaaccgaccccaac
Database-Potra ACCCACAAAAAGTTAGCAGAACCAGTAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAGCAGCAA 320
Database-Potrs ACCCACAAAAAGTTAGCAGAACCAGTAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAG...CAA 317
allele-1-exon2 ACCCACAAAAAGTTAGCAGAACCAGCAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAG...CAA 317
allele-2-exon2 ACCCACAAAAAGTTAGCAGAACCAGTAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAGCAGCAA 320
Consensus acccacaaaaagttagcagaaccag aaaacgtgaagtgagaagaaggacaagcaacaacgacaaaaacaccag caa
Database-Potra CAAGAACCAGAGCAAAGGCAAGGCTAAACCAGTGCCGGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG 400
Database-Potrs CAAGAACCAGAGCAGAGGCAAGGCTAAACCAGTGCCGGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG 397
allele-1-exon2 CAAGAACCAGAGCAAAGGCAAGGCTAAACCAGTGCCGGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG 397
allele-2-exon2 CAAGAACCAGAGCAAAGGCAAGGCTAAACCAGTGCCGGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG 400
Consensus ca gaaccagagca aggcaaggctaaaccagtgccggctgaaaagcataagatccacctccctaaggctgttaggtttg
Database-Potra CTTCTTTATCCTTGCCGAGAAACGACAGTTTTGGAAGCAGTACAACTACTAGCCTGTCTCCAAGCCAGGGAAGGGATCAA 480
Database-Potrs CTTCTTTATCCTTGCCGAGAAATGACAGTTTTGGAAGCAGTACAACTACTAGCCTGTCTCCAAGCCAAGGAAGGGATCAA 477
allele-1-exon2 CTTCTTTATCCTTGCCGAGAAATGACAGTTTTGGAAGCAGTACAACTACTAGCCTGTCTCCAAGCCATGGAAGGGATCAA 477
allele-2-exon2 CTTCTTTATCCTTGCCGAGAAACGACAGTTTTGGAAGCAGTACAACTACTAGCCTGTCTCCAAGCCAGGGAAGGGATCAA 480
Consensus cttctttatccttgccgagaaa gacagttttggaagcagtacaactactagcctgtctccaagcca ggaagggatcaa
Database-Potra GGGTACTGTTTCGGTTCAGAACTTGCAGTGGATGTTTCTTGGAGTAATTTTAAGGATTGTGATAATGGGGTTGCCTTTTT 560
Database-Potrs GGGTACTGTTTCGGTTCAGAACTTGCAGTGGATGTTTCTTGGAGTAATTTCAAGGATTGTGATAATGGGGTTGCCTTTTT 557
allele-1-exon2 GGGTACTGTTTCGGTTCAGAACTTGCAGTGGATGTTTCTTGGAGTAATTTTAAGGATTGTGATAATGGGGTTGCCTTTTT 557
allele-2-exon2 GGGTACTGTTTCGGTTCAGAACTTGCAGTGGATGTTTCTTGGAGTAATTTTAAGGATTGTGATAATGGGGTTGCCTTTTT 560
Consensus gggtactgtttcggttcagaacttgcagtggatgtttcttggagtaatttaaggattgtgataatggggttgccttttt
Database-Potra TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACACCTTAGAGAAGC 640
Database-Potrs TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACACCTTAGAGAAGC 637
allele-1-exon2 TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACACCTTAGAGAAGC 637
allele-2-exon2 TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACAACTTAGAGAAGC 640
Consensus tgttggagatgttgatcttgttaatggttcagatcttgaatgccaatctcttctgcctacgacagaca cttagagaagc
Database-Potra TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCTAGATCAAGTAGAATTAAACTCCTTCACCGAGTCACTT 720
Database-Potrs TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCAAGATCAAGTAGAATTAAACTCCTTCACCGAGTCACTT 717
allele-1-exon2 TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCTAGATCAAGTAGAGTTAAACTCCTTCACCGAGTCACTT 717
allele-2-exon2 TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCTAGATCAAGTAGAGTTAAACTCCTTCACCGAGTCACTT 720
Consensus tatatgaggagtatcttcaggtcctgaacacaaatgagcatc agatcaagtaga ttaaactccttcaccgagtcactt
Database-Potra CTGGTATGA 800
Database-Potrs CTGGTATGA 726
allele-1-exon2 CTGGTATGA 726
allele-2-exon2 CTGGTATGA 729
Consensus ctggtatga

C

```
Database-tremula    ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCGGAGAAGATGCATTGCTTACCGA     80
Database-alba       ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCAGAGAAGATGCATTGCTTACAGA     80
tremula-exon1       ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCGGAGAAGATGCATTGCTTACCGA     80
alba-exon1          ATGGGAAGGGCTCCTTGTTGCTCTAAGGTGGGTTTGCGTAGAGGACCATGGACTCCCAGAGAAGATGCATTGCTTACAGA     80
Consensus           atgggaagggctccttgttgctctaaggtgggtttgcgtagaggaccatggactccc gagaagatgcattgcttac ga

Database-tremula    GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCAAAAAAATCT                                 132
Database-alba       GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCTAAAAAATCT                                 132
tremula-exon1       GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCAAAAAAATCT                                 132
alba-exon1          GTATATCCAGGCTCACGGCGAGGGCCATTGGAGGTCTTTGCCTAAAAAATCT                                 132
Consensus           gtatatccaggctcacggcgagggccattggaggtctttgcc aaaaaatct
```

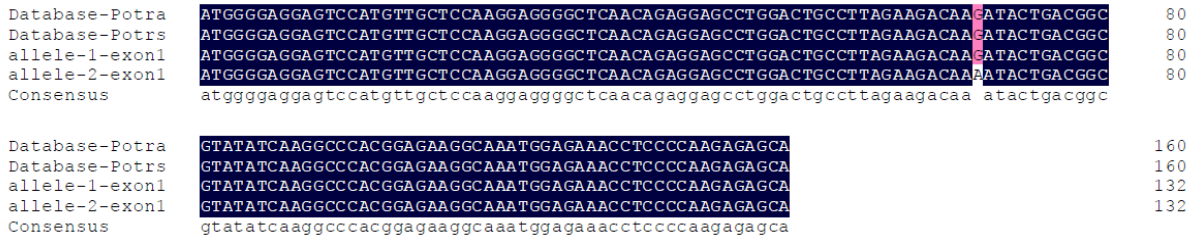
D

```
Database_tremula    GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC     80
Database_alba       GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC     80
tremula-exon2       GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC     80
alba-exon2          GGTCTCCTCAGATGTGGAAAAAGTTGTAGGCTAAGATGGATGAATTATTTACGACCAGACATCAAGAGAGGGAATATAAC     80
Consensus           ggtctcctcagatgtggaaaaagttgtaggctaagatggatgaattatttacgaccagacatcaagagagggaatataac

Database_tremula    TCCTGACGAGGATGATCTCATTATTAGAATGCACTCCCTACTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG    160
Database_alba       TCCTGACGAGGATGATCTCATTATTAGAATGCACTCCCTGCTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG    160
tremula-exon2       TCCTGACGAGGATGATCTCATTATTAGAATGCACTCCCTACTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG    160
alba-exon2          TCCTGACGAGGATGATCTCATTATTAGAATGCACTCCCTGCTCGGTAACCGGTGGTCTCTCATTGCTGGAAGACTTCCAG    160
Consensus           tcctgacgaggatgatctcattattagaatgcactccct ctcggtaaccggtggtctctcattgctggaagacttccag

Database_tremula    GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGCAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC    240
Database_alba       GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGCAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC    240
tremula-exon2       GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGCAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC    240
alba-exon2          GTCGAACCGATAACGAGATCAAGAATCATTGGAACACCCATCTCAGCAAGAGACTTAGAAGCCAAGGAACCGACCCCAAC    240
Consensus           gtcgaaccgataacgagatcaagaatcattggaacacccatctcagcaagagacttagaagccaaggaaccgaccccaac

Database_tremula    ACCCACAAAAAGTTAGCAGAACCAGTAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAGCAGCAA    320
Database_alba       ACCCACAAAAAGTTAGCAGAACCAGTAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAGCA...A    317
tremula-exon2       ACCCACAAAAAGTTAGCAGAACCAGTAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAGCAGCAA    320
alba-exon2          ACCCACAAAAAGTTAGCAGAACCAGTAAAACGTGAAGTGAGAAGAAGGACAAGCAACAACGACAAAAACACCAGCA...A    317
Consensus           acccacaaaaagttagcagaaccagtaaaacgtgaagtgagaagaaggacaagcaacaacgacaaaaacaccagca   a

Database_tremula    CAAGAACCAGAGCAAAGGCAAGGCTAAACCAGTGCCGGGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG    400
Database alba       CAAGAACAGAGCAAAGGCAAGGCTAAACCAGTGCCAGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG    397
tremula-exon2       CAAGAACCAGAGCAAAGGCAAGGCTAAACCAGTGCCGGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG    400
alba-exon2          CAAGAACAGAGCAAAGGCAAGGCTAAACCAGTGCCAGCTGAAAAGCATAAGATCCACCTCCCTAAGGCTGTTAGGTTTG    397
Consensus           caagaa cagagcaaaggcaaggctaaaccagtgcc gctgaaaagcataagatccacctccctaaggctgttaggtttg

Database_tremula    CTTCTTTATCCTTGCCGAGAAACGACAGTTTTGGAAGCAGTACAACTACTAGCCTGTCTCCAAGCCAGGGAAGGGATCAA    480
Database_alba       CTTCTTTATCCTTGCCGAGAAATGACAGCTTTGGAAGCAGTACAGCTACTAGCCTGTCTCCAAGCCAAGGAAGGGATCAA    477
tremula-exon2       CTTCTTTATCCTTGCCGAGAAACGACAGTTTTGGAAGCAGTACAACTACTAGCCTGTCTCCAAGCCAGGGAAGGGATCAA    480
alba-exon2          CTTCTTTATCCTTGCCGAGAAATGACAGCTTTGGAAGCAGTACAGCTACTAGCCTGTCTCCAAGCCAAGGAAGGGATCAA    477
Consensus           cttctttatccttgccgagaaa gacag tttggaagcagtaca ctactagcctgtctccaagcca ggaagggatcaa

Database_tremula    GGGTACTGTTTCGGTTCAGAACTTGCAGTGGATGTTTCTTGGAGTAATTTTAAGGATTGTGATAATGGGGTTGCCTTTTA    560
Database_alba       GGGTACTGTTTCGGTTCAGAACTTCCAGTGGATGTTTCTTGGAGTAATTTTAAGGATTGTGATAATGGGGTTGCCTTTTT    557
tremula-exon2       GGGTACTGTTTCGGTTCAGAACTTGCAGTGGATGTTTCTTGGAGTAATTTTAAGGATTGTGATAATGGGGTTGCCTTTTA    560
alba-exon2          GGGTACTGTTTCGGTTCAGAACTTCCAGTGGATGTTTCTTGGAGTAATTTTAAGGATTGTGATAATGGGGTTGCCTTTTT    557
Consensus           gggtactgtttcggttcagaactt cagtggatgtttcttggagtaattttaaggattgtgataatggggttgccttttt

Database_tremula    TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACACCTTAGAGAAGC    640
Database_alba       TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACACCTTAGAGAAGC    637
tremula-exon2       TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACACCTTAGAGAAGC    640
alba-exon2          TGTTGGAGATGTTGATCTTGTTAATGGTTCAGATCTTGAATGCCAATCTCTTCTGCCTACGACAGACACCTTAGAGAAGC    637
Consensus           tgttggagatgttgatcttgttaatggttcagatcttgaatgccaatctcttctgcctacgacagacaccttagagaagc

Database tremula    TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCTAGATCAAGTAGAATTAAACTCCTTCACCGAGTCACTT    720
Database_alba       TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCAAGATCAAGTAGAATTAAACTCCTTCACCGAGTCACTT    717
tremula-exon2       TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCTAGATCAAGTAGAATTAAACTCCTTCACCGAGTCACTT    720
alba-exon2          TATATGAGGAGTATCTTCAGGTCCTGAACACAAATGAGCATCAAGATCAAGTAGAATTAAACTCCTTCACCGAGTCACTT    717
Consensus           tatatgaggagtatcttcaggtcctgaacacaaatgagcatc agatcaagtagaattaaactccttcaccgagtcactt

Database_tremula    CTGGTATGA                                                                            729
Database_alba       CTGGTATGA                                                                            726
tremula-exon2       CTGGTATGA                                                                            729
alba-exon2          CTGGTATGA                                                                            726
Consensus           ctggtatga
```

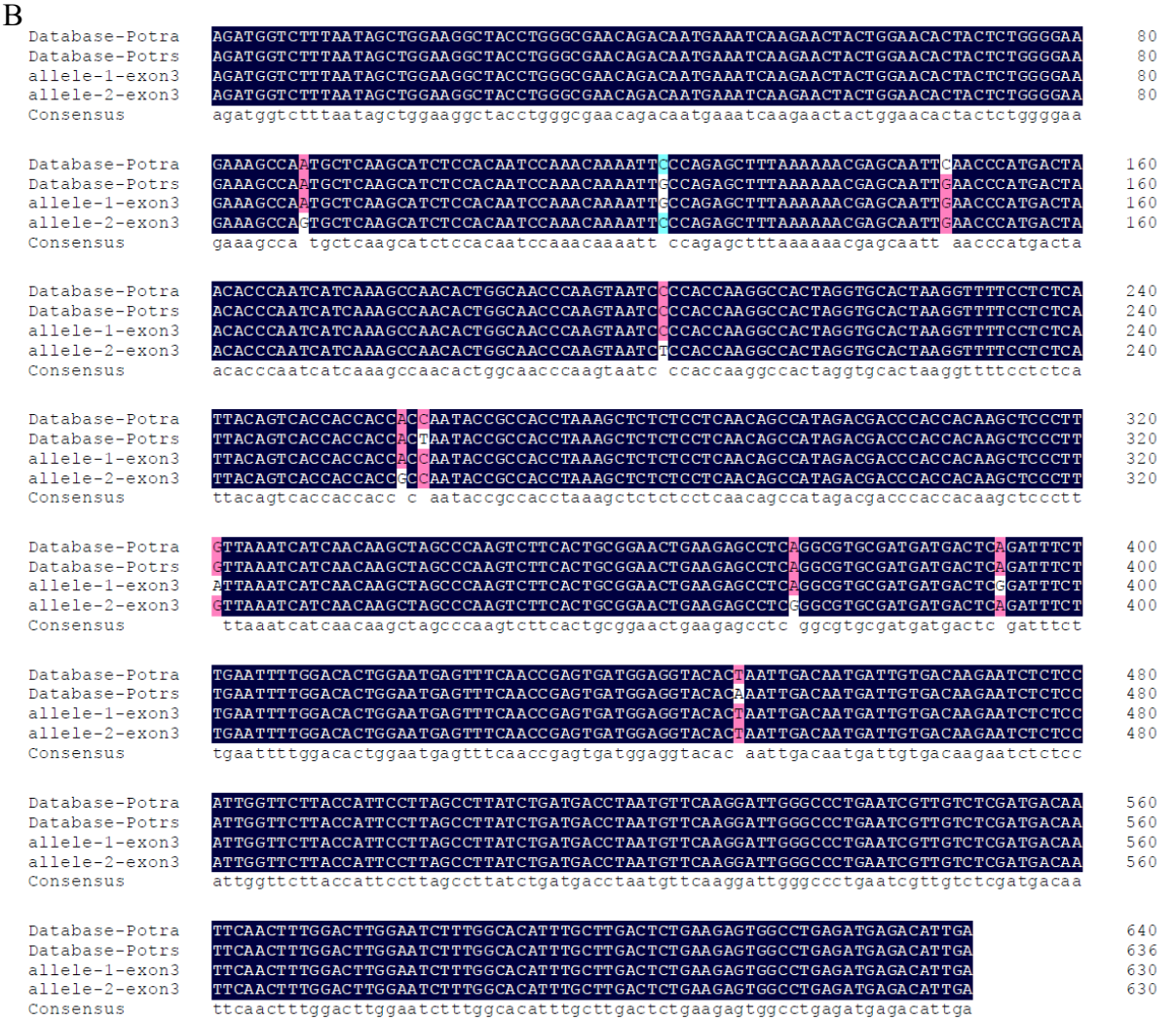
**Figure 3-1: Poplar MYB115 DNA sequence obtained from *P. tremula* × *P. tremuloides* hybrid INRA 353-38 (WT353) and *P. tremula* × *P. alba* hybrid INRA 717-1B4 (WT717).** (A) MYB115 exon1 sequence from WT353, (B) MYB115 exon2 sequence from WT353, (C) MYB115 exon1 sequence from WT717, (D) MYB115 exon2 sequence from WT717. SNPs are

highlighted in red or blue. Red = not matched with Database, blue = matched with Database. Database, sequence acquired from database; Potra, *P. tremula* allele; Potrs, *P. tremuloides* allele; alba, *P. alba* allele.

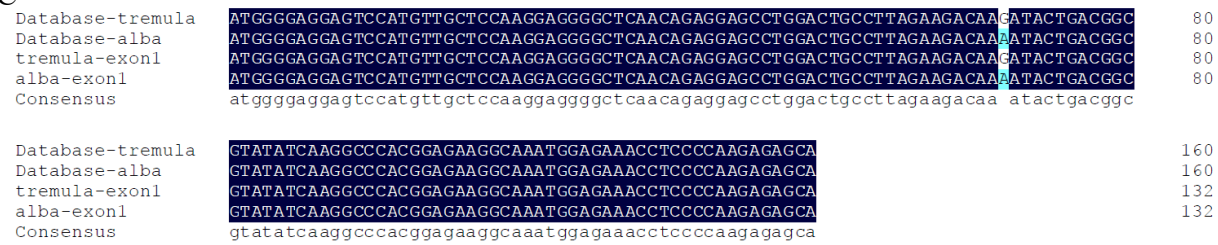
A




B



C




D

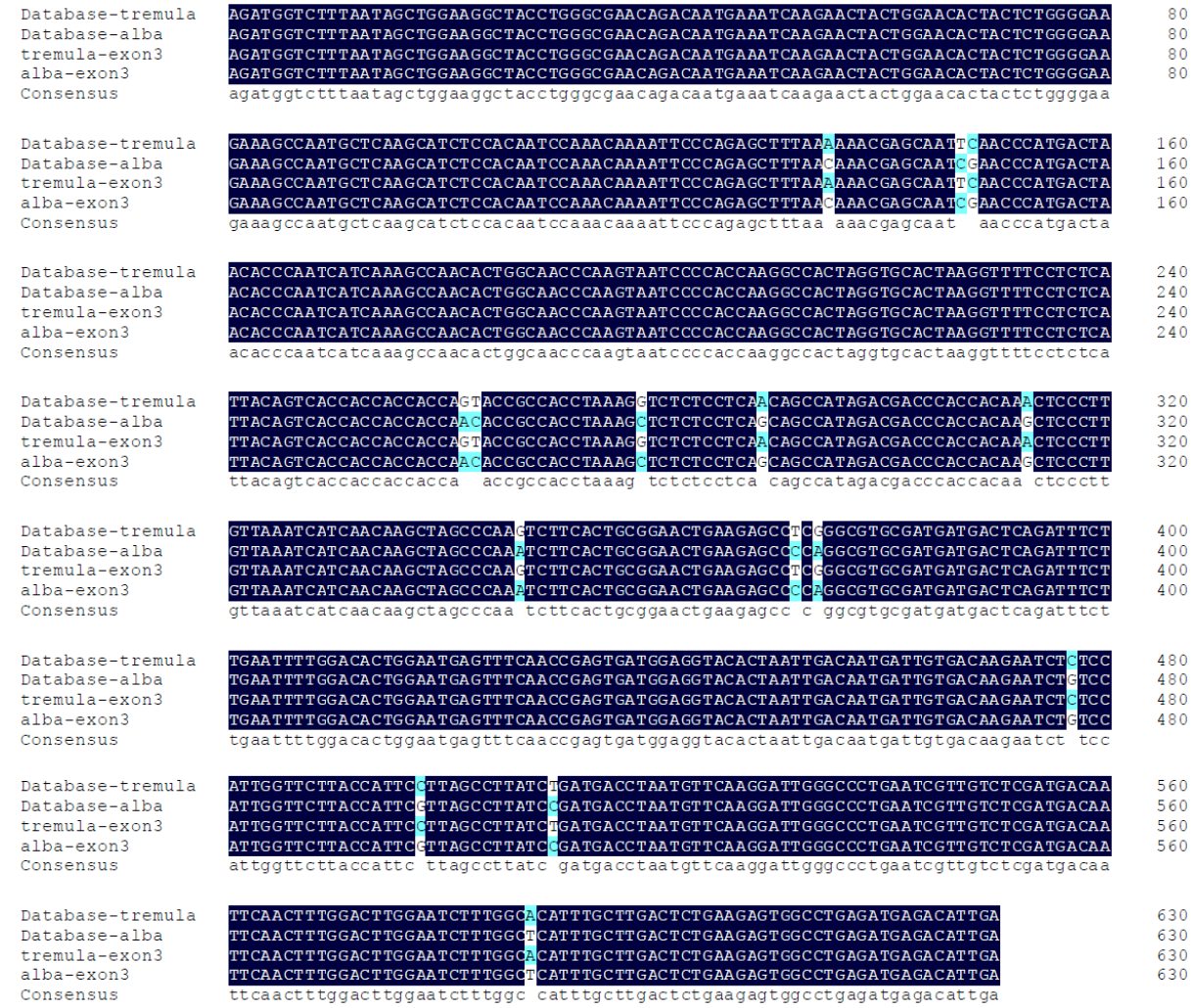



**Figure 3-2: Poplar MYB134 DNA sequence obtained from WT717 and WT353.** (A) MYB134 exon1 sequence from WT353, (B) MYB134 exon3 sequence from WT353, (C) MYB134 exon1 sequence from WT717, (D) MYB134 exon3 sequence from WT717. SNPs are highlighted in red or blue. Red = not matched with Database, blue = matched with Database. Database, sequence acquired from database; Potra, *P. tremula* allele; Potrs, *P. tremuloides* allele; alba, *P. alba* allele.

## 3.2 Generation of MYB115 and MYB134 CRISPR mutants in poplar hairy roots and whole plants

### 3.2.1 Design of gRNA and making constructs

In order to test the regulation roles and tissue specific priority of MYB115 and MYB134 in PA synthesis *in vivo*, CRISPR/Cas9 was applied to knock out these genes. The interspecific poplar hybrids I used for transformation are highly heterozygous, making selection of suitable gRNAs challenging, as chosen target sequences must be identical in both alleles due to the specificity of the CRISPR system. This means that chosen target sites should not contain any allele-specific single nucleotide polymorphisms (SNPs). With confirmed sequences of wild-type MYB115 and MYB134 genes as displayed in Figure 3-1 and Figure 3-2, potential targeted gRNA sequences without SNPs, either confirmed or updated, were selected from the pre-designed gRNAs dataset provided by Zhou *et al.* (2015). Two different gRNAs for both MYBs were chosen, approximately 700-800 bp apart within the coding sequence, to create larger deletions and more effectively disrupt the protein. The gRNAs designed are shown in Figure 3-3. Single gRNA1 and sgRNA2 were designed for MYB115 gene exon1 and exon2 region, respectively, and sgRNA3 and sgRNA4 were designed for MYB134 gene exon1 and exon3 region, respectively. Each gRNA was further compared to the poplar genome by BLAST to test for potential off-target sites within other genes. No matches were found by BLAST. Then each pair of two target gRNAs was synthesized and cloned into P201H/N or P201G plant transformation vectors created by Jacobs *et al.* (2015), which contain both Cas9 and a selective marker gene. The two types of plasmids used for transformation were described in Methods and the vector maps are shown in Figure 3-4. A double knockout construct was also created, which contains two gRNAs, sgRNA1 and sgRNA3, in order to mutate both MYB genes.

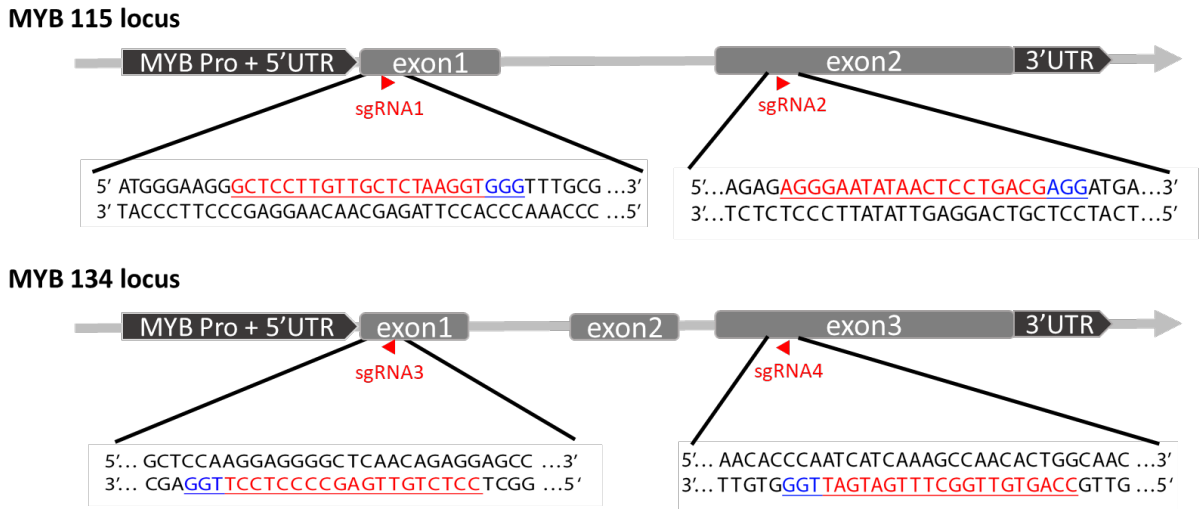
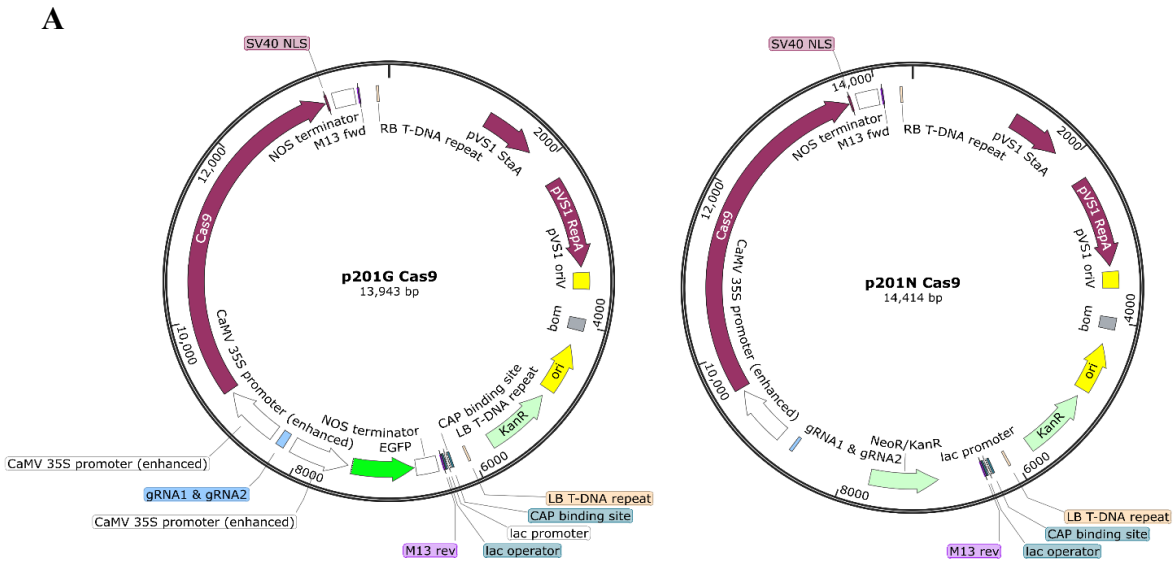



**Figure 3-3: Schematic representation of the MYB genes fragments, indicating the sgRNA target sites and sequences.** Exons and introns are represented by gray boxes and gray lines, respectively. The target sites for each CRISPR/Cas9 nuclease are indicated by red arrows, sgRNA target sequences are indicated in red and underlined, and protospacer adjacent motifs (PAM) are displayed in blue.



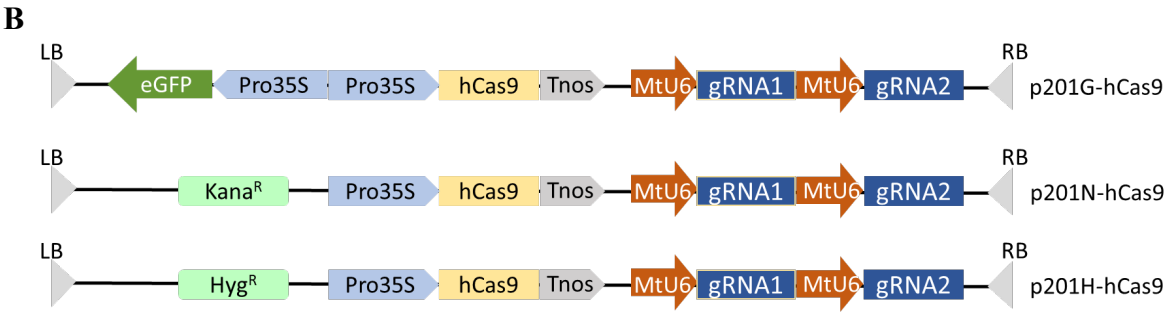



**Figure 3-4: CRISPR/Cas9 binary plasmids for targeted genome mutagenesis.** (A) Examples of two types of vectors targeting two loci simultaneously with eGFP marker in p201G vector and kanamycin resistance marker in p201N vector. (B) Linearized vector map. Pro35S, Cauliflower mosaic virus enhanced CaMV promoter; MtU6, *Medicago truncatula* U6.6 promoter; hCas9, the human codon-optimized Cas9 gene; LB, left T-DNA border; $Kana^R$, kanamycin resistance gene; $Hyg^R$, hygromycin resistance gene; eGFP, eGFP fluorescent marker gene; TNos, NOS terminator region; RB, right T-DNA border.

### 3.2.2 Positive transformants recovered with hairy roots and whole plants

Hairy root transformation is a faster and more efficient method to obtain transgenics material compared to whole plant transformation, and therefore this system is a good test of gRNA specificity and efficiency for whole plant CRISPR knockouts. Additionally, poplars accumulate large amounts of PAs in roots, an advantage of the hairy root expression system for exploring the roles of MYBs in PA biosynthesis. Simultaneously, whole plant transformation was conducted to generate low-PA lines and to explore the phenotypes in other tissues.

Positive hairy root transformants were screened by green fluorescence under UV light due to the eGFP marker in the p201G plasmid, indicating the presence of T-DNA, Cas9 gene and gRNAs. Hairy roots that were not successfully transformed with the construct showed no fluorescence, whereas hairy roots that were successfully transformed showed strong green fluorescence under UV light on the stereoscope (Figure 3-5). GFP-expressing roots were further sub-cultured into three biological replicates on antibiotic-containing solid medium. It took approximately 1 to 2 months for the root to grow and to obtain sufficient biomass for genotyping and phytochemical analysis. Whole plant mutants were selected using the resistance marker. Rooted plants grown on antibiotic containing medium were propagated as individual independently transformed whole plant lines and genotyped before being transplanted into the greenhouse.

Four independent transformation batches for each construct were generated in both hairy roots and whole plants. The numbers of GFP-positive hairy roots and total hairy roots as well as rooted plantlets and calli are presented in Tables 3-1 and 3-2. The ratio of positive transformed hairy roots or plantlets to total transformants were different among batches, and positive transformation ratio of hairy roots was higher than that of whole plants. Additionally, single knockouts were shown to possess higher transformation efficiency or viability than double knockouts.

**Table 3-1: Yield of GFP-positive hairy root lines and total transformants from 4 separate transformation batches.**

| hairy root lines | total GFP-positve hairy roots/ total transformants | batch A | batch B | batch C | batch D |
|---|---|---|---|---|---|
| MYB115-KO | 43/110 (39.1%) | 7/15 (46.7%) | 22/50 (44.0%) | 8/25 (32.0%) | 6/20 (30.0%) |
| MYB134-KO | 45/126 (35.7%) | 24/60 (40.0%) | 8/13 (61.5%) | 5/18 (27.8%) | 8/35 (22.9%) |
| double-KO | 37/139 (26.6%) | 15/70 (21.4%) | 9/23 (39.1%) | 5/22 (22.7%) | 8/24(33.3%) |

**Table 3-2: Yield of rooted plantlets and total calli grown on antibiotic containing medium from 4 separate transformation batches.**

| whole plant lines | total rooted plantlets/ total calli | batch A | batch B | batch C | batch D |
|---|---|---|---|---|---|
| MYB115-KO | 30/138 (21.7%) | 11/45 (24.4%) | 6/32 (18.75%) | 7/26 (26.9%) | 6/35 (17.1%) |
| MYB134-KO | 14/103 (13.6%) | 7/33 (21.2%) | 7/24 (29.2%) | 0/24 (0%) | 0/22 (0%) |
| double-KO | 13/122 (10.7%) | 2/20 (10.0%) | 8/52 (15.4%) | 2/38 (5.3%) | 1/12 (8.3%) |

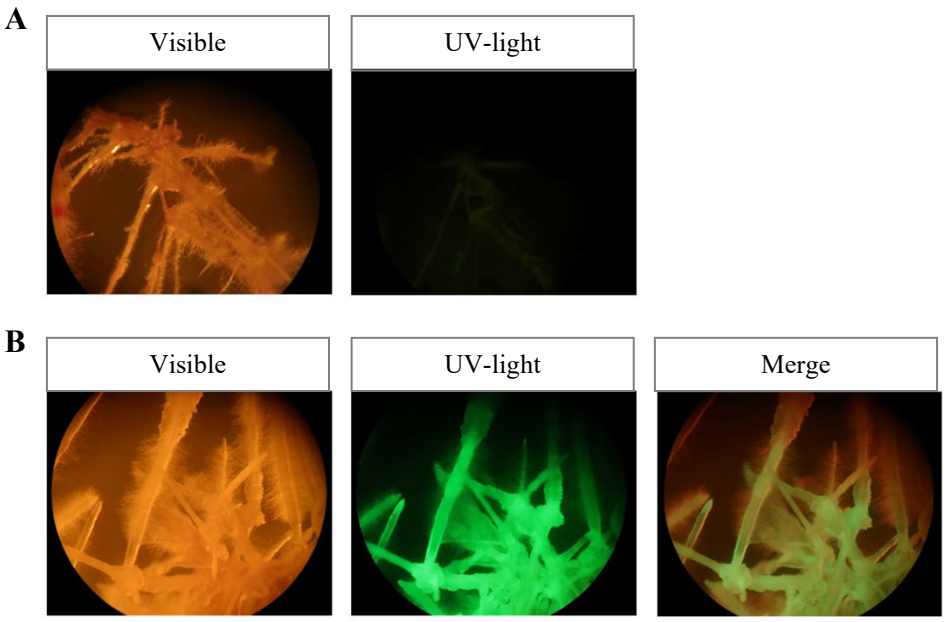



**Figure 3-5: Poplar hairy roots transformed with p201G construct expressing the eGFP marker.** (A) Non-transformed hairy roots with no fluorescence. (B) Successful transformants showing strong green fluorescence under UV light.

## 3.3 Genotyping analysis of CRISPR mutants in poplar hairy roots and whole plants

To check the CRISPR/Cas9 system knockout efficiency and characterize the mutation types, I first extracted genomic DNA from GFP-positive hairy root lines and rooted whole plant transformants and amplified the MYB115 and MYB134 gene target regions including the two gRNA sites. PCR products were first checked by gel electrophoresis to see if there were large deletions between the two gRNAs (Figure3-6). Bands with smaller size represent amplicons with large deletions, which can be found in one of the alleles for some lines. Two MYB134-KO lines contain two smaller sized bands which indicates both alleles contain large deletions, as shown in Figure 3-6(D). The lanes on the gel match the lines listed in Table 3-3.

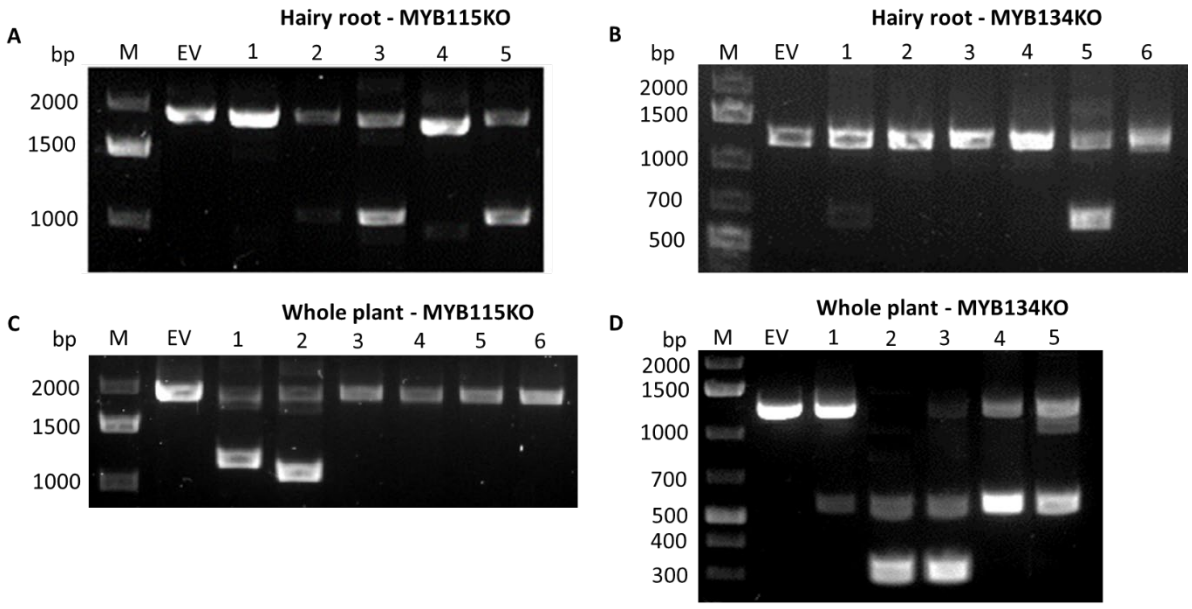



**Figure 3-6: Electrophoresis gels showing PCR amplicons for MYB115 and MYB134 knock-out (KO) transformants.** Detection of deletion mutations in MYB115-KO hairy roots (A) and MYB134-KO hairy roots (B) as well as in MYB115-KO whole plants (C) and MYB134-KO whole plants (D) respectively. The symbols above each lane indicate the transformant lines corresponding to the lines in Table 3-3. M, DNA size marker; EV, empty vector.

PCR amplicons were then sent for direct sequencing and analyzed using the web-based tool TIDE (https://tide.nki.nl/) to preliminarily identify mutations and distinguish bi-allelic from monoallelic mutants. Distinct allele-specific SNPs around the gRNA target region were used to confirm that both parental alleles had been amplified and mutated in this region. For ambiguous sequence results that could not determined by TIDE, PCR products were cloned into T-vectors and a number of replicate colonies were sequenced to ensure that mutations had been obtained for both alleles. Mutation types of both alleles for selected lines are summarized in Table 3-3, 3-4 and 3-5.

**Table 3-3: Prevalent mutation types of MYB115-KO and MYB134-KO hairy root transformants.**

| Target gene | Line | Allele | Indel[a] sgRNA1/3[b] | Indel[a] sgRNA2/4[c] | Frequency[d] |
|---|---|---|---|---|---|
| *MYB 115* | WT 717 | tremula | 0 | 0 | 3/6 |
| | | alba | 0 | 0 | 3/6 |
| | 115-A1 | tremula | 0 | +1 | 2/6 |
| | | alba[e] | -2 | +1 | 1/6 |
| | | alba[e] | 0 | +1 | 3/6 |
| | 115-C5 | tremula | 0 | +1 | 3/6 |
| | | alba | 0 | -2 | 3/6 |
| | 115-B23 | tremula | -1 | +1 | 2/6 |
| | | alba[e] | 0 | +1 | 2/6 |
| | | alba[e] | -806bp | | 2/6 |
| | 115-B17 | tremula | 0 | +1 | 2/4 |
| | | alba | -300 | +1 | 2/4 |
| | 115-B14 | tremula | -806bp | | 1/3 |
| | | alba | -11 | -1, edit 1 | 2/3 |
| *MYB 134* | WT 717 | tremula | 0 | 0 | 3/6 |
| | | alba | 0 | 0 | 3/6 |
| | 134-A1 | tremula | 1 | -2 | 1/6 |
| | | alba[e] | -1 | +1 | 2/6 |
| | | alba[e] | -656bp | | 3/6 |
| | 134-A2 | tremula | +1 | -2 | 2/4 |
| | | alba | -1 | +1 | 2/4 |
| | 134-A3 | tremula | +1 | -2 | 4/6 |
| | | alba | -1 | +1 | 2/6 |
| | 134-A10 | tremula | +1 | -2 | 2/5 |
| | | alba[e] | 0 | -5 | 1/5 |
| | | alba[e] | -1 | +1 | 2/5 |
| | 134-A20 | tremula | -656bp | | 2/6 |
| | | alba[e] | -1 | -5 | 2/6 |
| | | alba[e] | 0 | -5 | 2/6 |
| | 134-D2 | tremula | -27 | +1 | 3/6 |
| | | alba | -15 | +1 | 3/6 |

[a] Indel, insertion and or deletion. '+' indicates insertion and '-' indicates deletion. Numbers shown after '+' or '-' represent the inserted or deleted nucleotides length.

[b] Mutation types found at sgRNA1 target site in both alleles from MYB115 and at sgRNA3 target site in both alleles from MYB134.

[c] Mutation types found at sgRNA2 target site in both alleles from MYB115 and at sgRNA4 target site in both alleles from MYB134.

[d] At least four clones per PCR product were sequenced.

[e] Two different mutation types were found in the same allele, known as chimeric mutations.

**Table 3-4: Prevalent mutation types of MYB115-KO and MYB134-KO whole plant mutants.**

| Target gene | Line | Allele | Indel[a] sgRNA1/3[b] | Indel[a] sgRNA2/4[c] | Frequency[d] |
|---|---|---|---|---|---|
| *MYB 115* | WT 353 | allele1 | 0 | 0 | 3/6 |
| | | allele2 | 0 | 0 | 3/6 |
| | 115-E1 | allele1[e] | 0 | -32 | 2/4 |
| | | allele2[e] | deletion 717bp | | 2/4 |
| | 115-A5 | allele1 | -3 | +1 | 3/5 |
| | | allele2 | deletion 805bp | | 2/5 |
| | 115-C2 | allele1 | 0 | +1 | 2/6 |
| | | allele2 | 0 | +1 | 4/6 |
| | 115-C27 | allele1 | 0 | +1 | 3/5 |
| | | allele2 | | -1 | 2/5 |
| | 115-A1 | allele1 | +1 | +1 | 3/4 |
| | | allele2 | -2 | +1 | 1/4 |
| | 115-A4 | allele1 | +1 | 0 | 3/5 |
| | | allele2 | -2 | -4 | 2/5 |
| *MYB 134* | WT 353 | allele1 | 0 | 0 | 3/6 |
| | | allele2 | 0 | 0 | 3/6 |
| | 134-A1 | allele1[f] | deletion 695bp | | 2/7 |
| | | allele1[f] | -1 | +1 | 2/7 |
| | | allele2 | -1 | +1 | 3/7 |
| | 134-A7 | allele1 | deletion 695bp | | 2/4 |
| | | allele2 | deletion 941bp | | 2/4 |
| | 134-A5 | allele1[f] | deletion 695bp | | 3/8 |
| | | allele1[f] | -13 | +1 | 3/8 |
| | | allele2 | deletion 941bp | | 2/8 |
| | 134-B2 | allele1[f] | -4 | +1 | 3/7 |
| | | allele1[f] | deletion 695bp | | 2/7 |
| | | allele2 | deletion 695bp | | 2/7 |
| | 134-B3 | allele1[f] | -199bp | +1 | 2/6 |
| | | allele1[f] | deletion 695bp | | 2/6 |
| | | allele2 | -1 | +1 | 2/6 |

[a] Indel, insertion and or deletion. '+' indicates insertion and '-' indicates deletion. Numbers shown after '+' or '-' represent the inserted or deleted nucleotides length.

[b] Mutation types found at sgRNA1 target site in both alleles from MYB115 and at sgRNA3 target site in both alleles from MYB134.

[c] Mutation types found at sgRNA2 target site in both alleles from MYB115 and at sgRNA4 target site in both alleles from MYB134.

[d] At least four clones per PCR product were sequenced.

[e] Allele1 and allele2 are the two alleles found in WT353, which could not be assigned to a tremula or tremuloides allele based on sequences from databases.

[f] Two different mutation types were found in the same allele, known as chimeric mutations.

**Table 3-5: Prevalent mutation types of double-KO whole plant mutants.**

| Line | Allele | Indel[a] MYB115-sgRNA1 | Frequency | Indel MYB134-sgRNA3 | Frequency[a] |
|---|---|---|---|---|---|
| WT 353 | allele1 | 0 | 3/6 | 0 | 3/6 |
| | allele2 | 0 | 3/6 | 0 | 3/6 |
| d-B1[d] | allele1[b] | +1 T | 3/7 | -1 G | 2/5 |
| | allele2[b] | +1 A | 4/7 | -1 A | 3/5 |
| d-B2 | allele1 | +1 T | 2/6 | +1 A | 5/6 |
| | allele2 | -2 AA | 4/6 | -1 A | 1/6 |
| d-B7 | allele1 | 0 | 3/6 | +1 A | 2/4 |
| | allele2 | +1 G | 3/6 | -1 A | 2/4 |
| d-E | allele1[c] | +1 T | 2/4 | +1 A | 3/6 |
| | allele1[c] | - | - | -1 A | 1/6 |
| | allele2 | +1 A | 2/4 | -1 A | 2/6 |
| d-A/B[d] | allele1[c] | +1 T | 2/5 | -1 G | 4/8 |
| | allele1[c] | - | - | -1 G | 2/8 |
| | allele2 | +1 A | 3/5 | -1 G | 2/8 |

[a] Indel, insertion and or deletion. '+' indicates insertion and '-' indicates deletion. Numbers and letters shown after '+' or '-' represent the nucleotides inserted or deleted.

[b] Allele1 and allele2 are the two alleles found in WT353, which could not be assigned to *tremula* or *tremuloides* alleles based on sequences from databases.

[c] Two different mutation types were found in the same allele, known as chimeric mutations.

[d] d represents double-KO. d-A/B represent double-KO-A and double-KO-B two lines. These two lines have same genotypes.

The rates of edited hairy roots or whole plants were calculated on the basis of the PCR subcloning sequencing results. For hairy root mutants, preliminary results identified seven mutants with only one allele altered and eight transgenics with both alleles altered in all 17 MYB115 GFP-positive lines, and five with only one allele altered and seven with both two alleles altered in all 13 genotyped MYB134 lines (Table 3-6). Altogether, 82.4-92.3% of all GFP-positive hairy roots are edited. However, only 30.8-35.3% of all transgenic events led to biallelic mutations, most of which consisted of 1-5 indel base pairs. Thus, most mutants are monoallelic and some are chimeric with three mutated alleles or with two mutated alleles and one WT allele simultaneous in the same tissue. For both MYB115 and MYB134 hairy root mutants, I identified more indels located in exon2/3 than that in exon1 region. Mutations occurring at gRNA1 and gRNA2 targeted sites accounted for 71.4 and 100% of the total edited hairy roots respectively, while gRNA3 and gRNA4 possess 91.7% and 100% efficiency respectively, as shown in Table 3-7 and Table 3-8, suggesting that gRNA1 and gRNA3 were not as efficient as gRNA2 and gRNA4.

**Table 3-6: Numbers of edited hairy roots and types of mutations in MYB115 and MYB134 knockouts.**

| lines | edited hairy roots/ total genotyped GFP-positve hairy roots | plants with both alleles altered (no WT) | | | plants with one allele altered and one WT | | plants with no alleles altered (WT/WT) |
|---|---|---|---|---|---|---|---|
| | | homozygous (A1/A1[a]) | bi-allelic (A1/A2[a]) | chimeric (A1/A2/A3[a]...) | monoallelic (WT[a]/A1) | chimeric (WT/A1/A2...) | |
| MYB115 KO[b] | 14/17 (82.4%) | 0 | 6 (35.3%)[b] | 2 (11.8%) | 5 (29.4%) | 2 (11.8%) | 2 (11.8%) |
| MYB134 KO[b] | 12/13 (92.3%) | 0 | 4 (30.8%)[b] | 3 (23.1%) | 4 (30.8%)[b] | 1 (7.7%) | 1 (7.7%) |

[a] A1, allele1 with mutations, A2, allele2 with mutations, A3, allele3 with mutations, WT, wild-type allele without mutations.

[b] MYB115-KO and MYB134-KO lines are highlighted in blue and yellow respectively, and the mutation types with highest ratio are also highlighted with the corresponding color.

**Table 3-7: Editing frequency of sgRNA1 and sgRNA2 in edited MYB115-KO hairy roots.**

| lines | total edited hairy roots | types of mutations in sgRNA1 | | | types of mutations in sgRNA2 | | |
|---|---|---|---|---|---|---|---|
| | | none[a] | one allele[b] | both alleles[c] | none | one allele | both alleles |
| MYB115-KO | 14 | 4 | 6 | 4 | 0 | 7 | 7 |
| | | | 10/14 (71.4%) | | | 14/14 (100%) | |

[a] No alleles were edited.

[b] One of the two alleles was edited, and the mutation types can be either indels or large deletions.

[c] Both two alleles were edited, and the mutation types can be either indels or large deletions.

**Table 3-8: Editing frequency of sgRNA3 and sgRNA4 in edited MYB134-KO hairy roots.**

| lines | total edited hairy roots | types of mutations in sgRNA3 | | | types of mutations in sgRNA4 | | |
|---|---|---|---|---|---|---|---|
| | | none[a] | one allele[b] | both alleles[c] | none | one allele | both alleles |
| MYB134-KO | 12 | 1 | 5 | 6 | 0 | 6 | 6 |
| | | | 11/12 (91.7%) | | | 12/12 (100%) | |

[a] No alleles were edited.

[b] One of the two alleles was edited, and the mutation types can be either indels or large deletions.

[c] Both two alleles were edited, and the mutation types can be either indels or large deletions.

For whole plant transgenics, 75 - 100% of all the 24 independent transgenic plants were edited. One out of 12 MYB115 transgenics exhibited monoallelic mutation, another one exhibited chimeric mutation containing two types of indels within the same allele, and 10 out of 12 (83.3%) were biallelic. Seven lines with both alleles altered were found in all 12 genotyped MYB134 transgenic plantlets, most of which are chimeras and only two are bi-allelic (Table 3-9). Again, I found more indels located in gRNA2 than in gRNA1 for MYB115 transgenic whole plants, confirming that gRNA2 possessed higher efficiency than gRNA1 for the MYB115 target gene (Table 3-10). Although MYB115 exhibited higher editing rates than MYB134, the MYB134 CRISPR construct generated more chromosomal fragment deletions between the two gRNAs than MYB115 in transgenic whole plants and both gRNAs designed for MYB134 exhibit 100% editing frequency (Table 3-11). In contrast, short indel mutations were the most prevalent mutations identified in MYB115 transgenics.

**Table 3-9: Numbers of edited whole plants and types of mutations in MYB115 and MYb134 knockouts.**

| lines | edited plants/ total genotyped independent transgenic lines | plants with both alleles altered (no WT) | | | plants with one allele altered and one WT | | plants with no alleles altered (WT/WT) |
|---|---|---|---|---|---|---|---|
| | | homozygous (A1/A1[a]) | bi-allelic (A1/A2[a]) | chimeric (A1/A2/A3[a]...) | monoallelic (WT[a]/A1) | chimeric (WT/A1/A2...) | |
| MYB115 KO[b] | 12/12 (100%) | 0 | 10 (83.3%) | 1 (8.3%) | 1 (8.3%) | 0 | 0 |
| MYB134 KO[b] | 9/12 (75%) | 0 | 2 (16.7%) | 5 (41.7%) | 0 | 2 (16.7%) | 3 (25%) |

[a] A1, allele1 with mutations; A2, allele2 with mutations; A3, allele3 with mutations; WT, wild-type allele without mutations.

[b] MYB115-KO and MYB134-KO lines are highlighted in blue and yellow respectively, and the mutation types with highest ratio are also highlighted with corresponding color.

**Table 3-10: Editing frequency of sgRNA1 and sgRNA2 in edited MYB115-KO whole plants.**

| whole plant lines | total edited plantlets | types of mutations in sgRNA1 | | | types of mutations in sgRNA2 | | |
|---|---|---|---|---|---|---|---|
| | | none[a] | one allele[b] | both alleles[c] | none | one allele | both alleles |
| MYB115-KO | 12 | 7 | 2 | 3 | 0 | 3 | 9 |
| | | | 5/12 (41.7%) | | | 12/12 (100%) | |

[a] No alleles were edited.

[b] One of the two alleles was edited, and the mutation types can be either indels or large deletions.

[c] Both two alleles were edited, and the mutation types can be either indels or large deletions.

**Table 3-11: Editing frequency of sgRNA3 and sgRNA4 in edited MYB134-KO whole plants.**

| whole plant lines | total edited plantlets | types of mutations in sgRNA3 | | | types of mutations in sgRNA4 | | |
|---|---|---|---|---|---|---|---|
| | | none[a] | one allele[b] | both alleles[c] | none | one allele | both alleles |
| MYB134-KO | 9 | 0 | 2 | 7 | 0 | 3 | 6 |
| | | | 9/9 (100%) | | | 9/9 (100%) | |

[a] No alleles were edited.

[b] One of the two alleles was edited, and the mutation types can be either indels or large deletions.

[c] Both two alleles were edited, and the mutation types can be either indels or large deletions.

For double knockout plants, no potential mutants were found in hairy roots (Table 3-12). Twenty-two individual GFP-positive hairy root lines have been sequenced and analyzed by TIDE, however, more than half of them exhibited mutations in neither MYB115 nor MYB134. Very few lines contained monoallelic mutations in one of the MYB genes, while none contained biallelic mutations or contained any type of mutation in both genes. For genotyped individual whole plant lines, only nine of all 22 whole plants were mutated for both MYB115 and MYB134 genes (Table 3-12). Among these nine lines, only five possessed bi-allelic mutations in both MYB115 and MYB134 target regions, analyzed by TIDE and manually compared to empty-vector sequences. Numbers of edited double KO plants and gRNAs efficiency for both genes were recorded in Table 3-12 and Table 3-13.

**Table 3-12: Numbers of edited plants and type of mutations in hairy root (HR) and whole plant (WP) double KO lines.**

| lines | edited/ total[a] | both genes[b] | mutation types | | | | | no genes[e] |
|---|---|---|---|---|---|---|---|---|
| | | | 115 mono[c] | 115 bi[d] | 134 mono[c] | 134 bi[d] | 134 chimeric | |
| Hairy root | 16/22 (72.7%) | 0 | 6 (27.3%) | 0 | 10 (45.5%) | 0 | 0 | 6 (27.3%) |
| Whole plant | 20/22 (90.9%) | 9 (40.9%) | 5 (22.7%) | 7 (31.8%) | 11 (50%) | 3 (13.6%) | 6 (27.3%) | 2 (9.1%) |

[a] The ratio equals to edited plants per total genotyped independent transgenic lines.

[b] The number of plants with both MYB115 and MYB134 genes mutated.

[c] Mono-allelic mutation type.

[d] Bi-allelic mutation type.

[e] The number of plants with no genes mutated.

**Table 3-13: Editing frequency of sgRNA1 in MYB115 gene and sgRNA3 in MYB134 gene for edited double-KO lines.**

| double KO lines | total edited plantlets | types of mutations in MYB115-sgRNA1 | | | types of mutations in MYB134-sgRNA3 | | |
|---|---|---|---|---|---|---|---|
| | | none[a] | one allele[b] | both alleles[c] | none | one allele | both alleles |
| hairy root double-KO | 16 | 10 | 6 | 0 | 6 | 10 | 0 |
| | | | 6/16 (37.5%) | | | 10/16 (62.5%) | |
| whole plant double-KO | 20 | 9 | 5 | 7 | 0 | 14 | 6 |
| | | | 12/20 (60%) | | | 20/20 (100%) | |

[a] No alleles were edited.

[b] One of the two alleles was edited, and the mutation types can be either indels or large deletions.

[c] Both two alleles were edited, and the mutation types can be either indels or large deletions.

## 3.4 Chemical and genetic characterization of hairy root and whole plant mutant phenotypes

### 3.4.1 PA, flavan-3-ol, and salicinoid composition in transformed hairy root lines

To characterize the chemical phenotypes of MYB115-KO and MYB134-KO hairy roots, PA concentrations were measured using the butanol-HCl assay. No obvious differences were found between controls (hairy roots transformed with the empty vector) and MYB115 knock-out

hairy roots (Figure 3-7). The PA content in one-month old hairy root was approximately 30 mg per gram tissue dry weight. Similar results were observed in MYB134 knock-out hairy root compared to the control lines.

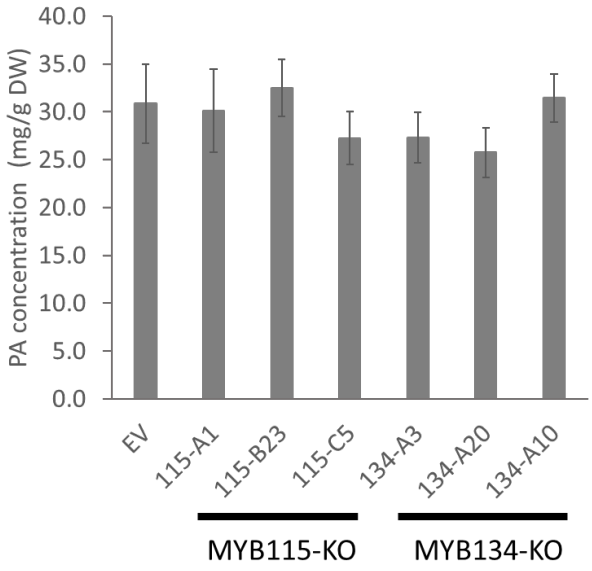



**Figure 3-7: PA concentrations in transformed hairy roots with confirmed genotypes.** One-month old hairy roots for all lines were harvested and phenolics were extracted as described in the Methods and Materials. All data points represent the means of three individual hairy root samples. Knock-out lines were not significantly different from controls (one-way ANOVA; $p < 0.05$). Error bars represent SE. EV, empty vector; 115, MYB115 knock-out lines; 134, MYB134 knock-out lines.

PAs are polymers of flavan-3-ol subunits and the most common subunits in poplar include catechin, epicatechin, epigallocatechin and gallocatechin (James et al., 2017). As these precursors were not detectable in wildtypes by HPLC, we used UPLC-MS to quantify flavan-3-ol subunits in each line. Epicatechin was expressed at much lower level in our tissues compared to catechin. Therefore, we only measured the catechin level in hairy root mutants (Figure 3-8). Catechin was identified based on its retention time, UV spectrum at 280nm as well as MS detection compared to a catechin standard. An example trace of catechin is shown in Supplemental Figure 3.1. No difference in catechin content was observed between control and hairy root mutants, although substantial variability was found among hairy root individuals. This is probably due to the uneven growth, for example, age and size might be different across all samples, though they were propagated at the same time. Salicinoids, such as salicortin,

tremuloidin and tremulacin, were also detected by MS detection and quantified using standards. Hairy root lines showed substantial variability in salicinoids among individuals and none of the hairy root lines created showed significant differences in salicinoid content compared to the empty vector control lines (Supplemental Table 3.1).

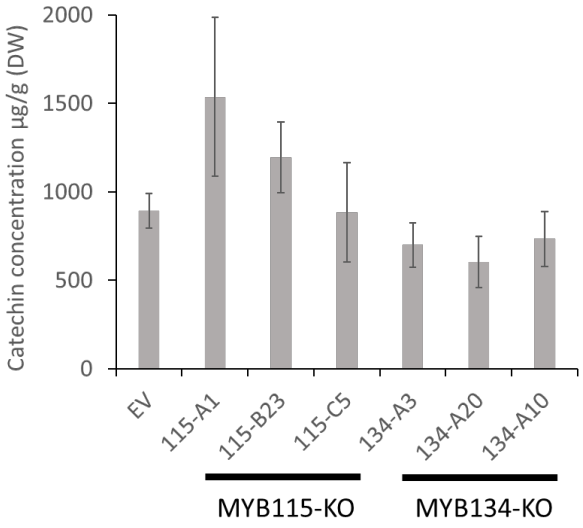



**Figure 3-8: Catechin concentration in CRISPR mutants and control poplar hairy root cultures.** Catechin was quantified based on the UV response at 280 nm. Bars represent means ± SE of three biological replicates per line. Data were analyzed by one-way ANOVA and no significant difference were found among these lines. EV, empty vector; 115, MYB115 knock-out lines; 134, MYB134 knock-out lines.

### 3.4.2 Analysis of gene expression for key flavonoid and PA biosynthetic and transcription factor genes in hairy root lines

To confirm the effect of the MYB knockouts on PA biosynthesis, three knock-out hairy root lines for each construct were selected for gene expression analysis by qPCR. The expression of four flavonoid genes as well as one MYB gene was assayed.

QPCR analysis showed a 1.21-1.57 fold change in MYB134 expression of the three MYB115-KO hairy root lines and only one line (115B23) was significantly affected compared to the control (one-way ANOVA, $p < 0.05$) (Figure 3-9). F3'5'H1, an upstream flavonoid gene, showed some increase in MYB115-KO hairy root lines compared to the control, but the effect was not significant. Similar results were observed for general flavonoid pathway genes DFR2

and PA-specific genes LAR3, and no change for ANR1 expression was observed in MYB115-KO lines (Supplemental Table 4.1).

For MYB134-KO hairy roots, MYB115 gene expression showed a 0.55-0.85 fold change and the effect was significant in two lines as shown in Figure 3-9B (one-way ANOVA, $p < 0.01$). F3'5'H1 expression was not affected in all MYB134-KO hairy roots. DFR2, LAR3 and ANR1 genes expression showed some reduction in MYB134-KO hairy roots compared to the control, but the effect was not significant (Supplemental Table 4.1).

RT-qPCR quantification showed a pattern of flavonoid genes relative expression, though no significant effect was found (Supplemental Table 4.1). All selected MYB115-KO lines showed slightly higher expression level of flavonoid genes which is likely due to the increase of MYB134 transcripts. In contrast, MYB115 transcripts declined in MYB134-KO lines, and all flavonoid genes showed slightly lower expression. MYB expression in hairy roots was not consistent with the results in whole plant root tissues.

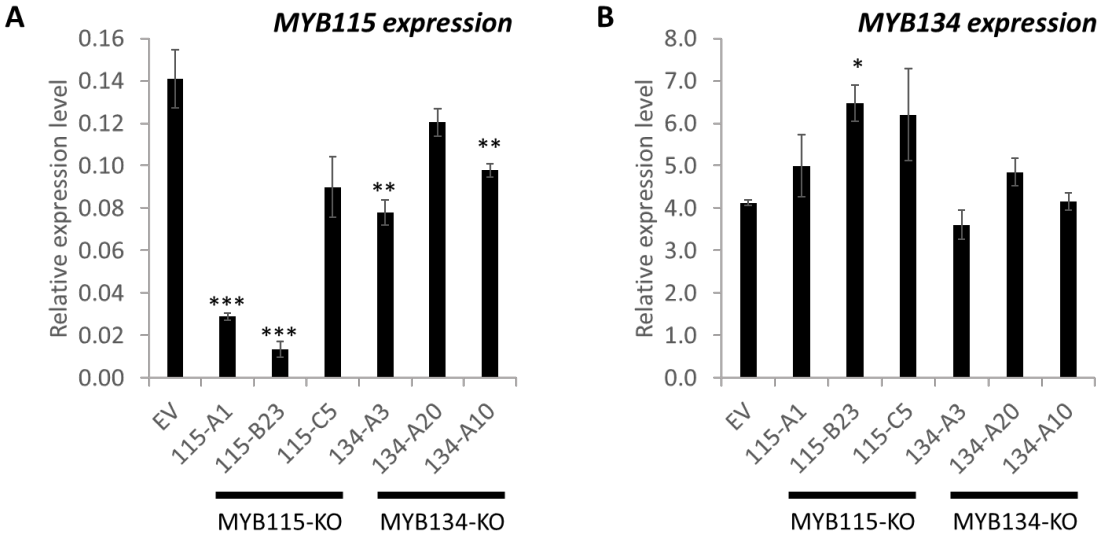



**Figure 3-9: Expression analysis of MYB genes for hairy root mutants.** (A) Relative expression of MYB115 in MYB115-KO and MYB134-KO hairy roots. (B) relative expression of MYB134 in MYB115-KO and MYB134-KO hairy roots. The bars represent means ± SE of at least three individual biological replicates per hairy root line. Asterisks represent statistically significant difference (P-value < 0.05) between CRISPR knockout hairy root lines and empty vector (EV) hairy roots using t-test.

### 3.4.3 Phenotypes of whole plant knock-out mutants treated by high-light after being grown in greenhouse

As *in vitro* plantlets typically do not accumulate measurable PAs in shoots, whole plant mutants and controls from the greenhouse were subsequently treated with high-light stress in a growth chamber to induce PA synthesis (Mellway et al., 2009; Gourlay and Constabel, 2019). The treatment lasted approximately 20 days (Supplemental Table 2.1) and phenotypes of plants were observed. All treated plants were clearly affected by the high light, showing reddish color on top leaves (Supplemental Figure 2.1). Necrosis occurred and severely affected plants lost a substantial number of leaves. There were differences in branch number and necroses level between double knock-out lines and wild-type. Branch number for each plant was counted, and necroses level of the bottom old leaves was assessed. These phenotypes were recorded immediately at the end of the experiment. (Supplemental Figure 2.1, Supplemental Figure 2.2, Supplemental Table 2.2).

### 3.4.4 Whole plant knock-out mutants show reduced tannins and tannin precursor catechin when grown in the greenhouse

To characterize the PA phenotypes of knock-out mutants, they were acclimated in the greenhouse. Young plants from the mist chamber were first stained for PA detection using DMACA (4-dimethylaminocinnamaldehyde). Double-KO and MYB134-KO plants showed less colour than MYB115-KOs and wild-types in stems and leaves, indicating less PA accumulation (Supplemental Figure 6.1). PA concentrations were measured using the butanol HCl assay. As a result of the experimental design and lack of plants, only leaf and root tissues were harvested for MYB115-KO plants, while four tissues including apex, stem, leaf and root tissues were available for MYB134-KO and double-KO plants. The tissues harvested were described in Method and Materials. No differences were observed between wild-type and MYB115-KO mutants in roots. By contrast, some reductions in PAs were detected in leaves for MYB115 knockouts and PA content was significantly reduced in one line 115-A1 (ANOVA, $p \leq 0.05$) (Figure 3-10A). Similar results were observed in MYB134-KO mutants, where PA concentrations in stems and leaves were more affected compared to apex and young roots (Figure 3-10B). Interestingly, all double knock-out mutants showed a strong reduction of PA content in tissues including shoot

(apex, stem, leaf) and young root (Figure 3-10C). In wild-type young roots, PA concentrations reached between 80-90 mg tannin/g dry weight (D.W.), while double knock-out lines accumulated only 20-40 mg tannin/g D.W., which is approximate 25-50% of the wild-type. Stronger effects on PA reduction were observed in MYB134-KO leaves compared to those of MYB115-KO. For further study, we selected three knock-out lines for each construct.

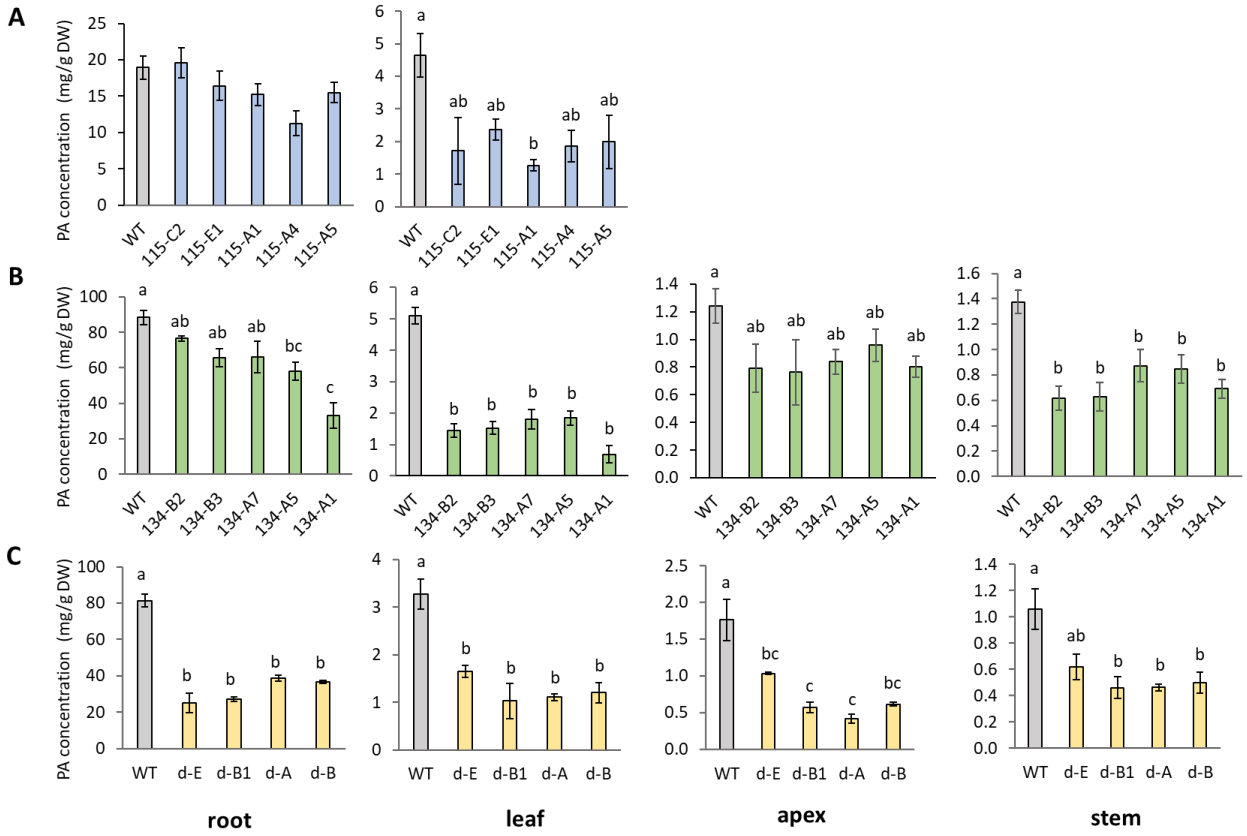



**Figure 3-10: PA concentrations in different tissues for high light exposed poplar plants.** PA content in roots and leaves of MYB115-KO poplars with confirmed genotypes (A), in four different tissues of MYB134-KO poplars with confirmed genotypes (B), and in four different tissues of double-KO poplar (with confirmed genotypes (C) respectively. Three-month old plants in greenhouse were treated with high light for an additional 20 days and tissues were harvested at the same time. Phenolics were extracted as described in the Methods. The bars represent means ± SE of three or four individual biological replicates per line. Data were analyzed by one-way ANOVA followed by Tukey's HSD post hoc test. Different letters on the bar indicate statistically significant differences of PA content according to Tukey's HSD test at 95% confidence (P-value < 0.05). Colors of bars represent different lines: wild-type plants are shown in grey, MYB115-KO lines are shown in blue, MYB134-KO lines are shown in green, double-KO lines are shown in yellow.

Using similar methodology as for the hairy roots, UPLC was used to quantify the PA precursor, catechin, in each whole plant knock-out line. In medium leaf tissues, significant reduction ($p < 0.05$) in catechin content was observed between control and MYB115-KO plants, and more dramatic effects ($p < 0.001$) were found between control and MYB134-KO plants as well as control and double-KO plants (Figure 3-11A). In wild-type plants, catechin concentrations reached between 1.2-2.0 mg/g D.W., while knock-out lines accumulated 0.1-0.5 mg/g D.W. (Figure 3-11A). In root tissues, only double-KO plants and one MYB115-KO line were significantly affected ($p < 0.05$) in terms of catechin content compared to wild-type plants (Figure 3-11A). These results for catechin match the observation for PA quantification in both leaf and root tissues but were more marked for MYB134-KO and double-KO plants in leaves.

The other major flavan-3-ol subunit in poplar, epicatechin, which is a stereoisomer of catechin, was quantified and found to present at a lower level compared to catechin, especially in root tissues (Figure 3-11B). In young root tissues, epicatechin content was only reduced in double-KO plants but the effect was not statistically significant. In leaf tissues, three double-KO lines and two MYB134-KO lines had significantly reduced epicatechin.

Salicinoids, including salicin, salicortin, tremuloidin and tremulacin, were also quantified based on area per dry weight. Only double-KO mutants showed significant differences in all the four salicinoids content compared to the control plants (Figure 3-12). Opposite patterns were observed in the two tissues analysed, as salicinoid content increased in root tissues but decreased in leaf tissues for all double-KO plants. Salicinoid content in MYB115-KO and MYB134-KO was not changed compared to the control plants (Supplemental Figure 5.1).

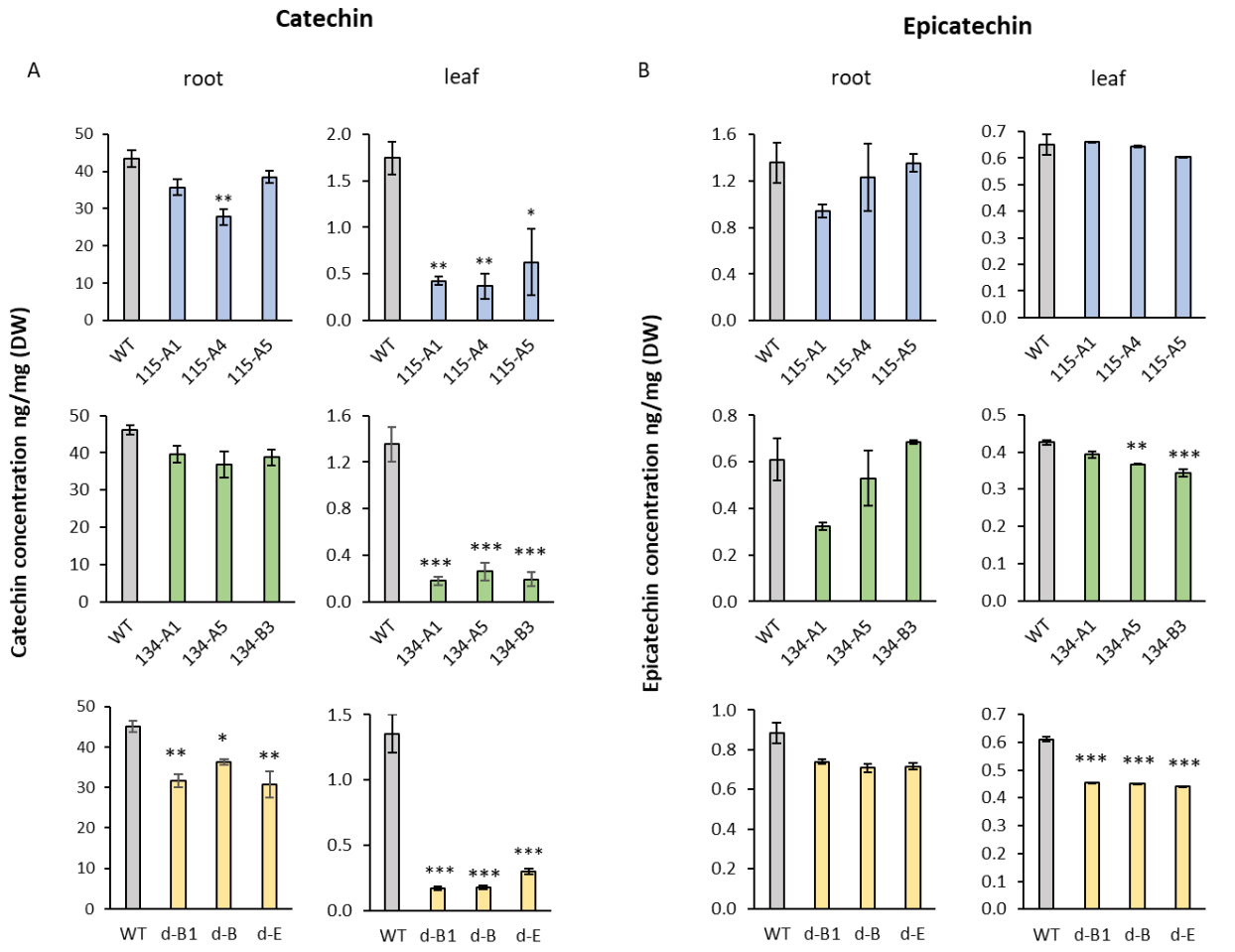



**Figure 3-11: Concentrations of PA precursors in knock-out and wild-type tissues as determined using UPLC in high light exposed poplar plants.** (A) catechin concentrations in leaf and root tissues of MYB115-KO, MYB134-KO and double-KO plants with confirmed genotypes, (B) epicatechin quantification in medium leaf and young root tissues of WT and knockouts with confirmed genotypes. The bars represent means ± SE of at least three biological replicates per line. Significant differences from wild-types were determined using a one-way ANOVA and are indicated by asterisks (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). Colors of bars represent different lines, wild-type plants are shown in grey, MYB115-KO lines are shown in blue, MYB134-KO lines are shown in green, double-KO lines are shown in yellow.

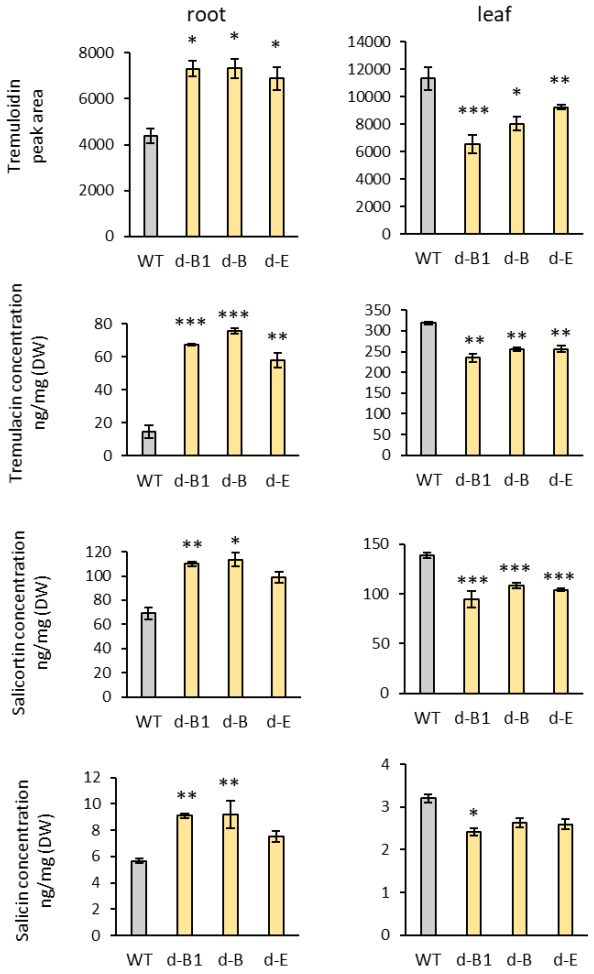



**Figure 3-12: Concentrations of key salicinoids in double-KO tissues as determined using UPLC in high light exposed poplar plants.** The bars represent means ± SE of at least three biological replicates per line. Significant differences from wild-types were determined using a one-way ANOVA and are indicated by asterisks (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). Colors of bars represent different lines, wild-type plants are shown in grey, and double-KO lines are shown in yellow.

### 3.4.5 Gene expression analysis of key flavonoid biosynthetic genes and transcription factors in whole plant mutants

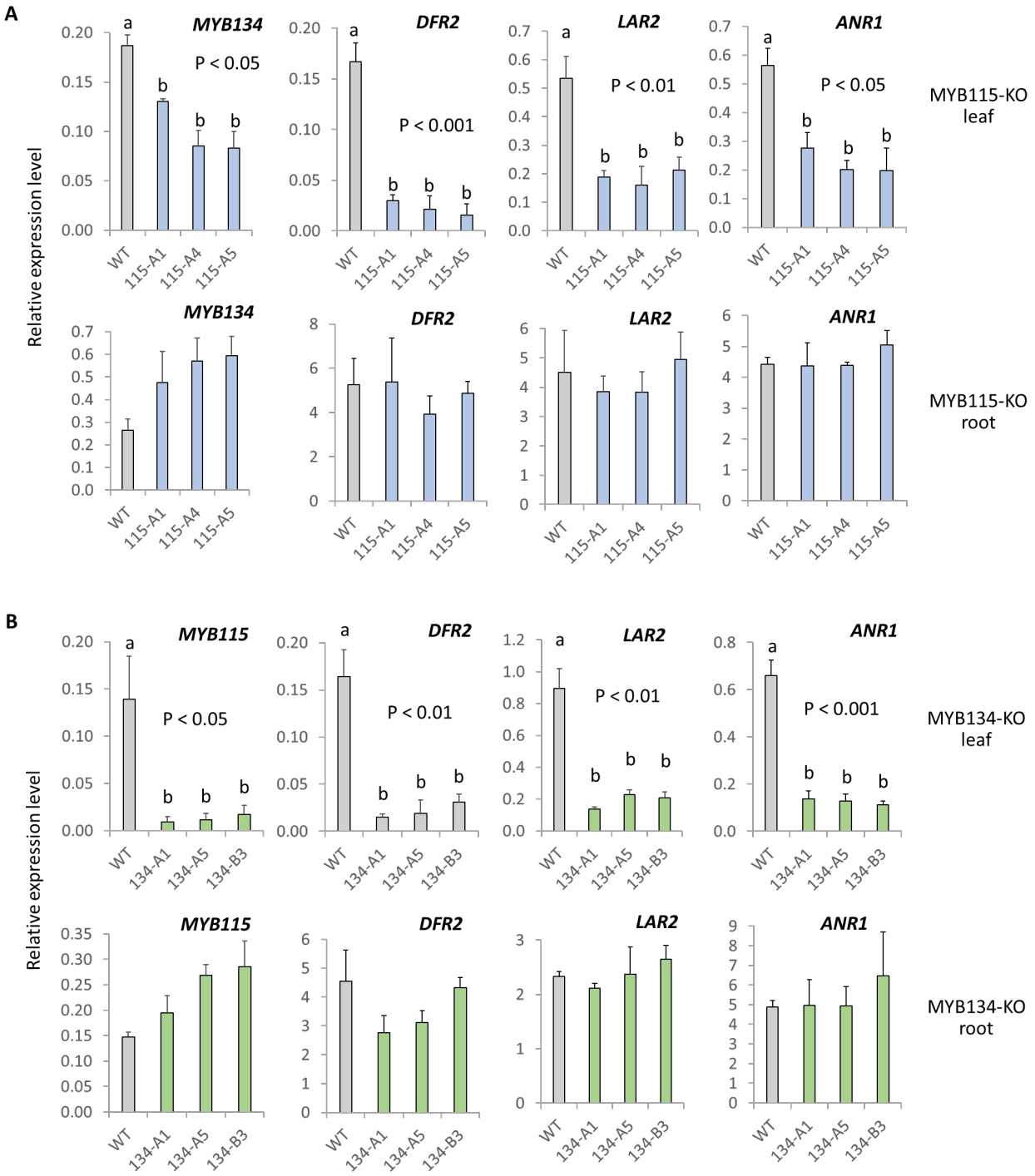
To confirm the effect of the CRISPR construct on PA biosynthesis genes, three knock-out lines for each construct were selected for flavonoid biosynthetic gene expression analysis by RT-qPCR. DFR (dihydroflavonol reductase) is a late flavonoid biosynthetic gene, which transforms dihydroflavonols into leucoanthocyanidins. LAR (leucoanthocyanidin reductase) and ANR (anthocyanin reductase) are two PA-specific enzymes. ANR converts anthocyanidins to epicatechin, the major flavan-3-ol component of PAs, while LAR converts leucoanthocyanidins

to catechin, a secondary flavan-3-ol precursor for PAs. As regulators of PA synthesis in poplar leaves, MYB134 was tested in MYB115-KO lines and MYB134 was quantified in MYB134-KO lines, respectively.

For MYB115-KO lines, qPCR analysis showed significant reduction in DFR2, ANR1 and LAR2 compared to wild-types in leaves (Figure 3-13A), which is consistent with the reduced PA content in MYB115-KO leaves (Figure 3-10A). By contrast, no difference was observed for any flavonoid genes in root tissues. MYB134 gene expression was also checked in MYB115-KO lines. MYB134 gene was downregulated 25-50% for three MYB115-KO lines in leaves ($p < 0.05$), while upregulated in roots though not significantly.

For MYB134-KO lines, DFR2, ANR1 and LAR2 were significantly downregulated in leaves (Figure 3-13B). Stronger effect of reduction for flavonoid genes expression was found in leaves for MYB134-KO lines compared to MYB115-KO lines, which match the observation that there was more reduction in PA content for MYB134-KO lines compared to that for MYB115-KO lines in leaves (Figure 3-10A, B). Again, no dramatic difference was observed for flavonoid genes in roots. MYB115 was downregulated in MYB134-KO leaves ($p < 0.05$) but was upregulated in roots though the effect was not significant (Figure 3-13B). The results of flavonoid genes expression in MYB115-KO and MYB134-KO roots are expected and match our observations that PA content showed no change in root tissues for these lines (Figure 3-10A, B). The upregulated MYB134 and MYB115 expression in MYB115-KO and MYB134-KO roots, respectively, (though not significantly) may explain the unchanged flavonoid gene expression and flavonoid concentration.

For double-KO lines, a reduction in DFR2 and ANR1 expression was observed in leaves (Figure 3-13C). The expression of LAR2 and LAR3 was also down-regulated, however, no statistical significance was observed for LAR3 due to its low expression level in leaf tissues (Figure 3-13C). The expression level of LAR3 and ANR1 correlated with the quantification of catechin and epicatechin respectively in double-KO leaves (Figure 3-11A, B). In roots, LAR2 and LAR3 were downregulated ($p < 0.05$), while DFR2 and ANR1 expression was not affected (Figure 3-13C), which match the reduced catechin and unchanged epicatechin in root tissues for double-KO lines (Figure 3-11A,B). The results of flavonoid gene expression in leaves and roots were consistent with reduced PA content in both tissues for double-KO lines (Figure 3-10C).


A
Relative expression level
MYB134
P < 0.05
DFR2
P < 0.001
LAR2
P < 0.01
ANR1
P < 0.05
MYB115-KO leaf
MYB134
DFR2
LAR2
ANR1
MYB115-KO root
WT
115-A1
115-A4
115-A5
B
Relative expression level
MYB115
P < 0.05
DFR2
P < 0.01
LAR2
P < 0.01
ANR1
P < 0.001
MYB134-KO leaf
MYB115
DFR2
LAR2
ANR1
MYB134-KO root
WT
134-A1
134-A5
134-B3

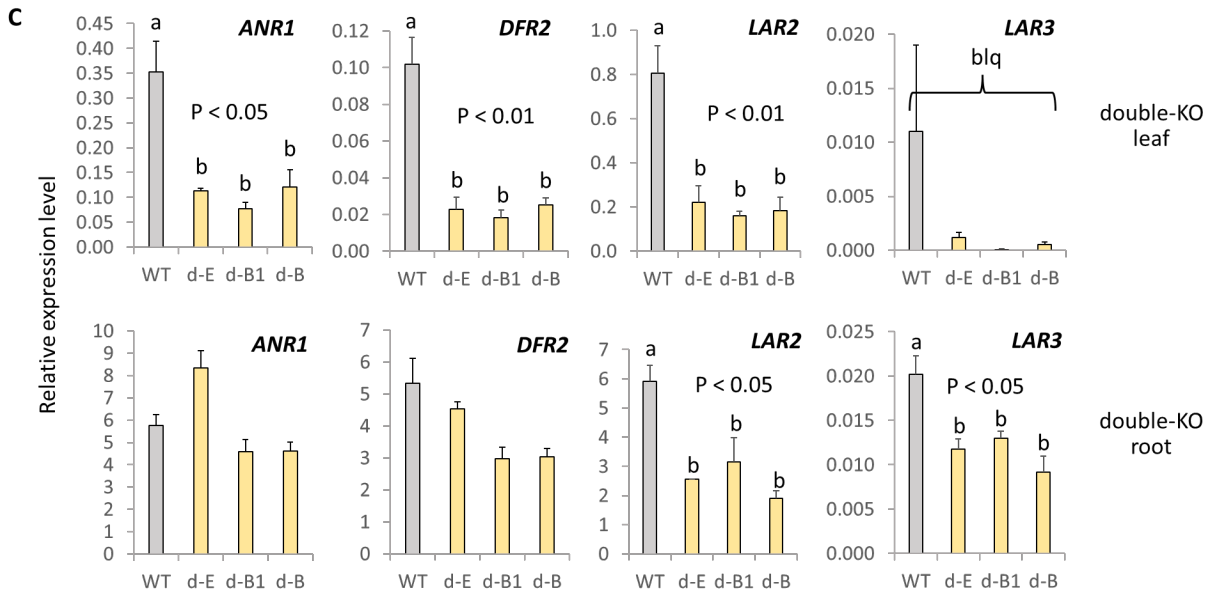



**Figure 3-13: Expression analysis of MYB115-KO, MYB134-KO and double-KO plants treated by high-light.** Relative expression of key flavonoid genes and MYB genes in MYB115-KO (A), MYB134-KO (B) and double-KO (C) leaves and roots as quantified by qPCR. After exposing plants to high light stress for 20 days, tissue was harvested and RNA was extracted and prepared for qPCR as described in the Methods. Transcript levels were determined using qPCR and were normalized against housekeeping genes (EF1-β and actin) as outlined in the Methods. Error bars indicate SE, at least three biological replicates per line were analyzed. Data were analyzed by one-way ANOVA followed by Tukey's HSD post hoc test. Different letters on the bar indicate statistically significant differences of PA content according to Tukey's HSD test at 95% confidence (P-value < 0.05). Colors of bars represent different lines, wild-type plants are shown in grey, MYB115-KO lines are shown in blue, MYB134-KO lines are shown in green, double-KO lines are shown in yellow. blq, below limit quantification.

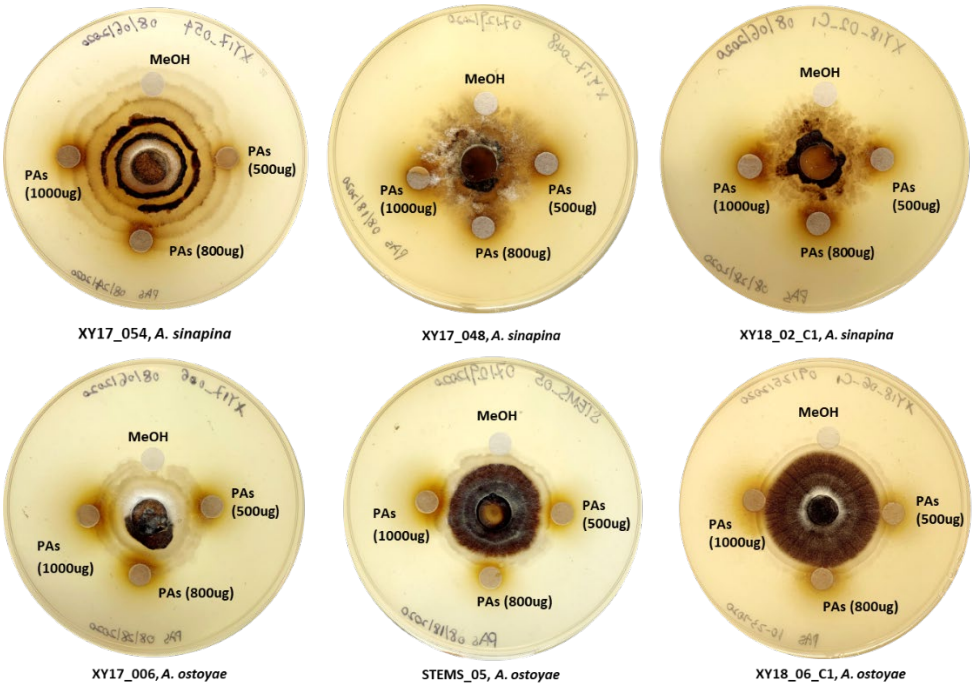
## 3.5 Disc diffusion assay for testing inhibitory effects of purified PAs and catechin on *Armillaria* sp. and *Laccaria bicolor*

To test the toxicity of PAs and catechin on fungi *in vitro*, hyphal inhibition by PAs and catechin was tested through disc diffusion assays using the pathogens *Armillaria* (*sinapina & ostoyae*) and the ectomycorrhizal fungus *Laccaria bicolor*. Hyphal inhibition was assessed by measuring the zone of inhibition surrounding each disc with the compounds at different concentrations after an incubation of 48h. *Armillaria* growth was inhibited by the positive control hygromycin as shown by arrows in Figure 3-14 B. However, no hyphal inhibition was observed for any of the *Armillaria* isolates tested with PAs/catechin concentration between 50 to 400 µg/disc after a 48h period compared to the negative control 100% methanol (Figure 3-

14A,B). The concentration was increased to 500, 800 and 1000 µg/disc and procedure was repeated. Again, no obvious inhibition was observed (Figure 3-14A,B).

Slight inhibition was observed with disc assays in *L. bicolor*. Mycelial growth was only slightly affected when discs were treated with low PA concentrations at 50 and 100 µg/disc but was strongly inhibited at 200 and 400 µg/disc (Figure 3-15A). The positive control hygromycin exerted a slight inhibition towards the *L. bicolor* fungus colony compared to the negative control 100% methanol (Figure 3-15A), but the inhibition was not as strong as in *Armillaria*. Catechin disc assays in *L. bicolor* showed strong pigment around the discs, but no inhibition effect was observed (Figure 3-15B). The inhibition zone was measured as shown in Figure 3-15C.

A. PA Disc Assays




B. Catechin Disc Assays

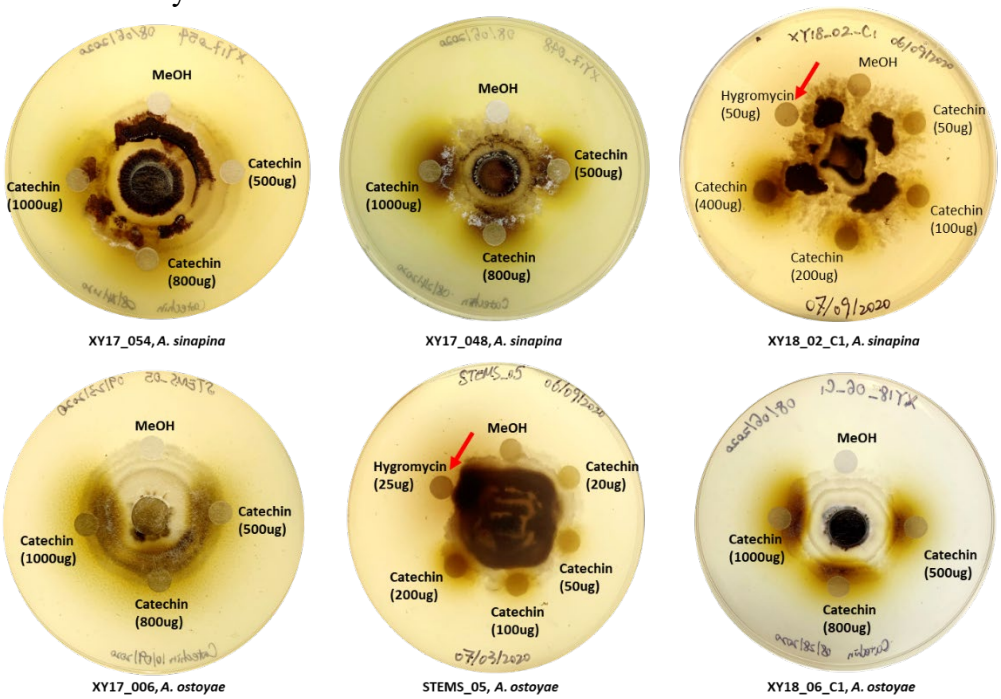



**Figure 3-14: Disc diffusion assay to evaluate hyphal growth inhibition.** Colonies of *A. sinapina* and *A. ostoyae* treated by purified PAs discs (A) and catechin discs (B). Each plate contained a disc with methanol only control. Six different isolates collected from different locations in BC were tested, including three *A. sinapina* isolates (XY17_054, XY17_048, XY18_02_C1) and three *A. ostoyae* isolates (XY17_006, STEMS_05, XY18_06_C1).

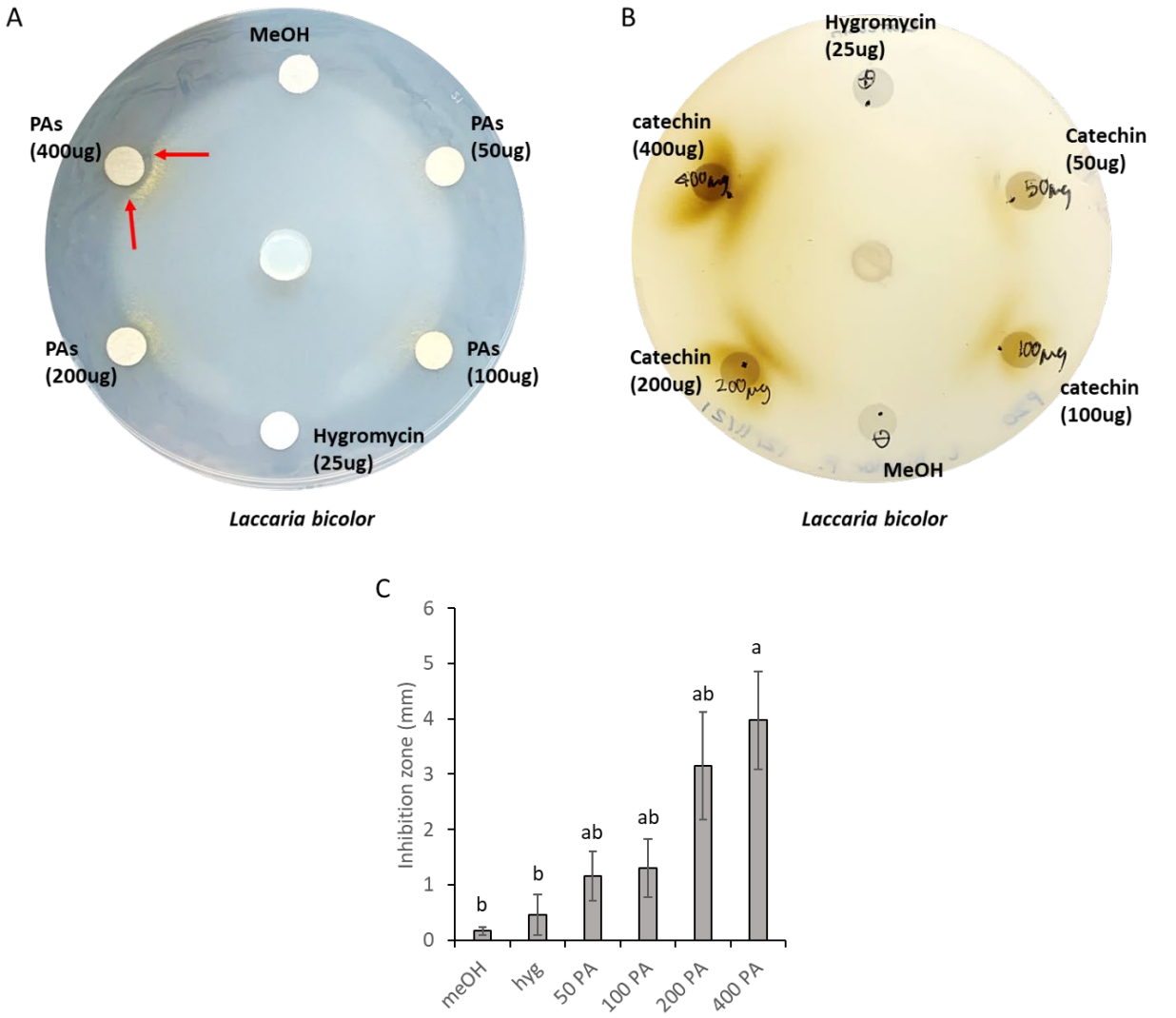



**Figure 3-15: Disc diffusion assay to evaluate hyphal growth inhibition of *L. bicolor* by purified PAs.** (A) PA disc assay, (B) catechin disc assay, (C) inhibition zone measurement. Each plate contained a disc with 25 ug hygromycin and methanol only control. Red arrows point out the edge of *L. bicolor* mycelium. The bars represent means ± SE of inhibition zone radius in three repeated PA disc assay. Data were analyzed by one-way ANOVA followed by Tukey's HSD post hoc test. Different letters on the bar indicate statistically significant differences of PA content according to Tukey's HSD test at 95% confidence (P-value < 0.05).

### 3.6 Bioassays of roots with fungi via sandwich co-culture system between *in vitro* poplar roots and mycorrhizae *L. bicolor*.

In collaboration with Simon Petley, Constabel Lab Honours student, I tested the co-culture method by Felten *et al.* (2009) in order to explore interactions of poplar roots with mycorrhizae *L. bicolor*. I first adapted the method to pathogens *Armillaria* and attempted to

quantify infection with the inoculated poplar roots. However, the staining method for *Armillaria* failed to detect any hyphae, due to unmatched fluorescent stereoscope laser or filter. I also tested the specificity of qPCR primers for housekeeper genes in both poplar and *Armillaria* genome. Genomic DNA was extracted from *Armillaria* mycelium and poplar roots respectively using DNeasy Plant Kit and I succeeded to amplify *Armillaria* housekeeping gene *EF1α*, which was specific to *Armillaria* and could not be amplified from poplar genomic DNA (Supplementary Figure 7.1). A poplar *actin* gene was also tested and was not found in products amplified from *Armillaria* genomic DNA (Supplementary Figure 7.1). This result confirmed the specificity of these two genes as well as the qPCR primers.

The sandwich inoculation between poplar hairy roots and mycorrhiza was successfully set up and ectomycorrhizal formation was observed under the stereo microscope. Infected root tips were distinct from uninfected ones in morphology, where infected root tips appeared to be round and brownish and were easily distinguished from normal roots under the microscope (Figure 3-16). To quantify ectomycorrhizal formation, mycorrhizal root tips and total root tips were counted under the stereo microscope. The ratio of infected roots per total root tips was further calculated and the ratio of double-KO inoculated roots was compared to the wild-type inoculated roots (Figure 3-17).

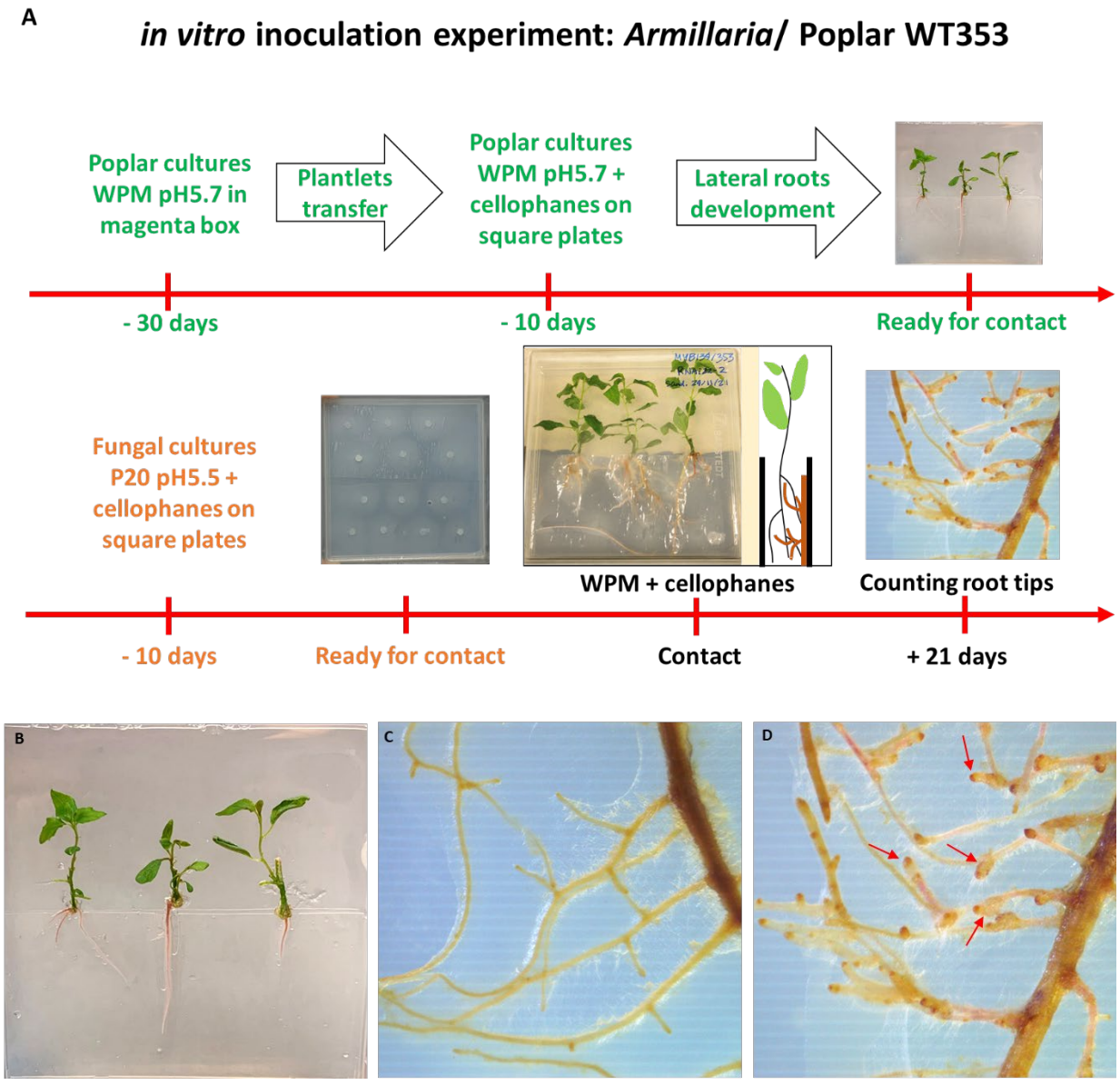



**Figure 3-16: *In vitro* model system for poplar–*L. bicolor* ectomycorrhizal symbiosis.** (A) The *in vitro* co-culture system for ectomycorrhizal symbiosis between poplar plantlets and *L. bicolor* (B) non-inoculated poplar roots, (C) uninfected inoculated poplar roots, (D) infected inoculated poplar roots with ectomycorrhizal formation. Red arrows point out the ectomycorrhizal root tips.

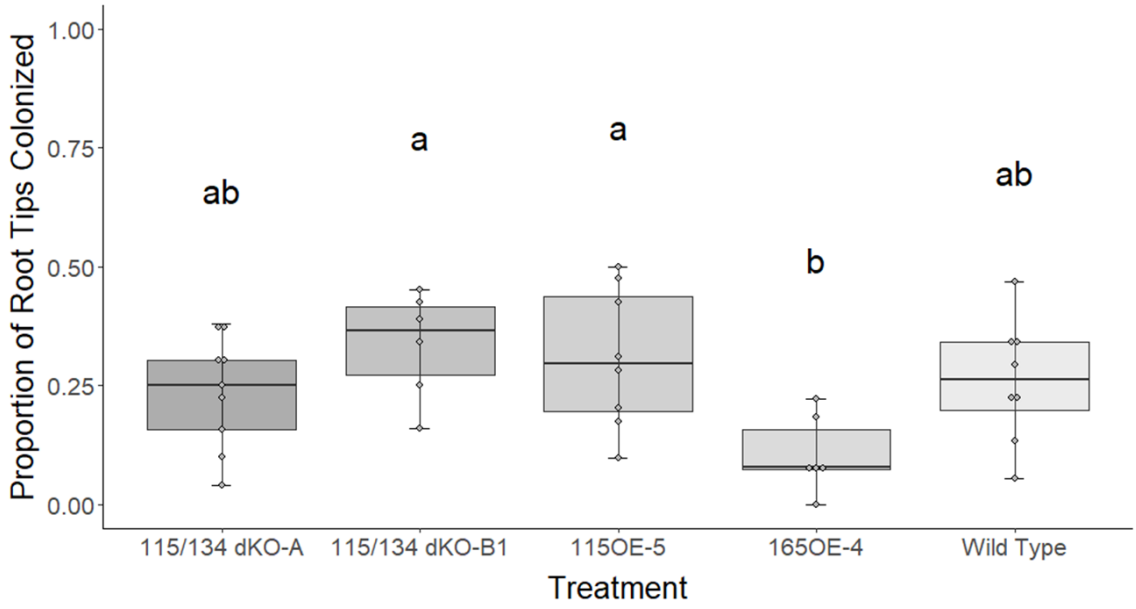



**Figure 3-17: Ratio of colonized root tips to total root tips in poplar transgenic lines and wild-types.** After inoculation with *Laccaria bicolor* for three weeks, the root tips of each plantlet were visually identified as colonized (mycorrhizal) or non-mycorrhizal. Different letters on the box represent statistically significant differences (P-value < 0.05) between the mean ratio of colonized to total root tips among all lines as determined by one-way ANOVA and followed by post-hoc Tukey's test. The number of biological replicates for each line are displayed as closed points (n>4 for each).

# Chapter 4: Discussion

## 4.0 Summary of key results

This study explored the regulation of MYB115 and MYB134 in PA biosynthesis by creating CRISPR knockouts. Genotyping analysis of CRISPR knockouts showed high editing efficiency. Different mutation types were found in knock-out lines, including monoallelic, biallelic and chimeric mutations. Small indels were detected as the most common mutations and large deletions between two target sites were also found in several lines. No obvious phenotypes were found in knock-out hairy roots. In whole plant knockouts, CRISPR disruption of MYB115 and MYB134 caused the greatest reduction of PA in leaves. In contrast, only double-KO whole plants showed a reduction of PA in roots. Single-KO whole plants displayed no difference in PA content in root tissues. The same trend was found with catechin content. All the single-KO lines synthesized low levels of epicatechin, but only double-KO and MYB134 KO lines showed reduced epicatechin in leaves. Interestingly, salicinoid content was altered but only in double-KO lines, with higher concentrations in roots and the opposite trend in leaves. QPCR analysis of key flavonoid gene expression showed consistent results, and key flavonoid genes were down-regulated in leaves but not changed in roots. In root tissues, only LAR2 and LAR3 were down-regulated in double-KO whole plants.

## 4.1 MYB115 and MYB134 sequences in *P. tremula* × *alba* 717-1B4 and *P. tremula* × *tremuloides* 353-38 genotype show slight differences compared to genome databases

Because of the frequent occurrence of SNPs in the highly heterozygous poplar genome, the sequences of MYB115 and MYB134 genes were confirmed and aligned to the sequences obtained from Popgenie and AspenDB genome databases. Sequences from *P. tremula* × *alba* 717-1B4 background matched with AspenDB sequence data (Figure 3-2). Sequences from *P. tremula* × *tremuloides 353-38* background mostly matched with Popgenie, but several SNPs did not match, and new SNPs were found within exon regions of the genes (Figure 3-1). A study of poplar CRISPR knockouts assessed the SNP interference of gRNA specificity and found that the CRISPR/Cas9 system is highly sensitive to SNPs (Zhou et al., 2015). Zhou and colleagues designed three gRNAs for the 4CL gene, two of which were without SNPs, and one contained

one SNP in each allele. This study showed that the cleavage for the third 4CL gene was abolished, with no editing events detected due to the single SNPs in the target sequence. To achieve efficient genome editing, gRNA design should ensure the absence of SNPs. Hence, updated MYB115 and MYB134 sequences were used to select SNP-free pre-designed gRNAs as target sites for CRISPR Cas9.

### 4.2 Editing efficiency of CRISPR Cas9 in poplar

For both MYB115 and MYB134 genes, I obtained high percentages of edited hairy roots as well as edited whole plants. Editing efficiency (all mutation types including monoallelic, biallelic and chimeric mutation types) for knock-out hairy roots was between 82.4-92.3%, while the editing ratio for whole plant mutants was 75 - 100%. Despite the high editing efficiency, only 30.8-35.3% of all transgenic events led to biallelic mutations in hairy root lines. A few mutants were monoallelic and some were chimeric. Some lines were mutated with one large deletion, while no homozygous mutations were found. As expected from the CRISPR/Cas9 system, we found not only short deletions and insertions, but also very large deletions between two gRNAs, though it was difficult to get chromosomal fragment deletions in both alleles. An advantage of the CRISPR/Cas9 system is the capacity of Cas9 to create large deletions between two target cut sites, which was also demonstrated in petunia transgenic plants and eucalyptus hairy roots (Zhang et al., 2016; Dai et al., 2020).

No homozygous mutations were detected in any transgenic lines, which was unexpected especially when compared to other studies in poplar. For example, a study by Fan et al. (2015) applied CRISPR in *Populus tomentosa* Carr. and targeted genomic sites for the phytoene desaturase gene 8 (PtoPDS). Although they only obtained a total mutation rate of 51.7%, 93.2% of the edited plants were homozygous mutants, and only 6.7% were heterozygous mutants (Fan et al., 2015). A study by Elorriaga et al. (2018) used the CRISPR/Cas9 system to alter the flowering related gene PLFY with two sgRNAs in poplar 717-1B4 and 353-53 and found a high knockout rate (with both alleles altered) in both clones (77.0% and 69.7% respectively). This study also found that small indel mutations were prevalent among mutated alleles of events when only one sgRNA was present, while large deletions were prevalent among alleles targeted by two sgRNAs (Elorriaga et al., 2018). Homozygous mutations were found to be around 10-20% of

total events with both alleles altered (Elorriaga et al., 2018). Similar editing rates were found in rice, where 54.9% biallelic, 24.7% homozygous, 5.8% heterozygous (monoallelic) mutations and only one chimeric mutation were detected (Ma et al., 2015). Another paper tested 11 target genes in two rice subspecies, and found that the most abundant T0 genotype was chimera (40.4%), as well as 10.5% homozygotes (Zhang et al., 2014). This research further genotyped the next generation (T1) and found no detectable new mutations or reversions (Zhang et al., 2014). These studies suggest different editing efficiency and variable mutation types among species and different genes. Our poplar knockouts showed high editing efficiency and high ratio of biallelic mutants, which is comparable to other studies. However, the complete lack of homozygous mutation is uncommon in other experiments.

Based on the differences observed between the MYB115- and the MYB134-edited mutants, the editing efficiency of CRISPR/Cas9 system appeared to be both gene- and gRNA-dependent. This might be due to factors including cell type and phase of the cell-division cycle (branching/proliferation of hairy roots). This might also indicate that the Cas9 capacity may be different in these two backgrounds. Therefore, we also compared the efficiency of Cas9 in two backgrounds of hybrid poplars and hairy roots generated in hybrid 7171-B4 which seemed to have more chimeras than whole plants generated in hybrid 353-38. The study by Elorriaga et al. (2018) obtained high knockout rate in poplar 717-1B4, suggesting that different hybrid backgrounds may not be the major factor for editing efficiency. The hairy root system may affect the editing efficiency, however. The low ratio of biallelic mutations and the high occurrence of chimeras in poplar hairy roots were unexpected when compared to hairy roots system in other species such as soybean, tomato or chicory, where high frequencies of biallelic mutations and low number of chimeras were reported in general (Ron et al., 2014; Cai et al., 2015; Jacobs et al., 2015; Sun et al., 2015; Bernard et al., 2019). The yields for double-KO lines were different in the whole plant and hairy root system. Though editing efficiency was high, no hairy root lines were mutated in both MYB genes. It was also difficult to obtain double-KO whole plants with both genes mutated and biallelic mutations for each gene. The low yields suggest that knocking out both MYB genes may have an influence on plant development which results in low vitality.

In addition, more indels were identified within exon2 of MYB115 gene and exon3 of MYB134 gene than that in exon1 of either MYB gene. These results suggests that the gene editing efficiency may be also influenced by the properties of the two paired gRNAs as well as

the location of the target sites within the coding sequences. A recent paper revealed that the structure of gRNA, including GC content, purine residues in the gRNA end, and the free accessibility of its seed region seemed to be relevant to genome editing efficiency, based on studies of nine different poplar genes (Bruegmann et al., 2019).

### 4.3 Biochemical and gene expression analysis of MYB115 and MYB134 single knock-out as well as double knock-out plants reveals that the transcriptional regulation of PA synthesis may differ in leaf and root tissues

The biochemical and genetic phenotypes of knockouts differed in leaves and roots, suggesting that MYB115 and MYB134 are important regulators of PA pathway in both tissues but with some differences. In leaves, single-KO lines and double-KO lines all showed a reduction in PA content, but a stronger effect was observed in double-KO and MYB134-KO lines compared to MYB115-KO lines (Figure 3-10). The stronger effect was also supported by leaves qPCR data in MYB134-KOs compared to MYB115-KOs (Figure 3-13A, B). This indicates that MYB134 is more important than MYB115 for regulating PA biosynthesis in leaves. In roots, double-KO lines showed a reduction in PA content while no change was observed in PA content of single-KO lines (Figure 3-10). This result demonstrates that MYB115 and MYB134 work together in regulating PA pathway in roots. The PA concentration was only reduced by one-half in roots and leaves of double-KOs. The remaining PA suggests the existence of other PA regulators.

The trends in catechin accumulation showed a reduction in all knock-out leaves and a reduction only in double-KO roots (Figure 3-11A), which is consistent with the PA data (Figure 3-10). QPCR analysis showed significantly down-regulated DFR2 in leaves for all knock-out lines. As DFR is a late flavonoid biosynthetic gene, this result matches the reduction of PA content in leaves. The LAR and ANR expression data on qPCR (Figure 3-13) mostly correlated with differences in catechin and epicatechin quantification. LAR and ANR are required for producing the PA precursors catechin and epicatechin, respectively. QPCR quantification showed significant reduction in ANR1 expression level in leaves for all knockouts (Figure 3-13). However, only double-KO lines and two MYB134-KO lines contained significantly reduced epicatechin in leaves (Figure 3-11B). MYB115-KO lines showed no change in epicatechin concentration which is not consistent with the down-regulated ANR1 expression (Figure 3-13A).

In parallel, all knock-out lines showed dramatically reduced catechin concentration in leaves (Figure 3-11A), which did not match the unchanged LAR3 gene expression level in MYB115-KO and MYB134-KO leaves (data not shown). The expression of LAR3 was down-regulated in double-KO leaves though the effect was not statistically significant. Due to the low expression level of LAR3 in leaf tissues, LAR3 might not be the major enzyme for catechin synthesis. Therefore, I also checked LAR2 expression. LAR2 was down-regulated in leaves for all knock-out lines (Figure 3-13), which can explain the reduction of catechin (Figure 3-11A). Catechin content was also reduced in double-KO root tissues, which matches the significant decrease of LAR2 and LAR3 expression in roots (Figure 3-13C). No significant difference in catechin concentration was found in roots of MYB115-KO and MYB134-KO plants (Figure 3-11A). This is consistent with the lack of difference in qPCR analysis for LAR genes (Figure 3-13). Likewise, no significant changes in epicatechin concentrations were found in roots of any knockouts (Figure 3-11B). This corresponds to the unaffected ANR1 gene expression level in roots (Figure 3-13).

Interestingly, the salicinoid pathway was affected in double-KO lines, since salicinoids increased in roots and decreased in leaves. These opposite trends indicate that the regulatory pathways by MYB115 and MYB134 are quite different for above- and under-ground tissues. Reduced salicinoid content was also observed in MYB115 and MYB134 overexpressors leaves (James et al., 2017), which might be due to the induction of repressor MYBs, as the overexpressors of repressor MYB165- and MYB194 show a strong reduction in salicinoid content in leaves (Ma et al., 2018). The increased salicinoids in double-KO roots might be due to the decreased repressors, while the decreased salicinoids in double-KO leaves were unexpected. Hence, the expression levels of repressor MYBs should be quantified in knock-out lines in future experiments to help explain the change in salicinoids.

## 4.4 The functional interaction of MYB115 and MYB134 in regulating PA biosynthesis

Our previous study by James et al. (2017) found that both MYB115 and MYB134 activated their own and each other's promoters, demonstrated by *in vivo* transient expression experiments. Microarray data and transcriptome data also suggested positive and negative

feedback loops in the regulation of PA biosynthesis (James et al., 2017; Ma et al., 2018). In overexpressing lines of MYB115 and MYB134, induction of MBW cofactors and MYB repressors occurred simultaneously. To determine to what extent the functions of MYB115 and MYB134 overlap, MYB134-RNAi lines and CRISPR mutants were established (Gourlay et al., 2020). Surprisingly, MYB134-RNAi lines showed a significant downward change for MYB115 expression in leaves (Gourlay et al., 2020). Similarly, the MYB134-KO lines showed dramatically decreased MYB115 expression in leaves (0.07-0.12 fold, $p < 0.05$). These results suggest that MYB115 was partly regulated by MYB134, as predicted by our model (James et al., 2017). By contrast, MYB115 was more strongly expressed in MYB134-KO roots, though not significantly, perhaps due to a compensation mechanism for PA regulation. A similar result was seen in MYB115-KOs, where MYB134 was down-regulated in leaves (0.4-0.7 fold, $p < 0.05$), while up-regulated in roots (though not significantly). Therefore, when knocking out one MYB gene, the other MYB can be induced in roots but down-regulated in leaves. These opposite behaviour in the expression of the other MYB gene in MYB-KOs revealed the complex relationship between MYB115 and MYB134 genes, and the relationship differs in leaf and root tissues.

Leaf PA content was generally more affected than root PA in all the knock-out lines. Additionally, young roots accumulate more PAs than old roots (Westley, 2015), and roots accumulate more PAs than shoots (Figure 3-10). This suggest that in roots, PA might be controlled by additional developmental genes in addition to the stress responsive MYB115 and MYB134. The repressor MYBs could also be part of this developmental regulation in roots where one KO led to the upregulation of the other MYB. To test if repressor MYBs are involved in root PA regulation during development is a direction for future experiments. MYB115 was found to be expressed at a very low level in poplar leaf and root tissues, compared to MYB134 (James et al., 2017). Thus, other transcription factors may also be important in root PA regulation. For example, a recent study by Li et al. (2022) characterized a CCCH type transcription factor, PuC3H35, which positively regulates the expression of genes involved in PA and lignin biosynthesis as a response to drought stress in the roots of *Populus ussuriensis*. Although PuC3H35 is also expressed in leaf veins and stems, it is mostly expressed in roots as the drought stress affects root the most and starts from roots. Therefore, drought stress

responsive PuC3H35 might be one of the regulators for root PA synthesis and interact with MYBs as well.

The relative expression level of MYB115 in roots was at the same level compared to that in leaves, while the relative expression level of MYB134 in roots is much lower than in leaves (James et al., 2017). MYB134 was expressed at higher levels and showed more effects on PA levels in leaves when being knocked out, while MYB115 was expressed at low levels in wild-type and has fewer influences on PA levels in leaves when being knocked out. Therefore, MYB134 may be the primary driver of PA synthesis in leaves. When it comes to roots, the other MYB can be induced as compensation when knocking out one MYB gene, perhaps explaining the lack of change of root PA. However, when knocking out both MYBs, root PA was affected, though the effect was small compared to leaves. This observation indicates that these two MYBs have redundant functions in leaves and roots, and they may be equally important for PA synthesis in root tissues. In either tissue of double-KO plants, there are still lots of PAs, suggesting there must be other transcription factors. A PA regulatory pathway network was described in a tentative model (Figure 4-1). In summary, PAs are synthesized at different concentrations in leaves and roots, and they are regulated differently.

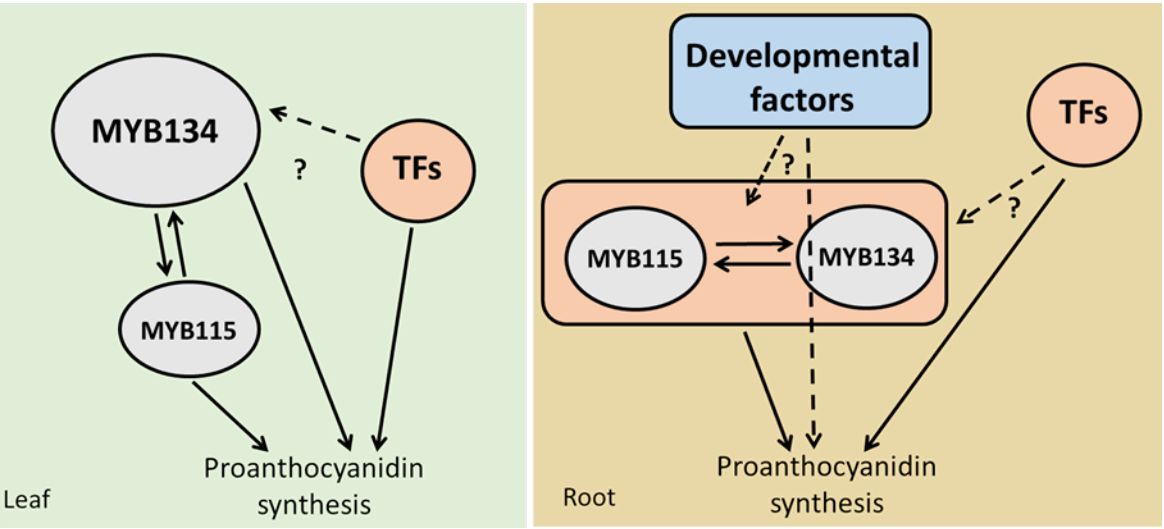



**Figure 4-1: A tentative model for PA regulatory network for leaf and root.**

Some models have been supposed that the effects of MYBs are synergistic rather than additive, such as the transcriptional network of AtTT2/MYB5/MYB23 regulating the biosynthesis of mucilage, tannin, jasmonic acid and wax (cuticle) via multiple tiers of regulatory

hierarchy in *Arabidopsis* (Li et al., 2020). Other examples include the model of VvMYBPA1, VvMYB5a and VvMYB5b regulating PA biosynthesis in developing grape berries (Deluc et al., 2008), and the mysterious interaction among MtMYBPAR, MtMYB5 and MtMYB14 in *M. trunculata* (Liu et al., 2014). In summary, the relationship and interaction between TT2-type (MYB134) and PA1-type (MYB115) MYB activators still needs more work to integrate.

### 4.5 Phenotypes of double knock-out plants suggest phytohormone profiles or sugar availability change

The physical phenotypes of double-KO lines were distinct from wild-type plants, while single-KO lines showed no differences. Double-KO lines had a number of branches sprouting from the middle of the stem and more severe necrosis was found on double-KO bottom leaves. To control branch number, the growing shoot apical meristem inhibits the growth of axillary buds further down the stem by producing auxin. This phenomenon is known as apical dominance (Phillips, 1975). If the apical meristem is removed (decapitated) or auxin is removed from the terminal bud, the dormant axillary buds become activated and can develop into branches. Shoot branching is controlled by complex interactions among hormones, nutrients, and environmental cues. Plant hormones including auxin, strigolactones, and cytokinins (CKs) are involved in the regulation of bud outgrowth (Ferguson and Beveridge, 2009). Auxin upregulates strigolactone-biosynthesis genes and represses CK-biosynthesis genes in the stem, mediating the branching inhibition (Tanaka et al., 2006; Dun et al., 2012; Saeed et al., 2017). There is evidence that flavonoids (such as kaempferol and quercetin) can modulate auxin transport, which is tissue specific occurring at the root and shoot apices (Peer et al., 2004; Peer and Murphy, 2007). Absence of flavonoids can enhance auxin transport from shoot apex to root tip and excess flavonoids can inhibit polar auxin transport (Peer et al., 2004). This mechanism may explain the branch phenotypes in double-KO plants with reduced flavonoids. Sugars also play an important role in the induction of shoot branching in several species, including *Rosa* sp., leafy spurge (*Euphorbia esula* L.) and pea (*Pisum sativum* L.) (Girault et al., 2010; Rabot et al., 2012; Barbier et al., 2015; Chao et al., 2016; Barbier et al., 2019; Bertheloot et al., 2020). Plants response to decapitation by releasing signaling regulators, such as sugar, that moves from the leaf to the lateral bud (Mason et al., 2014; Fichtner et al., 2017). A study on potato (*Solanum tuberosum* L.) indicate that sucrose promotes bud outgrowth through cytokinin (CK) -induced vacuolar

invertase, an enzyme that contributes to sugar sink strength (Salam et al., 2021). Hence, auxin and sugar availability can control stem branching. The phenotypes of more branching in our double-KO lines indicate a profile change in phytohormone or sugar mobilization in shoots which needs further evidence. This change might be due to the reduction in PA pathway, and the extra carbon flowing into sugar transport, metabolism, and signaling.

We observed more severe necrosis in double-KO plants which is consistent with previous studies. High-PA transgenics showed lower rates of leaf necrosis as well as a delay in necrotic development compared with wild-types during drought stress (Gourlay et al., 2022). By contrast, MYB134-RNAi plants showed more rapid necrosis compared to controls during drought stress (Gourlay et al., 2022). The severe necrosis might be due to the negative oxidative effects of abiotic stresses, including high-light stress. This process was aggravated by the reduction of PA in double-KO leaves, as PAs serve as cellular antioxidants in poplar leaves (Gourlay et al., 2022).

### 4.6 PAs effect on *Armillaria* and *L. bicolor in vitro* and *in vivo*

The disc assay on two types of fungi *in vitro* showed that high concentration of PAs can inhibit fungal-*L.bicolor* growth but have no effect on the pathogen fungus *Armillaria*. This opposite effect suggests the specificity of PAs antimicrobial function towards fungus species. However, the concentration of PAs on the disc may not be comparable to root PA concentration as a result of the complicated distribution of PAs in root zones. The antifungal function of PAs in roots needs more evidence from co-culture inoculation assay *in vivo*. However, a tissue culture assay still lacks a real soil environment with rhizosphere microbial diversity, which might reveal the actual function of PAs in root. So far, no significant changes in mycorrhizal formation among wild-types, OE and KO lines were observed in preliminary co-culture sandwich assay between poplar roots and mycorrhiza *L. bicolor*.

PAs serve as not only antimicrobial agent but may also be a carbon source for *L. bicolor* during beneficial interaction (unpublished data from Dr. Felten's Lab, Umeå Plant Science Center). Mycorrhiza can induce polyphenyloxidase (PPO) in plants leading to tannins degradation. For example, tomato plants inoculated with arbuscular mycorrhizal fungi (AMF) and plant growth promoting rhizobacteria (PGPR) induced potent activity of phenolics and defensive enzymes including PPO (1.35 fold) (Sharma and Sharma, 2017). Another study on

inoculation of chickpea with AMF significantly increased PPO activities compared with non-inoculated chickpea (Sohrabi et al., 2012). It seems that the fungus can use the existing tannins (probably from the epidermal layer of young roots) as a carbon source, as mycorrhiza are thought to remove PAs and degrade pectin to penetrate between the apoplastic space during early stages of ECM colonization. Transcriptome data also revealed that PA genes in mycorrhizal roots are downregulated when compared to uncolonized roots (unpublished data from Dr. Judith Felten Lab, Umea University, Sweden).

# Chapter 5: Overall Conclusions and Future Directions

This study confirmed the sequences of MYB115 and MYB134 genes and found new SNPs compared to genome databases. MYB115 and MYB134 genes were knocked out in poplar using CRISPR/Cas9 technique to explore the regulation and function of PA biosynthesis, particularly in roots. High editing efficiency was seen in CRISPR knockouts and mutation types varied among mutants. No obvious phenotypes were observed in knock-out hairy roots. MYB134 was more likely the primary driver of PA synthesis in leaves. When knocking out one MYB gene, the other MYB was induced as a compensation in roots. CRISPR disruption of MYB115 and MYB134 together did lead to significant reduction of PA in leaves, however, root PA was less affected, suggesting the redundant functions of two MYBs in leaves and roots. The opposite trends in salicinoid content of double-KOs as well as the opposite trends of the other MYB gene expression in single-KO leaves and roots indicate that the regulatory pathways by MYB115 and MYB134 are quite different for above- and under-ground tissues. The remaining PAs in either tissue of double-KOs revealed the present of other regulators as well as the complex relationship between MYB115 and MYB134 genes in different tissues. Double-KO plants were the only lines showing phenotypes (more branches and necroses), suggesting a connection between PAs and development.

This study successfully applied CRISPR Cas9 system to generate targeted mutagenesis in the genome of poplar and demonstrated high editing efficiency of this system in poplar based on genotyping results. Though biallelic mutation rates could be improved, the hairy root system is an efficient system to quickly screen the knockouts and estimate the editing efficiency. Generating stable whole plant mutants is also a powerful tool to functionally characterize genes for research. The knockout plants of MYB115 and MYB134, both regulators of the PA pathway in poplar, help explain the regulation of PA biosynthesis in poplar leaves and roots. This study further consolidates the regulatory network of flavonoids and PAs and adds more information to the positive and negative interactions among transcription factors based on phenotypes of MYB overexpressors and RNAi lines. This project also provides preliminary insight into the functions of PAs in roots, organs which accumulate large quantities of PAs. Though PAs function in roots has only been studied superficially, the *in vitro* assays pave a way to testing their role in underground interactions by using the knock-out lines created here.

Although my work has shown that the regulatory pathways by MYB115 and MYB134 are quite different for above- and under-ground tissues, there are still a few questions left to be answered. Future research could try different environment stress instead of high-light treatment to induce PA accumulation in leaves. Experimental design should also contain untreated lines to check the baseline of PA content in shoots and should include more replicates for each line. Transcriptomes in different tissues would be necessary to check the expression level of important genes, including MYB repressors, sugar availability relevant genes, phytohormone genes as well as developmental relevant genes. Other transcription factors would also be worth investigating as there might be a redundancy for the regulation of PA biosynthesis, in particular in poplar roots. As the salicinoid pathway was also affected in double-KOs, analysis of salicinoid biosynthetic genes should be considered as well. The species *L. bicolor* would be a good choice for this work as a recognized model species for the study of host-plant interactions (Felten et al., 2009). Hence, inoculation of poplar roots by fungus in the soil environment would be illuminating. Additionally, the quantification of root infection by staining or qPCR would be more accurate and would explain the interaction between poplar roots and rhizosphere fungus.

# References

**Aggarwal PR, Nag P, Choudhary P, Chakraborty N, Chakraborty S** (2018) Genotype-independent *Agrobacterium rhizogenes*-mediated root transformation of chickpea: a rapid and efficient method for reverse genetics studies. Plant Methods **14**: 13

**Anderson RC, Vodovnik M, Min BR, Pinchak WE, Krueger NA, Harvey RB, Nisbet DJ** (2012) Bactericidal effect of hydrolysable and condensed tannin extracts on *Campylobacter jejuni in vitro*. Folia Microbiol (Praha) **57**: 253–258

**Aron PM, Kennedy JA** (2008) Flavan-3-ols: Nature, occurrence and biological activity. Mol Nutr Food Res **52**: 79–104

**Bai QX, Duan BB, Ma JC, Fen YN, Sun SJ, Long QM, Lv JJ, Wan DS** (2020) Coexpression of PalbHLH1 and PalMYB90 genes from *Populus alba* enhances pathogen resistance in poplar by increasing the flavonoid content. Front Plant Sci **10**: 14

**Bai S, Tao R, Yin L, Ni J, Yang Q, Yan X, Yang F, Guo X, Li H, Teng Y** (2019) Two B-box proteins, PpBBX18 and PpBBX21, antagonistically regulate anthocyanin biosynthesis via competitive association with *Pyrus pyrifolia* ELONGATED HYPOCOTYL 5 in the peel of pear fruit. Plant J **100**: 1208–1223

**Bailey JK, Deckert R, Schweitzer JA, Rehill BJ, Lindroth RL, Gehring C, Whitham TG** (2005) Host plant genetics affect hidden ecological players: links among *Populus*, condensed tannins, and fungal endophyte infection. Can J Bot **83**: 356–361

**Barbehenn R, Cheek S, Gasperut A, Lister E, 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 V., Jaros A, Lee G, Mozola C, Weir Q, Salminen JP** (2009) Hydrolyzable tannins as "quantitative defenses": Limited impact against *Lymantria dispar* caterpillars on hybrid poplar. J Insect Physiol **55**: 297–304

**Barbehenn R V., Jones CP, Hagerman AE, Karonen M, Salminen JP** (2006) Ellagitannins have greater oxidative activities than condensed tannins and galloyl glucoses at high pH: Potential impact on caterpillars. J Chem Ecol **32**: 2253–2267

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

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

**Barbier F, Péron T, Lecerf M, Perez-Garcia MD, Barrière Q, Rolčík J, Boutet-Mercey S, Citerne S, Lemoine R, Porcheron B, et al** (2015) Sucrose is an early modulator of the key hormonal mechanisms controlling bud outgrowth in *Rosa hybrida*. J Exp Bot **66**: 2569–2582

**Barbier FF, Dun EA, Kerr SC, Chabikwa TG, Beveridge CA** (2019) An update on the signals controlling shoot branching. Trends Plant Sci **24**: 220–236

**Bardon C, Piola F, Haichar F el Z, Meiffren G, Comte G, Missery B, Balby M, Poly F**

(2016) Identification of B-type procyanidins in *Fallopia* spp. involved in biological denitrification inhibition. Environ Microbiol **18**: 644–655

**Bernard G, Gagneul D, dos Santos HA, Etienne A, Hilbert JL, Rambaud C** (2019) Efficient genome editing using CRISPR/Cas9 technology in chicory. Int J Mol Sci **20**: 1155

**Bernays EA, Chamberlain DJ, McCarthy P** (1980) The differential effects of ingested tannic acid on different species of Adridoidea. Entomol Exp Appl **28**: 158–166

**Bertheloot J, Barbier F, Boudon F, Perez-Garcia MD, Péron T, Citerne S, Dun E, Beveridge C, Godin C, Sakr S** (2020) Sugar availability suppresses the auxin-induced strigolactone pathway to promote bud outgrowth. New Phytol **225**: 866–879

**Beyeler M, Heyser W** (1997) The influence of mycorrhizal colonization on growth in the greenhouse and on catechin, epicatechin and procyanidin in roots of *Fagus sylvatica* L. Mycorrhiza **7**: 171–177

**Boeckler GA, Towns M, Unsicker SB, Mellway RD, Yip L, Hilke I, Gershenzon J, Constabel CP** (2014) Transgenic upregulation of the condensed tannin pathway in poplar leads to a dramatic shift in leaf palatability for two tree-feeding Lepidoptera. J Chem Ecol **40**: 150–158

**Bogs J, Jaffe FW, Takos AM, Walker AR, Robinson SP** (2007) The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. Plant Physiol **143**: 1347–1361

**Brahem M, Bornard I, Renard CMGC, Le Bourvellec C** (2020) Multiscale localization of procyanidins in ripe and overripe perry pears by light and transmission electron microscopy. J Agric Food Chem **68**: 8900–8906

**Bruegmann T, Deecke K, Fladung M** (2019) Evaluating the efficiency of gRNAs in CRISPR/Cas9 mediated genome editing in poplars. Int J Mol Sci **20**: 3623

**Burda S, Oleszek W** (2001) Antioxidant and antiradical activities of flavonoids. J Agric Food Chem **49**: 2774–2779

**Cai Y, Chen L, Liu X, Sun S, Wu C, Jiang B, Han T, Hou W** (2015) CRISPR/Cas9-mediated genome editing in soybean hairy roots. PLoS One **10**: e0136064

**Chao WS, Dořramaci M, Horvath DP, Anderson J V., Foley ME** (2016) Phytohormone balance and stress-related cellular responses are involved in the transition from bud to shoot growth in leafy spurge. BMC Plant Biol **16**: 1–21

**Chin L, Leung DWM, Harry Taylor H** (2009) Lead chelation to immobilised *Symphytum officinale* L. (comfrey) root tannins. Chemosphere **76**: 711–715

**Close DC, McArthur C** (2002) Rethinking the role of many plant phenolics - protection from photodamage not herbivores? Oikos **99**: 166–172

**Constabel CP, Yoshida K, Walker V** (2014) Diverse ecological roles of plant tannins: plant defense and beyond. Recent Adv Polyphen Res **4**: 115–142

**Dai Y, Hu G, Dupas A, Medina L, Blandels N, Clemente HS, Ladouce N, Badawi M, Hernandez-Raquet G, Mounet F, et al** (2020) Implementing the CRISPR/CAS9 technology in eucalyptus hairy roots using wood-related genes. Int J Mol Sci **21**: 3408

**Debeaujon I, Peeters AJM, Léon-Kloosterziel KM, Koornneef M** (2001) The TRANSPARENT TESTA12 gene of arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. Plant Cell **13**: 853–871

**DeGabriel JL, Moore BD, Foley WJ, Johnson CN** (2009) The effects of plant defensive chemistry on nutrient availability predict reproductive success in a mammal. Ecology **90**: 711–719

**Deluc L, Bogs J, Walker AR, Ferrier T, Decendit A, Merillon JM, Robinson SP, Barrieu F** (2008) The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. Plant Physiol **147**: 2041–2053

**Deveau A, Palin B, Delaruelle C, Peter M, Kohler A, Pierrat JC, Sarniguet A, Garbaye J, Martin F, Frey-Klett P** (2007) The mycorrhiza helper Pseudomonas fluorescens BBc6R8 has a specific priming effect on the growth, morphology and gene expression of the ectomycorrhizal fungus *Laccaria bicolor* S238N. New Phytol **175**: 743–755

**Dixon RA, Xie DY, Sharma SB** (2005) Proanthocyanidins - a final frontier in flavonoid research? New Phytol **165**: 9–28

**Dun EA, Germain A de Saint, Rameau C, Beveridge CA** (2012) Antagonistic action of strigolactone and cytokinin in bud outgrowth control. Plant Physiol **158**: 487–498

**Elorriaga E, Klocko AL, Ma C, Strauss SH** (2018) Variation in mutation spectra among CRISPR/Cas9 mutagenized poplars. Front Plant Sci **9**: 594

**Engels C, Knödler M, Zhao YY, Carle R, Gänzle MG, Schieber A** (2009) Antimicrobial activity of gallotannins isolated from mango (*Mangifera indica* L.) kernels. J Agric Food Chem **57**: 7712–7718

**Fan D, Liu T, Li C, Jiao B, Li S, Hou Y, Luo K** (2015) Efficient CRISPR/Cas9-mediated targeted mutagenesis in populus in the first generation. Sci Rep **5**: 1–7

**Fellenberg C, Corea O, Yan LH, Archinuk F, Piirtola EM, Gordon H, Reichelt M, Brandt W, Wulff J, Ehlting J, et al** (2020) Discovery of salicyl benzoate UDP-glycosyltransferase, a central enzyme in poplar salicinoid phenolic glycoside biosynthesis. Plant J **102**: 99–115

**Felten J, Kohler A, Morin E, Bhalerao RP, Palme K, Martin F, Ditengou FA, Legué V** (2009) The ectomycorrhizal fungus *Laccaria bicolor* stimulates lateral root formation in poplar and *Arabidopsis* through auxin transport and signaling. Plant Physiol **151**: 1991–2005

**Ferguson BJ, Beveridge CA** (2009) Roles for auxin, cytokinin, and strigolactone in regulating shoot branching. Plant Physiol **149**: 1929–1944

**Feucht W, Schmid PPS, Christ E** (1986) Distribution of flavanols in meristematic and mature tissues of *Prunus-avium* shoots. J Plant Physiol **125**: 1–8

**Fichtner F, Barbier FF, Feil R, Watanabe M, Annunziata MG, Chabikwa TG, Höfgen R, Stitt M, Beveridge CA, Lunn JE** (2017) Trehalose 6-phosphate is involved in triggering axillary bud outgrowth in garden pea (*Pisum sativum* L.). Plant J **92**: 611–623

**Foss SR, Nakamura C V., Ueda-Nakamura T, Cortez DAG, Endo EH, Dias Filho BP** (2014) Antifungal activity of pomegranate peel extract and isolated compound punicalagin against dermatophytes. Ann Clin Microbiol Antimicrob **13**: 1–6

**Fu C, Loo AEK, Chia FPP, Huang D** (2007) Oligomeric proanthocyanidins from mangosteen pericarps. J Agric Food Chem **55**: 7689–7694

**Furuno K, Akasako T, Sugihara N** (2002) The contribution of the pyrogallol moiety to the superoxide radical scavenging activity of flavonoids. Biol Pharm Bull **25**: 19–23

**Gesell A, Yoshida K, Tran LT, Constabel CP** (2014) Characterization of an apple TT2-type R2R3 MYB transcription factor functionally similar to the poplar proanthocyanidin regulator PtMYB134. Planta **240**: 497–511

**Girault T, Abidi F, Sigogne M, Pelleschi-Travier S, Boumaza R, Sakr S, Leduc N** (2010) Sugars are under light control during bud burst in Rosa sp. Plant Cell Environ **33**: 1339–1350

**Giuliani C, Tani C, Maleci Bini L, Fico G, Colombo R, Martinelli T** (2018) Localization of phenolic compounds in the fruits of *Silybum marianum* characterized by different silymarin chemotype and altered colour. Fitoterapia **130**: 210–218

**Gourlay G, Constabel CP** (2019) Condensed tannins are inducible antioxidants and protect hybrid poplar against oxidative stress. Tree Physiol **39**: 345–355

**Gourlay G, Hawkins BJ, Albert A, Schnitzler JP, Peter Constabel C** (2022) Condensed tannins as antioxidants that protect poplar against oxidative stress from drought and UV-B. Plant Cell Environ **45**: 362–377

**Gourlay G, Ma D, Schmidt A, Constabel CP** (2020) MYB134-RNAi poplar plants show reduced tannin synthesis in leaves but not roots, and increased susceptibility to oxidative stress. J Exp Bot **71**: 6601–6611

**Hagerman AE, Butler LG** (1991) Chapter 10 - Tannins and lignins. In: Rosenthal, G.A. and Berenbaum, M.R., Eds., Herbivores: their interactions with secondary plant metabolites, Vol. I: The Chemical participants, Academic Press, New York, 355-388.

**Hagerman AE, Riedl KM, Jones GA, Sovik KN, Ritchard NT, Hartzfeld PW, Riechel TL** (1998) High molecular weight plant polyphenolics (tannins) as biological antioxidants. J Agric Food Chem **46**: 1887–1892

**Hammouda H, Alvarado C, Bouchet B, Kalthoum-Chérif J, Trabelsi-Ayadi M, Guyot S** (2014) Tissue and cellular localization of tannins in tunisian dates (*Phoenix dactylifera* L.) by light and transmission electron microscopy. J Agric Food Chem **62**: 6650–6654

**Hancock KR, Collette V, Fraser K, Greig M, Xue H, Richardson K, Jones C, Rasmussen S** (2012) Expression of the R2R3-MYB transcription factor TaMYB14 from *Trifolium arvense* activates proanthocyanidin biosynthesis in the legumes *Trifolium repens* and *Medicago sativa*. Plant Physiol **159**: 1204–1220

**Hatano T, Kira R, Yoshizaki M, Okuda T** (1986) Seasonal changes in the tannins of *Liquidambar formosana* reflecting their biogenesis. Phytochemistry **25**: 2787–2789

**Hervás G, Pérez V, Giráldez FJ, Mantecón AR, Almar MM, Frutos P** (2003) Intoxication of

sheep with quebracho tannin extract. J Comp Pathol **129**: 44–54

**Hoffmann T, Friedlhuber R, Steinhauser C, Tittel I, Skowranek K, Schwab W, Fischer TC** (2012) Histochemical screening, metabolite profiling and expression analysis reveal Rosaceae roots as the site of flavan-3-ol biosynthesis. Plant Biol **14**: 33–40

**Hunter MD** (2001) Out of sight, out of mind: the impacts of root-feeding insects in natural and managed systems. Agric For Entomol **3**: 3–9

**Inoue KH, Hagerman AE** (1988) Determination of gallotannin with rhodanine. Anal Biochem **169**: 363–369

**Jacobs TB, LaFayette PR, Schmitz RJ, Parrott WA** (2015) Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol **15**: 1–10

**James AM, Ma DW, Mellway R, Gesell A, Yoshida K, Walker V, Tran L, Stewart D, Reichelt M, Suvanto J, et al** (2017) Poplar MYB115 and MYB134 transcription factors regulate proanthocyanidin synthesis and structure. Plant Physiol **174**: 154–171

**Jones DL, Prabowo AM, Kochian L V** (1996) Kinetics of malate transport and decomposition in acid soils and isolated bacterial populations: The effect of microorganisms on root exudation of malate under Al stress. Plant Soil **182**: 239–247

**Jourdes M, Pouységu L, Deffieux D, Teissedre PL, Quideau S** (2013) Hydrolyzable tannins: gallotannins and ellagitannins. In: Ramawat K., Mérillon JM. (eds) Natural Products. Springer, Berlin, Heidelberg. pp 1975–2010

**Kao YY, Harding SA, Tsai CJ** (2002) Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen. Plant Physiol **130**: 796–807

**Kaplan I, Halitschke R, Kessler A, Sardanelli S, Denno RF** (2008) Constitutive and induced defenses to herbivory in above- and belowground plant tissues. Ecology **89**: 392–406

**Karamać M** (2009) Chelation of Cu(II), Zn(II), and Fe(II) by tannin constituents of selected edible nuts. Int J Mol Sci **10**: 5485–5497

**Karowe DN** (1989) Differential effect of tannic acid on two tree-feeding Lepidoptera: implications for theories of plant anti-herbivore chemistry. Oecologia **80**: 507–512

**Kimura M, Wada H** (1989) Tannins in mangrove tree roots and their role in the root environment. Soil Sci Plant Nutr **35**: 101–108

**Kochian L V, Hoekenga OA, Pineros MA** (2004) How do crop plants tolerate acid soils? - Mechanisms of aluminum tolerance and phosphorous efficiency. Annu Rev Plant Biol **55**: 459–493

**Kraus TEC, Dahlgren RA, Zasoski RJ** (2003) Tannins in nutrient dynamics of forest ecosystems - a review. Plant Soil **256**: 41–66

**Laitinen ML, Julkunen-Tiitto R, Rousi M** (2002) Foliar phenolic composition of European white birch during bud unfolding and leaf development. Physiol Plant **114**: 450–460

**Laplaze L, Gherbi H, Frutz T, Pawlowski K, Franche C, Macheix JJ, Auguy F, Bogusz D, Duhoux E** (1999) Flavan-containing cells delimit Frankia-infected compartments in *Casuarina glauca* nodules. Plant Physiol **121**: 113–122

**Lattanzio V, Terzano R, Cicco N, Cardinali A, Di Venere D, Linsalata V** (2005) Seed coat tannins and bruchid resistance in stored cowpea seeds. J Sci Food Agric **85**: 839–846

**Lavola A** (1998) Accumulation of flavonoids and related compounds in birch induced by UV-B irradiance. Tree Physiol **18**: 53–58

**Lees GL, Gruber MY, Suttill NH** (1995a) Condensed tannins in sainfoin. II. Occurrence and changes during leaf development. Can J Bot **73**: 1540–1547

**Lees GL, Suttill NH, Gruber MY** (1993) Condesed tannins in sainfoin .1. A histological and cytological survey of plant tissues. Can J Bot Can Bot **71**: 1147–1152

**Lees GL, Wall KM, Beveridge TH, Suttill NH** (1995b) Localization of condensed tannins in apple fruit peel, pulp, and seeds. Can J Bot **73**: 1897–1904

**Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M** (2006) Genetics and biochemistry of seed flavonoids. Annu. Rev. Plant Biol. Annual Reviews, Palo Alto, pp 405–430

**Li C, Leverence R, Trombley JD, Xu S, Yang J, Tian Y, Reed JD, Hagerman AE** (2010) High molecular weight persimmon (*Diospyros kaki* L.) proanthocyanidin: A highly galloylated, a-linked tannin with an unusual flavonol terminal unit, myricetin. J Agric Food Chem **58**: 9033–9042

**Li C, Pei J, Yan X, Cui X, Tsuruta M, Liu Y, Lian C** (2021) A poplar B-box protein PtrBBX23 modulates the accumulation of anthocyanins and proanthocyanidins in response to high light. Plant Cell Environ **44**: 3015–3033

**Li D, Yang J, Pak S, Zeng M, Sun J, Yu S, He Y, Li C** (2022) PuC3H35 confers drought tolerance by enhancing lignin and proanthocyanidin biosynthesis in the roots of *Populus ussuriensis*. New Phytol **233**: 390–408

**Li SF, Allen PJ, Napoli RS, Browne RG, Pham H, Parish RW** (2020) MYB-bHLH-TTG1 regulates *Arabidopsis* seed coat biosynthesis pathways directly and indirectly via multiple tiers of transcription factors. Plant Cell Physiol **61**: 1005–1018

**Li YG, Tanner G, Larkin P** (1996) The DMACA-HCl protocol and the threshold proanthocyanidin content for bloat safety in forage legumes. J Sci Food Agric **70**: 89–101

**Liu CG, Jun JH, Dixon RA** (2014) MYB5 and MYB14 play pivotal roles in seed coat polymer biosynthesis in *Medicago truncatula*. Plant Physiol **165**: 1424–1439

**Liu JY, Osbourn A, Ma PD** (2015) MYB transcription factors as regulators of phenylpropanoid metabolism in plants. Mol Plant **8**: 689–708

**Liu Y, Gao L, Xia T, Zhao L** (2009) Investigation of the site-specific accumulation of catechins in the tea plant (*Camellia sinensis* (L.) O. Kuntze) via vanillin HCI staining. J Agric Food Chem **57**: 10371–10376

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

**Ma DW, Constabel CP** (2019) MYB repressors as regulators of phenylpropanoid metabolism in plants. Trends Plant Sci **24**: 275–289

**Ma DW, Reichelt M, Yoshida K, Gershenzon J, Constabel CP** (2018) Two R2R3-MYB

proteins are broad repressors of flavonoid and phenylpropanoid metabolism in poplar. Plant J **96**: 949–965

**Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y, et al** (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Mol Plant **8**: 1274–1284

**Mace ME, Howell CR** (1974) Histochemistry and identification of condensed tannin precursors in roots of cotton seedlings. Can J Bot **52**: 2423-

**Major IT, Constabel CP** (2006) Molecular analysis of poplar defense against herbivory: comparison of wound- and insect elicitor-induced gene expression. New Phytol **172**: 617–635

**Marsh KJ, Foley WJ, Cowling A, Wallis IR** (2003) Differential susceptibility to Eucalyptus secondary compounds explains feeding by the common ringtail (*Pseudocheirus peregrinus*) and common brushtail possum (*Trichosurus vulpecula*). J Comp Physiol B Biochem Syst Environ Physiol **173**: 69–78

**Martius S, Hammer KA, Locher C** (2012) Chemical characteristics and antimicrobial effects of some *Eucalyptus kinos*. J Ethnopharmacol **144**: 293–299

**Mason MG, Ross JJ, Babst BA, Wienclaw BN, Beveridge CA** (2014) Sugar demand, not auxin, is the initial regulator of apical dominance. Proc Natl Acad Sci U S A **111**: 6092–6097

**McGivern BB, Tfaily MM, Borton MA, Kosina SM, Daly RA, Nicora CD, Purvine SO, Wong AR, Lipton MS, Hoyt DW, et al** (2021) Decrypting bacterial polyphenol metabolism in an anoxic wetland soil. Nat Commun **12**: 2466

**McKenzie BA, Peterson CA** (1995a) Root browning in *Pinus banksiana* Lamb. and *Eucalyptus pilularis* Sm. 1. Anatomy and permeability of the white and tannin zones. Bot Acta **108**: 127–137

**McKenzie BA, Peterson CA** (1995b) Root browning in *Pinus banksiana* Lamb. and *Eucalyptus pilularis* Sm. 1. Anatomy and permeability of the cork zone. Bot Acta **108**: 138–143

**McMahon LR, McAllister TA, Berg BP, Majak W, Acharya SN, Popp JD, Coulman BE, Wang Y, Cheng KJ** (2000) A review of the effects of forage condensed tannins on ruminal fermentation and bloat in grazing cattle. Can J Plant Sci **80**: 469–485

**McSweeney CS, Palmer B, McNeill DM, Krause DO** (2001) Microbial interactions with tannins: nutritional consequences for ruminants. Anim Feed Sci Technol **91**: 83–93

**Mellway RD, Tran LT, Prouse MB, Campbell MM, Peter Constabel C** (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

**Mila I, Scalbert A, Expert D** (1996) Iron withholding by plant polyphenols and resistance to pathogens and rots. Phytochemistry **42**: 1551–1555

**Mira L, Fernandez MT, Santos M, Rocha R, Florencio MH, Jennings KR** (2002) Interactions of flavonoids with iron and copper ions: A mechanism for their antioxidant

activity. Free Radic Res **36**: 1199–1208

**Miranda M, Ralph SG, Mellway R, White R, Heath MC, Bohlmann J, Constabel CP** (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

**Moilanen J, Karonen M, Tähtinen P, Jacquet R, Quideau S, Salminen JP** (2016) Biological activity of ellagitannins: effects as anti-oxidants, pro-oxidants and metal chelators. Phytochemistry **125**: 65–72

**Moilanen J, Salminen JP** (2008) Ecologically neglected tannins and their biologically relevant activity: chemical structures of plant ellagitannins reveal their in vitro oxidative activity at high pH. Chemoecology **18**: 73–83

**Mueller-Harvey I** (2006) Unravelling the conundrum of tannins in animal nutrition and health. J Sci Food Agric **86**: 2010–2037

**Munzenberger B, Heilemann J, Strack D, Kottke I, Oberwinkler F** (1990) Phenolics of mycorrhizas and non-mycorrhizal roots of Norway spruce. Planta **182**: 142–148

**Muoki RC, Paul A, Kumari A, Singh K, 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

**Nesi N, Jond C, Debeaujon I, Caboche M, 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

**Niemi K, Julkunen-Tiitto R, Haggman H, Sarjala T** (2007) *Suillus variegatus* causes significant changes in the content of individual polyamines and flavonoids in Scots pine seedlings during mycorrhiza formation in vitro. J Exp Bot **58**: 391–401

**Okuda T, Ito H** (2011) Tannins of constant structure in medicinal and food plants-hydrolyzable tannins and polyphenols related to tannins. Molecules **16**: 2191–2217

**Okuda T, Yoshida T, Hatano T, Ito H** (2009) Ellagitannins renewed the concept of tannins. Chem. Biol. Ellagitannins An Underestimated Cl. Bioact. Plant Polyphenols. pp 1–54

**Osawa H, Endo I, Hara Y, Matsushima Y, Tange T** (2011) Transient proliferation of proanthocyanidin-accumulating cells on the epidermal apex contributes to highly aluminum-resistant root elongation in camphor tree. Plant Physiol **155**: 433–446

**Osier TL, Lindroth RL** (2006) Genotype and environment determine allocation to and costs of resistance in quaking aspen. Oecologia **148**: 293–303

**Osier TL, Lindroth RL** (2001) Effects of genotype, nutrient availability, and defoliation on aspen phytochemistry and insect performance. J Chem Ecol **27**: 1289–1313

**Peer WA, Bandyopadhyay A, Blakeslee JJ, Makam SN, Chen RJ, Masson PH, Murphy AS** (2004) Variation in expression and protein localization of the PIN Family of auxin efflux facilitator proteins in flavonoid mutants with altered auxin transport in *Arabidopsis thaliana*. Plant Cell **16**: 1898–1911

**Peer WA, Murphy AS** (2007) Flavonoids and auxin transport: modulators or regulators? Trends

Plant Sci **12**: 556–563

**Peng S, Scalbert A, Monties B** (1991) Insoluble ellagitannins in *Castanea sativa* and *Quercus petraea* woods. Phytochemistry **30**: 775–778

**Perkovich C, Ward D** (2020) Protein:carbohydrate ratios in the diet of gypsy moth *Lymantria dispar* affect its ability to tolerate tannins. J Chem Ecol **46**: 299–307

**Peters DJ, Constabel CP** (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

**Peterson CA, Enstone DE, Taylor JH** (1999) Pine root structure and its potential significance for root function. Plant Soil **217**: 205–213

**Phillips IDJ** (1975) Apical dominance. Annu Rev Plant Physiol **26**: 341–367

**Porter LJ** (1992) Structure and chemical properties of the condensed tannins. *In* RW Hemingway, PE Laks, eds, Plant Polyphenols Synth. Prop. Significance. Springer US, Boston, MA, pp 245–258

**Porter LJ, Hrstich LN, Chan BG** (1986) The conversion of procyanidins and prodelphinidins to cyanidin and delphinidin. Phytochemistry **25**: 223–230

**Pourcel L, Routaboul JM, Kerhoas L, Caboche M, Lepiniec L, Debeaujon I** (2005) TRANSPARENT TESTA10 encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in *Arabidopsis* seed coat. Plant Cell **17**: 2966–2980

**Puupponen-Pimiä R, Nohynek L, Hartmann-Schmidlin S, Kähkönen M, Heinonen M, Määttä-Riihinen K, Oksman-Caldentey KM** (2005) Berry phenolics selectively inhibit the growth of intestinal pathogens. J Appl Microbiol **98**: 991–1000

**Quideau S, Deffieux D, Douat-Casassus C, Pouysegu L** (2011) Plant polyphenols: chemical properties, biological activities, and synthesis. Angew Chemie-International Ed **50**: 586–621

**Rabot A, Henry C, Baaziz K Ben, Mortreau E, Azri W, Lothier J, Hamama L, Boummaza R, Leduc N, Pelleschi-Travier S, et al** (2012) Insight into the role of sugars in bud burst under light in the rose. Plant Cell Physiol **53**: 1068–1082

**Rasmann S, Agrawal AA** (2008) In defense of roots: a research agenda for studying plant resistance to belowground herbivory. **146**: 875–880

**Ron M, Kajala K, Pauluzzi G, Wang D, Reynoso MA, Zumstein K, Garcha J, Winte S, Masson H, Inagaki S, et al** (2014) Hairy root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene expression and function using tomato as a model. Plant Physiol **166**: 455–469

**Saeed W, Naseem S, Ali Z** (2017) Strigolactones biosynthesis and their role in abiotic stress resilience in plants: A critical review. Front Plant Sci **8**: 1487

**Salam BB, Barbier F, Danieli R, Teper-Bamnolker P, Ziv C, Spíchal L, Aruchamy K, Shnaider Y, Leibman D, Shaya F, et al** (2021) Sucrose promotes stem branching through cytokinin. Plant Physiol **185**: 1708–1721

**Scalbert A** (1991) Antimicrobial properties of tannins. Phytochemistry **30**: 3875–3883

**Schimel JP, Cates RG, Ruess R** (1998) The role of balsam poplar secondary chemicals in controlling soil nutrient dynamics through succession in the alaskan taiga. Biogeochemistry **42**: 221–234

**Schutzendubel A, Polle A** (2002) Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization. J Exp Bot **53**: 1351–1365

**Schweitzer JA, Madritch MD, Bailey JK, LeRoy CJ, Fischer DG, Rehill BJ, Lindroth RL, Hagerman AE, Wooley SC, Hart SC, 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 AN, Schmidt MA, Moore MA, Lindroth RL, Wooley SC, Hagerman AE** (2011) Qualitative variation in proanthocyanidin composition of *Populus* species and hybrids: genetics is the key. J Chem Ecol **37**: 57–70

**Seyoum A, Asres K, El-Fiky FK** (2006) Structure–radical scavenging activity relationships of flavonoids. Phytochemistry **67**: 2058–2070

**Sharma IP, Sharma AK** (2017) Physiological and biochemical changes in tomato cultivar PT-3 with dual inoculation of mycorrhiza and PGPR against root-knot nematode. Symbiosis **71**: 175–183

**Smith AH, Zoetendal E, Mackie RI** (2005) Bacterial mechanisms to overcome inhibitory effects of dietary tannins. Microb Ecol **50**: 197–205

**Smith S, Read D** (2008) Mycorrhizal Symbiosis (Third Edition). Mycorrhizal Symbiosis. Academic Press, pp. 191-385.

**Sohrabi Y, Heidari G, Weisany W, Golezani KG, Mohammadi K** (2012) Changes of antioxidative enzymes, lipid peroxidation and chlorophyll content in chickpea types colonized by different Glomus species under drought stress. Symbiosis **56**: 5–18

**Solaiman ZM, Senoo K** (2018) Arbuscular mycorrhizal fungus causes increased condensed tannins concentrations in shoots but decreased in roots of *Lotus japonicus* L. Rhizosphere **5**: 32–37

**Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG** (1995) Chloroplast gene sequence data suggest a single origin of the predispostion for symbiotic nitrogen-fixation in angiosperms. Proc Natl Acad Sci U S A **92**: 2647–2651

**Stevens CE, Hume ID** (2004) Comparative physiology of the vertebrate digestive system. Cambridge Univ Press pp.11-23.

**Sukovata L, Jaworski T, Karolewski P, Kolk A** (2015) The performance of Melolontha grubs on the roots of various plant species. Turkish J Agric For **39**: 107–116

**Sun X, Hu Z, Chen R, Jiang Q, Song G, Zhang H, Xi Y** (2015) Targeted mutagenesis in soybean using the CRISPR-Cas9 system. Sci Reports 2015 51 **5**: 1–10

**Sutherland ORW, Hutchins RFN, Greenfield WJ** (1982) Effect of lucerne saponins and Lotus condensed tannins on survival of grass grub, *Costelytra zealandica*. New Zeal J Zool **9**: 511–514

**Svistoonoff S, Hocher V, Gherbi H** (2014) Actinorhizal root nodule symbioses: What is

signalling telling on the origins of nodulation? Curr Opin Plant Biol **20**: 11–18

**Tahara K, Hashida K, Otsuka Y, Ohara S, Kojima K, Shinohara K** (2014) Identification of a hydrolyzable tannin, Oenothein B, as an aluminum-detoxifying ligand in a highly aluminum-resistant tree, *Eucalyptus camaldulensis*. Plant Physiol **164**: 683–693

**Tam PCF, Griffiths DA** (1993) Mycorrhizal associations in Hong Kong Fagaceae - V. The role of polyphenols. Mycorrhiza **3**: 165–170

**Tanaka M, Takei K, Kojima M, Sakakibara H, Mori H** (2006) Auxin controls local cytokinin biosynthesis in the nodal stem in apical dominance. Plant J **45**: 1028–1036

**Tanner GJ, Francki KT, Abrahams S, Watson JM, Larkin PJ, Ashton AR** (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, Verries C, Cheynier V, Romieu C** (2009) Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. Plant Physiol **149**: 1028–1041

**Ullah C, Unsicker SB, Fellenberg C, Constabel CP, Schmidt A, Gershenzon J, Hammerbacher A** (2017) Flavan-3-ols are an effective chemical defense against rust infection. Plant Physiol **175**: 1560–1578

**Ullah C, Unsicker SB, Reichelt M, Gershenzon J, Hammerbacher A** (2019) Accumulation of catechin and proanthocyanidins in black poplar stems after infection by *Plectosphaerella populi*: hormonal regulation, biosynthesis and antifungal activity. Front Plant Sci **10**: 15

**Wallis IR, Edwards MJ, Windley H, Krockenberger AK, Felton A, Quenzer M, Ganzhorn JU, Foley WJ** (2012) Food for folivores: nutritional explanations linking diets to population density. Oecologia **169**: 281–291

**Wang L, Ran L, Hou Y, Tian Q, Li C, Liu R, Fan D, Luo K** (2017) The transcription factor MYB115 contributes to the regulation of proanthocyanidin biosynthesis and enhances fungal resistance in poplar. New Phytol **215**: 351–367

**Ward D, Young TP** (2002) Effects of large mammalian herbivores and ant symbionts on condensed tannins of *Acacia drepanolobium* in Kenya. J Chem Ecol **28**: 921–937

**Waterman P, Mole S** (1994) Analysis of phenolic plant metabolites. Blackwell Scientific Publications, Oxford, 73-99.

**Weiss M, Mikolajewski S, Peipp H, Schmitt U, Schmidt J, Wray V, Strack D** (1997) Tissue-specific and development-dependent accumulation of phenylpropanoids in larch mycorrhizas. Plant Physiol **114**: 15–27

**Weiss M, Schmidt J, Neumann D, Wray V, Christ R, Strack D** (1999) Phenylpropanoids in mycorrhizas of the Pinaceae. Planta **208**: 491–502

**Westley R** (2015) Investigating potential physiological roles of condensed tannins in roots of *Populus*: localization and distribution in relation to nutrient ion uptake.

**Xie DY, Dixon RA** (2005) Proanthocyanidin biosynthesis - still more questions than answers? Phytochemistry **66**: 2127–2144

**Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA** (2003) Role of anthocyanidin

reductase, encoded by BANYULS in plant flavonoid biosynthesis. Science (80- ) **299**: 396–399

**Xu WJ, Dubos C, Lepiniec L** (2015) Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. Trends Plant Sci **20**: 176–185

**Yoshida K, Iwasaka R, Kaneko T, Sato S, Tabata S, 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, Ma D, Constabel CP** (2015) The MYB182 protein down-regulates proanthocyanidin and anthocyanin biosynthesis in poplar by repressing both structural and regulatory flavonoid genes. Plant Physiol **167**: 693–710

**Yoshihara E, Minho AP, Yamamura MH** (2013) Anthelmintic effect of condensed tannins in gastrointestinal nematodes of sheep (*Ovis aries*). Semin Agrar **34**: 3935–3950

**Yuan L, Wang LJ, Han ZJ, Jiang YZ, Zhao LL, Liu H, Yang L, Luo KM** (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

**Zeng X, Du Z, Xu Y, Sheng Z, Jiang W** (2019) Characterization of the interactions between banana condensed tannins and biologically important metal ions ($Cu^{2+}$, $Zn^{2+}$ and $Fe^{2+}$). Food Research International **123**:518–528

**Zhang B, Yang X, Yang C, Li M, Guo Y** (2016) Exploiting the CRISPR/Cas9 system for targeted genome mutagenesis in *Petunia*. Sci Rep **6**: 20315

**Zhang H, Zhang J, Wei P, Zhang B, Gou F, Feng Z, Mao Y, Yang L, Zhang H, Xu N, et al** (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant Biotechnol J **12**: 797–807

**Zhou HC, Lin YM, Wei SD, Tam NFY** (2011) Structural diversity and antioxidant activity of condensed tannins fractionated from mangosteen pericarp. Food Chem **129**: 1710–1720

**Zhou X, Jacobs TB, Xue LJ, Harding SA, Tsai CJ** (2015) Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and redundancy. New Phytol **208**: 298–301

**Zifkin M, Jin A, Ozga JA, Irina Zaharia L, Schernthaner JP, Gesell A, Abrams SR, Kennedy JA, Peter Constabel C** (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

# Appendix

## Appendix 1

### Supplemental Table 1.1: List of primers for MYB115 and MYB134 cloning and sequencing.

| Amplicon name | Forward primer (5'-3') | Reverse primer (5'-3') | Amplicon length (bp) | Tm (°C) |
|---|---|---|---|---|
| MYB115_exon1 | GAGAAGGGCAGATAGGAAGAAGG | GCAACAGTTTCAAGTGCTCG | 563 | 58 |
| MYB115_exon2 | CATCTGACCTATCTCTCACCCTTG | ATTGTCATACCAGAAGTGACTCGG | 777 | 60 |
| MYB134_exon1 | GGCCTATAAAAGGGGAATAGACC | GCCTTCCAGCTATTAAAGACCATC | 757 | 58 |
| MYB134_exon3 | ATCTTAGACCGGACATCAAGAGAG | TTCTCGTCAATGTCTCATCTCAGG | 803 | 59 |

### Supplemental Table 1.2: List of primers for CRISPR constructs.

| Primer name | Sequence |
|---|---|
| 35S*Spe*I_MtU6F | CGTGCTCCACCATGTTGGGAATGCCTATCTTATATGATCAATGAGG |
| *Apa*I_ScaffoldR | GTGCTCCACCATGTTGGGCCAAAAAAAGCACCGACTCGGTG |
| *Swa*I_MtU6F | GATATTAATCTCTTCGATGAAATTTATGCCTATCTTATATGATCAATGAGG |
| *Spe*I_ScaffoldR | GTCATGAATTGTAATACGACTCAAAAAAAAGCACCGACTCGGTG |
| UNS1_MtU6F | CATTACTCGCATCCATTCTCATGCCTATCTTATATGATCAATGAGG |
| UNS1_ScaffoldR | GAGAATGGATGCGAGTAATGAAAAAAAGCACCGACTCGGTG |

### Supplemental Table 1.3: List of primers for qPCR analysis.

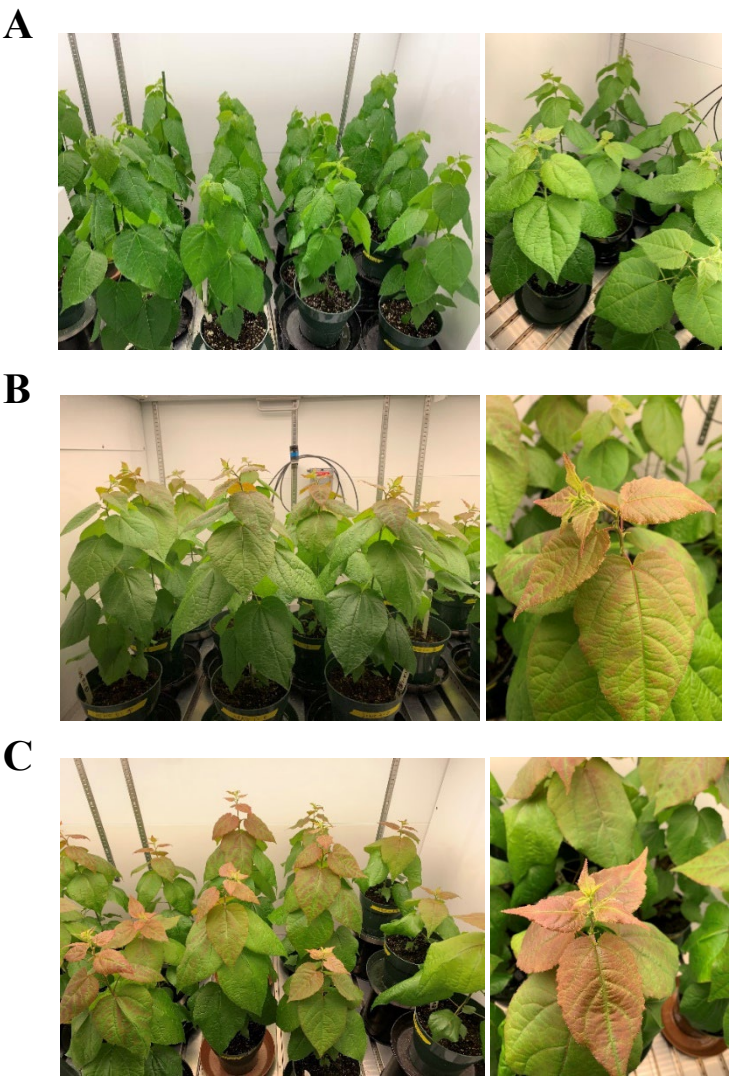
| Gene (Potri number) | Forward primer (5'-3') | Reverse primer (5'-3') | Amplicon (bp) | Tm (°C) |
|---|---|---|---|---|
| EF1b (Potri.001G224700) | AAGAGGACAAGAAGGCAGCA | CTAACCGCCTTCTCCAACAC | 145 | 58 |
| Act (Potri.001G309500) | CCCATTGAGCACGGTATTGT | TACGACCACTGGCATACAGG | 235 | 56 |
| MYB115 (Potri.002G173900) | GGATTGTGATAATGGGGTTGCC | GTGACTCGGTGAAGGAGTTT | 185 | 58 |
| MYB134 (Potri.006G221800) | GGACACTGGAATGAGTTTCAA | ATGTGCCAAAGATTCCAAGTC | 183 | 60 |
| F3'5'H1 (Potri.009G069100) | GCAACGGCTCATGAACGCAAGG | ATGCTCGAGGAAGTGTCAGTGC | 152 | 60 |
| DFR2 (Potri.002G033600) | CCAAGACTTTAGCAGAGCA | TGTTAGCATCATCCGAGTTG | 282 | 58 |
| ANR1 (Potri.004G030700) | GCTACCCTGCCTCCAAGACA | CGTGCGTGATTGAGATCGAGCC | 228 | 58 |
| LAR2 (Potri.010G129800.1) | ATGAACGAGCTTGCTTCTTTGTG | GCCCACATCCTTTTTACTGTCCT | 291 | 60 |
| LAR3 (Potri.015G050200) | CCTCGAATGTGGCCACCCCAC | GCTATGCTTGCACCACCAACAGC | 201 | 58 |

## Appendix 2

**Supplemental Table 2.1: Growth chamber conditions.**

| | |
|---|---|
| Temperature | 18-26°C |
| Day length | 16h (7am - 23 pm) |
| Light intensity (μmol $m^{-2}$ $s^{-1}$) | 1day - 5day: 600 |
| | 6day - 20day: 1000 |

**Supplemental Figure 2.1: Plants in high-light growth chamber.** Plants on the first day (A), at the first stage Day 3 (B), and at the second stage Day 15 (C).



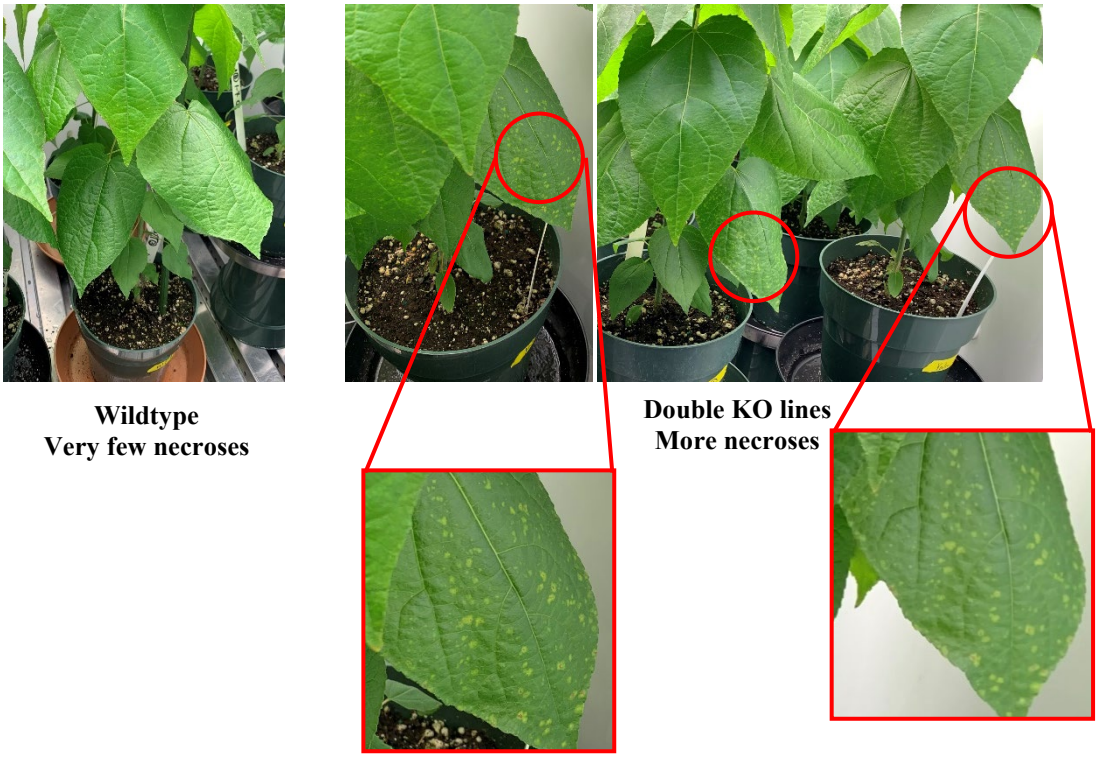
**Supplemental Figure 2.2: Phenotypes of double knockouts and wild-type plants treated by high-light for 20 Days.** (A) Branches[a], (B) Necrosis level.

**A**

**Wildtype**
**Almost no branches**

**Double KO lines**
**Lots of branches**

**B**



**Wildtype**
**Very few necroses**

**Double KO lines**
**More necroses**

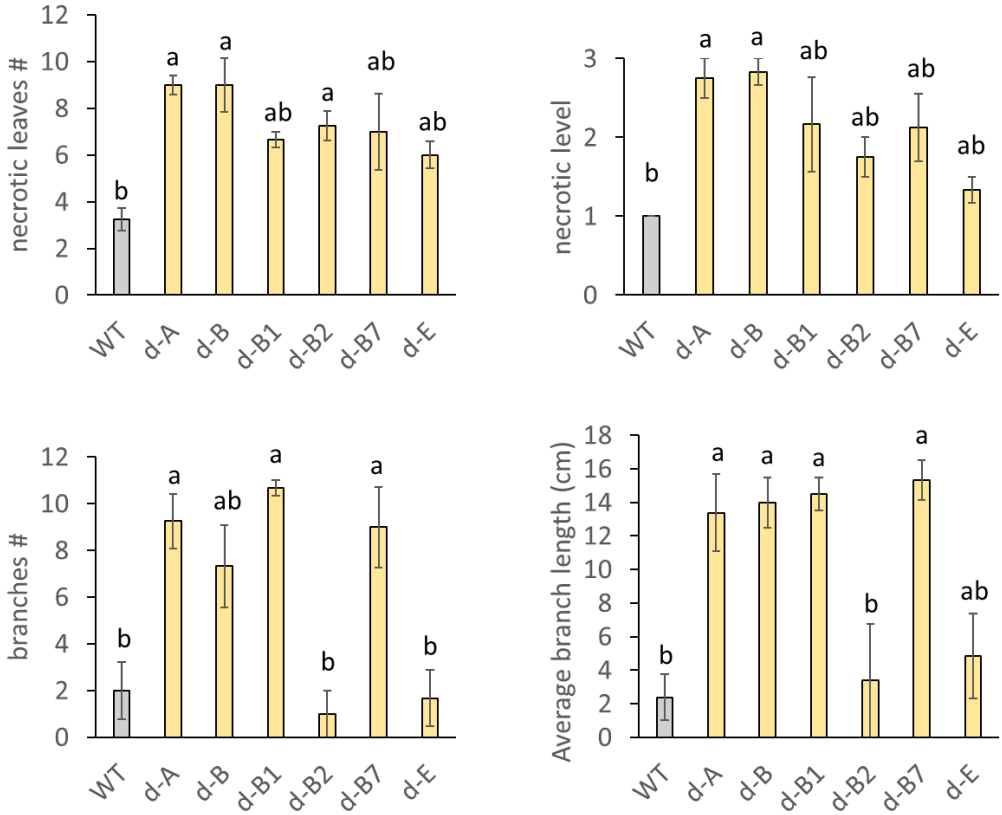
[a] Red arrows point out the branches outgrown from stems of double-KO lines.

**Supplemental Table 2.2: A summary of branch number and necrosis level for double knock-out plants and wild-types.**

| lines | necrotic leaves[a] | necrotic level[b] | branches[c] | length[d] |
|---|---|---|---|---|
| WT | 4 | 1 | 0 | 0 |
| WT | 2 | 1 | 0 | 0 |
| WT | 4 | 1 | 3 | 5 |
| WT | 3 | 1 | 5 | 4.5 |
| d-A | 10 | 3 | 11 | 15 |
| d-A | 9 | 2 | 11 | 16 |
| d-A | 9 | 3 | 9 | 16 |
| d-A | 8 | 3 | 6 | 6.5 |
| d-B | 9 | 3 | 4 | 12.5 |
| d-B | 11 | 3 | 8 | 17 |
| d-B | 7 | 2.5 | 10 | 12.5 |
| d-B1 | 7 | 2.5 | 11 | 16.5 |
| d-B1 | 7 | 3 | 11 | 13.5 |
| d-B1 | 6 | 1 | 10 | 13.5 |
| d-B2 | 9 | 2 | 0 | 0 |
| d-B2 | 7 | 2 | 4 | 13.5 |
| d-B2 | 6 | 1 | 0 | 0 |
| d-B2 | 7 | 2 | 0 | 0 |
| d-B7 | 7 | 2.5 | 6 | 13 |
| d-B7 | 3 | 1 | 9 | 17 |
| d-B7 | 7 | 2 | 12 | 16 |
| d-B7 | 11 | 3 | 0 | 0 |
| d-E | 5 | 1 | 0 | 0 |
| d-E | 6 | 1.5 | 1 | 6 |
| d-E | 7 | 1.5 | 4 | 8.5 |

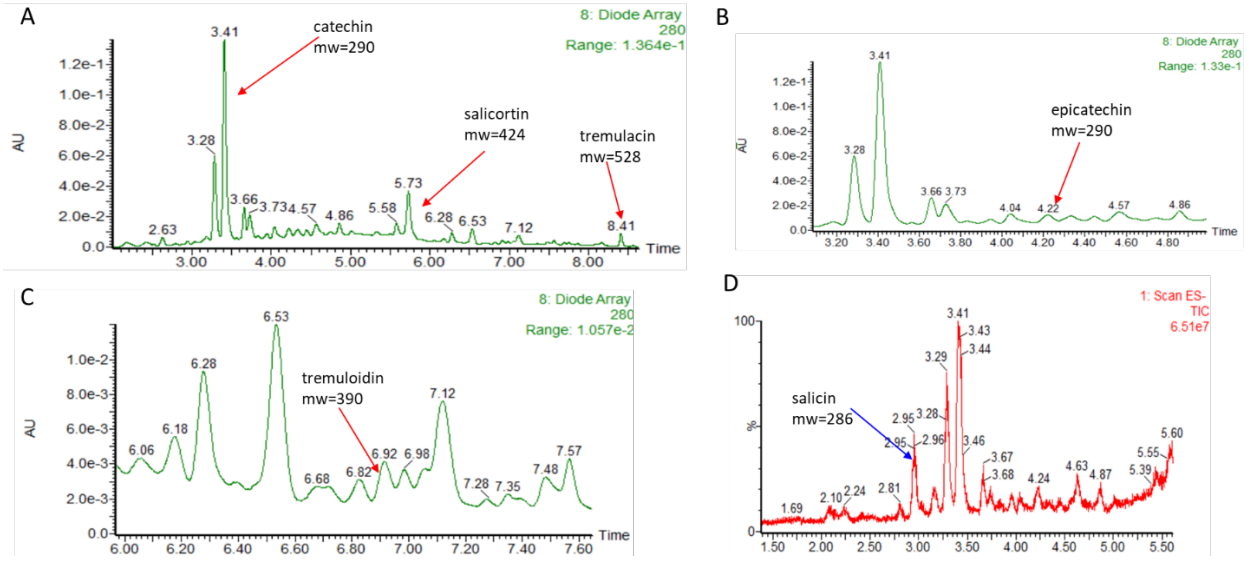
[a] Total number of necrotic leaves counted from bottom.
[b] Necroses level. 1 represents lowest (<5%), 1.5 represents low (~ 10 %), 2 represents medium (~30 %), 2.5 represents high (~50%), 3 represents severe (almost 70% of leaf area covered).
[c] Total number of branches outgrown from stems.
[d] The average length of branches. Values are represented as cm.

**Supplemental Figure 2.3: Branch number and necrosis level for double knock-out plants and wild-types.** The bars represent means ± SE of three or four individual biological replicates per line. Data were analyzed by one-way ANOVA followed by Tukey's HSD post hoc test. Different letters on the bar indicate statistically significant difference according to Tukey's HSD test at 95% confidence (P-value < 0.05).



## Appendix 3

**Supplemental Figure 3.1: UPLC analysis of soluble phenolics from wild-type poplar plants.** These peaks are identified by UV spectra at 280nm (shown in green colour, panel A-C) and MS profile (not shown) of representative standards and quantified in all knock-out lines. Negative ionization mode is shown in red colour (panel D).



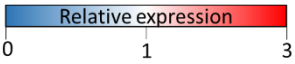

**Supplemental Table 3.1: Concentrations[a] of major poplar salicinoids in control and knock-out hairy root lines as determined by UPLC-MS.**

| Compound | EV | 115-A1 | 115-B23 | 115-C5 | 134-A3 | 134-A20 | 134-A10 |
|---|---|---|---|---|---|---|---|
| Tremulacin | 4.84 ± 1.14 | 1.41 ± 0.60 | 3.89 ± 1.62 | 3.55 ± 0.86 | 6.01 ± 1.36 | 2.14 ± 0.46 | 3.55 ± 0.86 |
| Tremuloidin | 4.13 ± 2.00 | 0.91 ± 0.68 | 2.88 ± 1.70 | 2.16 ± 1.47 | 2.90 ± 1.74 | 1.71 ± 0.24 | 3.59 ± 0.34 |
| Salicortin | 2.48 ± 0.26 | 0.88 ± 0.52 | 1.25 ± 0.37 | 1.95 ±0.63 | 1.83 ±0.45 | 0.92 ± 0.32 | 1.93 ± 0.54 |

[a] Concentrations are presented as µg $mg^{-1}$ dry weight ± SE. No significant difference from controls was determined using a one-way ANOVA.

## Appendix 4

**Supplemental Table 4.1: Heatmap of relative gene expression for the flavonoid and MYB genes quantified by qPCR in control, MYB115-KO and MYB134-KO hairy roots.**

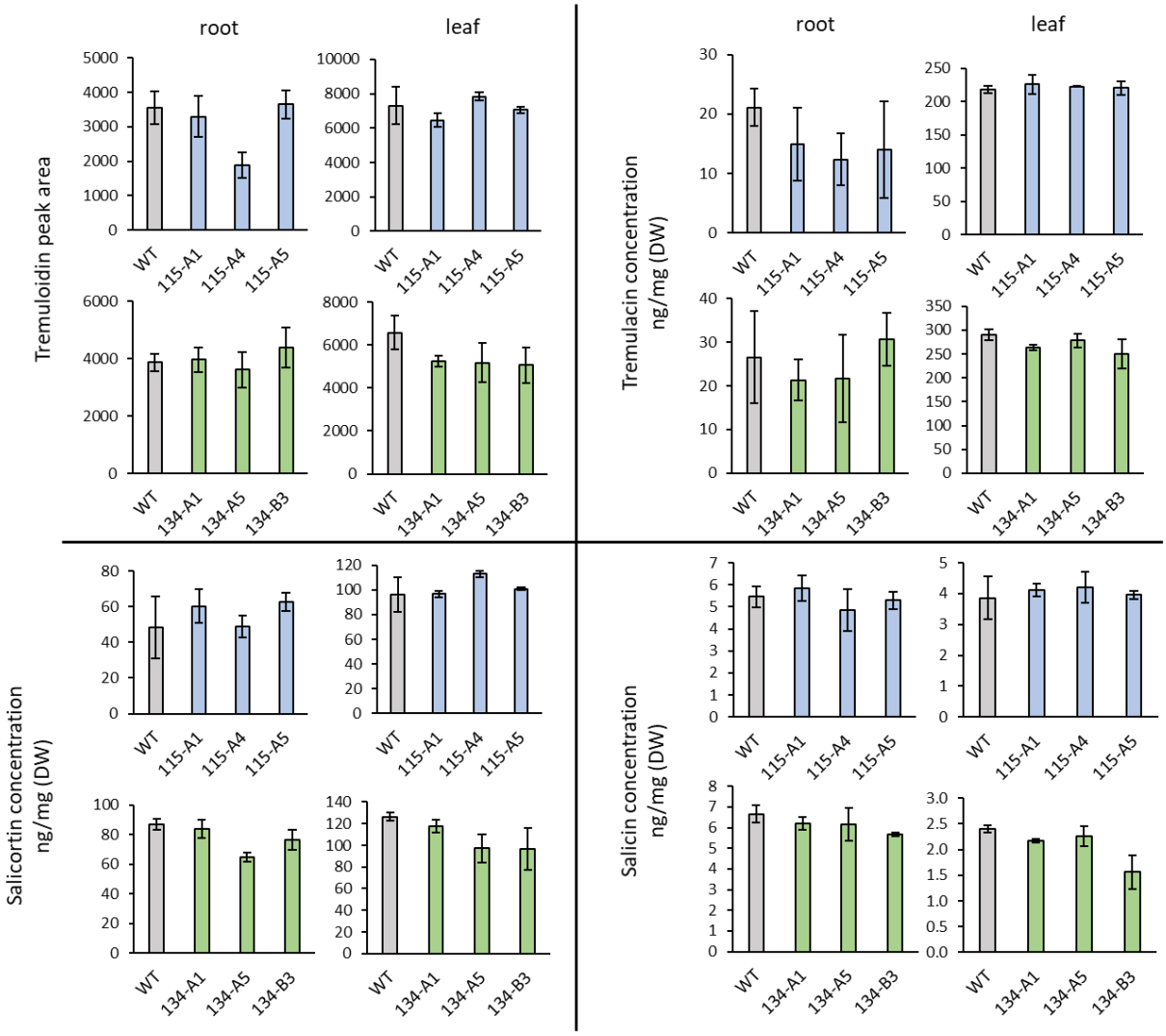



| | MYB115 | MYB134 | F3'5'H1 | DFR2 | ANR1 | LAR3 |
|---|---|---|---|---|---|---|
| EV | 1.00[a] | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 115-A1 | | 1.21 | 1.74 | 1.57 | 1.17 | 2.00 |
| 115-B23 | | 1.57 | 2.27 | 1.16 | 0.91 | 1.27 |
| 115-C5 | | 1.34 | 1.44 | 1.55 | 0.97 | 1.38 |
| 134-A3 | 0.55 | | 1.06 | 0.52 | 0.73 | 0.82 |
| 134-A20 | 0.85 | | 1.13 | 0.55 | 0.41 | 0.76 |
| 134-A10 | 0.69 | | 0.94 | 0.78 | 0.79 | 0.70 |

[a] Values represent relative gene expression level. No significant difference from controls was determined using a one-way ANOVA.

## Appendix 5

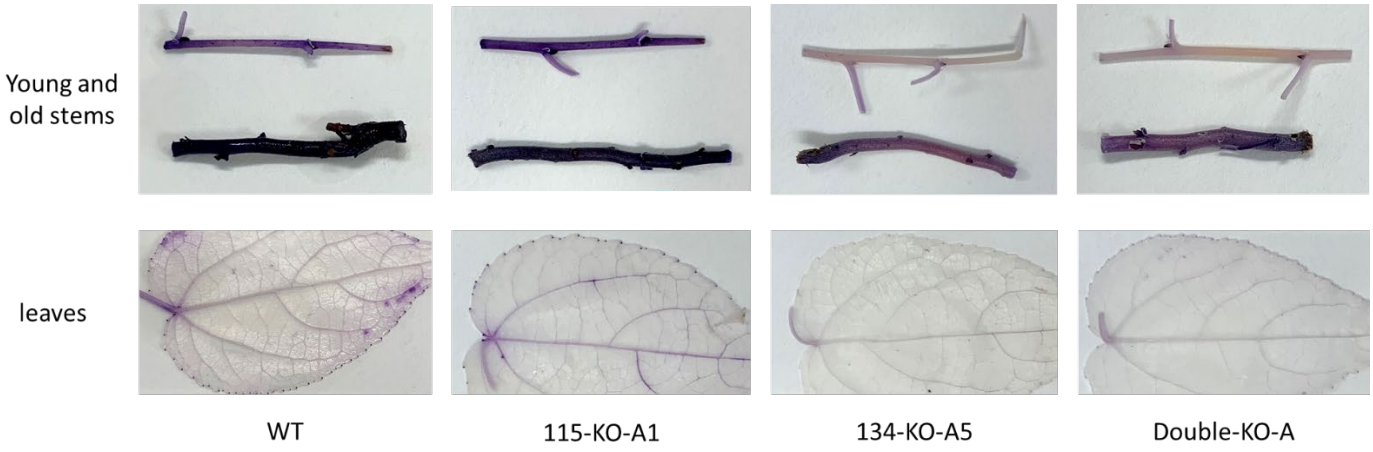
**Supplemental Figure 5.1: Concentrations of key salicinoids in leaves and roots of MYB115-KO and MYB134-KO plants as determined using UPLC in high light exposed poplar plants.**




Data points are presented as mean ± SE. At least three biological replicates per line were analyzed. No significant differences from controls were found using a one-way ANOVA. Colors of bars represent different lines, wild-type plants are shown in grey, MYB115-KO lines are shown in blue, MYB134-KO lines are shown in green.

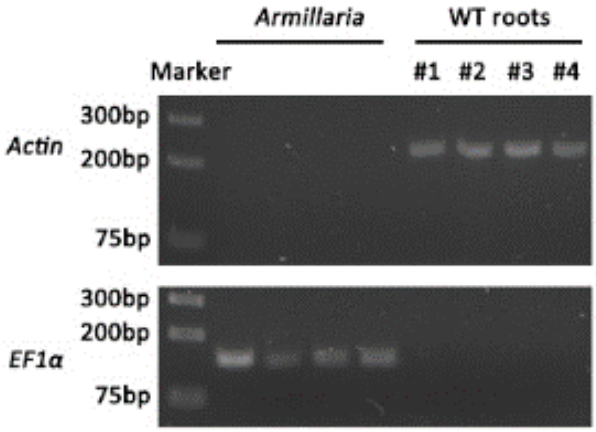
## Appendix 6

**Supplemental Figure 6.1: DMACA staining for different tissues of knock-out mutants and wild-type obtained from mist chamber.**



## Appendix 7

**Supplemental Figure 7.1: Specificity of *Armillaria* housekeeping gene *EF1α* and Poplar housekeeping gene *actin*.**




#1-4 represent four individual wild-type poplar root DNA.

Four isolates of *Armillaria* were used for this experiment.