UNEXPECTED CONNECTIONS:
SALICINOID BIOSYNTHESIS IN POPLAR

by

Harley O. W. Gordon

B.Sc. University of Victoria, 2015
M.Sc. University of Guelph, 2018

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

DOCTOR OF PHILOSOPHY

in the Department of Biology

Supervisory Committee

UNEXPECTED CONNECTIONS: SALICINOID BIOSYNTHESIS IN POPLAR

by

Harley O. W. Gordon
B.Sc. University of Victoria 2015
M.Sc. University of Guelph 2018

Supervisory Committee

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

Dr. Jürgen Ehlting, (Department of Biology)
Departmental Member

Dr. Patrick von Aderkas, (Department of Biology)
Departmental Member

Dr. Alisdair Boraston, (Department of Biochemistry and Microbiology)
Outside Member

Abstract

Populus is a genus distributed across the northern hemisphere. Poplars (Salicaceae) are subject to stresses in their environment such as herbivory, drought, and fire. These perennial hardwoods produce abundant phenylpropanoid derived anti-herbivory molecules called salicinoids. Understanding salicinoid function and biosynthesis is crucial for understanding the chemical ecology, carbon balance, and adaptability of poplars to changing ecosystems. The full biosynthetic pathway of salicinoids is unknown; however, recent progress has identified a biosynthetic gene for salicinoids. The gene is a UDP-dependent glycosyl transferase called UGT71L1.

Using CRISPR/Cas9 genome editing, UGT71L1 was disrupted in the hybrid poplar *Populus tremula x Populus alba*. Through metabolomic, transcriptomic, and biochemical techniques the co-dependent nature of growth, defence and salicinoid biosynthesis in poplars was explored. Following the elimination of UGT71L1, the exogenous application of deuterated benzenoids and mass spectroscopic analysis was used to examine biochemical connections across metabolic pathways. A carbon limited growth experiment was used to assess the capacity for glucosylated salicinoids to contribute to carbon reserves in resprouting trees. In addition, a second glycosyltransferase gene, UGT78M1, was disrupted in genome-edited poplars.

Interruption of UGT71L1 disrupted salicinoid biosynthesis. UGT71L1 knockout plants had small crinkled leaves, reduced growth, and were preferred by insect herbivores. Growth impacts were caused by the abundance of salicylic acid, which increased in concentration following salicinoid biosynthesis interruption. Furthermore, we determined that benzyl benzoate is a precursor to salicortin biosynthesis. Salicinoids are also an inaccessible carbon sink in poplar that cannot be remobilized during carbon starvation. The hypothesized salicinoid biosynthetic gene UGT78M1 does not contribute to salicinoid biosynthesis; however, UGT78M1 is crucial for salicyl benzoate glucoside homeostasis. This dissertation highlights the small molecule trichotomy of plant biochemistry and identifies connections between specialized metabolites, phytohormones, and primary metabolites.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I am grateful for the knowledge and passion shared by my supervisor Dr. Peter Constabel throughout my doctoral work. I extend my gratitude to the tutelage of my colleague Dr. Dawei Ma, whose insight and expertise were invaluable. I acknowledge and am grateful for all the Constabel lab members, past and present. My work has been conducted on the previous research of talented graduate, post-doctoral, and undergraduate members of the Constabel lab. I acknowledge the various organizations that have provided funding throughout the program. NSERC, University of Victoria, UVIC Centre for Forest Biology, the Max Planck Society, and the Canadian Society of Plant Biologists have all contributed to the completion of this dissertation.

Chapter 2 is a significant component of my research and was a collaborative project. I am grateful for the training provided by Dr. David Overy and Amanda Sproule at the Agri-food Canada experimental farm in Ottawa. I extend my gratitude to the biochemistry department at the Max-Planck institute for Chemical Ecology in Jena, Germany. In particular, the assistance of Drs. Nathalie Lackus, Tobias Köllner, Michael Reichelt, and Jonathan Gershenzon was academically inspiring. At the Uvic Centre for Forest Biology I thank Megan Loland for her efforts in generating rescue and overexpressor lines, Brad Binges for greenhouse management, Carmen Lea for providing tussock moth larvae, and Luis Barrena for his work with insect feeding assays.

I acknowledge Dr. Jeremy Wulff and his laboratory members for the synthesis of deuterated and novel dibenzenoid molecules for my labelling experiments and enzyme assays presented in Chapter 3 and Appendix 6. Chapter 3 would not have been possible without the substrates produced by J. Wulff. I would like to extend my thanks and acknowledgement to my collaborators Dr. Rachel Hillabrand, Dr. Simon Landhäusser, and Pak Chow at the University of Alberta for their insight and assistance in the study of carbon allocation in poplars as presented in Chapter 4. I thank Dr. Ori Granot for his assistance with the LC/MS analysis conducted throughout my dissertation.

The work begun by Dr. Christin Fellenberg was crucial for transgenic generation of UGT78M1 knockout trees. I thank Quinn Mckechnie for his assistance in their

generation. I thank my committee member, Dr. Jurgen Ehlting for his advice on genotyping. In the spirit of completion, I acknowledge the assistance of Dr. Yuriko Carrington, Dr. Benedicte Albrectsen, Yalin Liu, and Ahalya Rajendran for their contributions to information presented in the appendices.

Dedications

Nothing worth doing can be done alone.

My family and friends have helped more than they know.

With humility, thank you.

For Catriona

List of Abbreviations

3OHBB	3-hydroxybenzyl benzoate
4CL	4-coumaroyl-CoA ligase
ABA	abscisic acid
BA	benzoic acid
BA-Glc	benzoic acid glucoside
BEBT	benzoyl-CoA:benzyl alcohol O-benzoyltransferase
BG	benzyl gentisate
bp	base pairs
BS	benzyl salicylate
BS-Glc	benzyl salicylate glucoside
ºC	degrees Celsius
cDNA	coding DNA
CoA	coenzyme A
CYP	cytochrome P450
cm	centimeter
CM	chorismate mutase
CRISPR	clustered regularly interspaced short palindromic repeats
d	days
DAD	diode array detector
DMSO	dimethyl sulfoxide
DW	dry weight
ESI	electrospray ionization
eV	electron volts
Fig.	figure
FC	fold-change
FW	fresh weight
g	gram
g	gravity
GA	gibberellic acid
GB	gentisyl benzoate
GFS	glucose fructose sucrose
Glc	glucoside
gRNA	guide RNA
GO	gene ontology
GUS	β-glucuronidase
h	hours
HCH	6-hydoxy-2-cyclohexen-on-oyl
HPLC	high pressure liquid chromatography
HSD	honest significant difference
IAA	indole acetic acid
Indel	insertion and deletion
JA	jasmonic acid

JAZ	jasmonate ZIM domain
KO	knockout
L	litre
LB	Luria-Bertani broth
LC	liquid chromatography
LC/MS	liquid chromatography mass spectroscopy
m	metre
mAU	milli absorbance unit
min	minute
mg	milligram
mL	milliliter
mm	millimeter
mM	millimolar
mRNA	messenger RNA
MS	mass spectroscopy / Murashige and Skoog
MYB	myeloblastosis transcription factor
m/z	mass to charge
n	sample size
NAD	nicotinamide adenine dinucleotide
nd	not detected
NHEJ	non-homologous end joining
ng	nanogram
nm	nanometer
NMR	nuclear magnetic resonance
No.	number of
NPK	nitrogen phosphorous potassium
NPR	nonexpressor of pathogenesis-related genes
ns	non-significant / not significant
NSC	non-structural carbohydrates
OPDA	12-oxo-phytodienoic acid
PAL	phenylalanine ammonia lyase
PAM	protospacer adjacent motif
PAR	photosynthetically active region
PC	principal component
PCR	polymerase chain reaction
PFCW	protein free cell wall
psi	pressure per square inch
V	volts
v	volume
RNAi	RNA interference
rpm	revolutions per minute
s	seconds
SA	salicylic acid
SA-Glc	salicylic acid glucoside
SABT	benzoyl-CoA:salicyl alcohol O-benzoyltransferase
SAGT	salicylate glycosyltransferase

SB	salicyl benzoate
SB-Glc	salicyl benzoate glucoside
SCPL	serine carboxypeptidase-like
SE	standard error
Spp.	species
SS	salicyl salicylate
SS-Glc	salicyl salicylate glucoside
UDP	uridine diphosphate
UGT	UDP-dependent glycosyltransferase
UPLC	Ultra-high performance liquid chromatography
UV	ultraviolet
w	weight
W	watt
WT	wild-type
µg	microgram
µL	microliter
µm	micrometer
µM	micromolar
µmol	micromoles

Chapter 1 – General Introduction

1.1 Introduction

From an amoeba to an elephant, from a grove of aspens to a cluster of mushrooms, natural selection acts upon all life. Across a population, individuals may possess a combination of variable and heritable traits that can be evolutionary advantageous. For complex animals, adaptive success can be profoundly intuitive. The potential to succeed can often be measured through physical attributes such as size, speed, strength, diet, and sociability. In humans the capacity for collaboration, foresight, and problem solving has led to the global evolutionary success of our species. However, in organisms like plants which have no brain, no muscles, and on the surface – no way to communicate, it is not always clear what adaptations can be attributed to species success. Some plants persevere and succeed while others die; the mechanism by which survival occurs is often hidden. For a casual observer, many plants appear to be strikingly similar. Grasses (Poaceae) for example, have enormous species diversity, but for many they look the same and occupy the same environmental niche (Thompson et al., 2004). What contributes to fitness in a group of organisms that cannot run, hunt, or fly? What defines success for organisms that eat sunlight and air? How is it that certain plant species or families have come to dominate the terrestrial biosphere? The capacity for photosynthesis is central, but in the competition of plant against plant, and against all the herbivores that would do them harm, what tools are available for surviving and thriving in a competitive biosphere.

There are countless unseen interactions occurring constantly in plants and they are influenced by the surrounding environment. Rather than physical strength, intelligence, or speed many plant adaptations are biochemical. Fluctuations in abiotic and biotic conditions, from the sunrise to the seasonal migrations of herbivores, can cause plants to alter their chemistry – and the chemistry of plants is diverse. Interactions between plants and their environments often occurs on a molecular scale, and as such it can appear invisible. The importance of innocuous and unseen small molecules in plant biology is paramount (Heldt and Piechulla, 2021). This dissertation aims to highlight both the

importance, functional niche, and the beautiful complexity of one small group of molecules in one highly successful plant genus.

1.1.1 Plants interact and adapt to their environment through small molecules

Heat, drought, and herbivory are some of the stresses that plants may be subject to in a single day. As sessile organisms, relocating to a better environment is not an option nor is it required. Plants must survive in place and adapt to stress. However, as sessile autotrophs there are an enormous number of mechanisms plants have evolved for adaptation to environmental conditions. Many of these adaptations involve changes in the metabolism of small molecules (< 1,000 Daltons).

Small molecules in plants can be broadly categorized into three groups. One category is “primary metabolites” which consist of molecules that are conserved across many plant species, genera, and families. There are up to 10,000 unique primary metabolites shared among plants (Pichersky and Lewinsohn, 2011). The primary metabolite designation includes molecules such as glucose, nucleic acids, citric acid, and vitamin C. Primary metabolites such as glucose often contribute to the biosynthesis and structure of the next, much larger, category of molecules called specialized metabolites. Historically, specialized metabolites were called secondary metabolites, but the phrase is considered outdated. It originates from a time when researchers did not understand the biological and ecological importance of plant small molecule diversity. Specialized metabolite diversity is immense, the number of unique plant specialized metabolites is estimated to range from 200,000 to over 1,000,000 different compounds (Yonekura-Sakakibara and Saito, 2009; Afendi et al., 2012). Individual metabolites are often unique to species, genera, or families (Lou et al., 2022). Over 95% of the biochemical diversity in plants is specialized metabolites. In plants, the specialized metabolites can contribute to many roles such as stress adaptation, defense, and enzyme regulation (Bjarnholt et al., 2018; Erb and Kliebenstein, 2020; Yokoyama et al., 2021). A third category of plant small molecules, called hormones or regulators, are compounds that contribute to the regulation of broad processes throughout the plant. Disturbances in hormone biosynthesis and metabolism can impact all aspects of plant growth and development (Traw and

Bergelson, 2003; Gallego-Giraldo et al., 2011). Small molecule regulators like gibberellins and abscisic acid control mechanisms such as seed germination (Popko et al., 2010). Although most regulators and phytohormones are shared across plant families their broad effects and roles necessitate a separate functional group (Fig. 1.1).

Specialized metabolites can have multiple biological roles in plants, but many have evolved as defense toxins. Insect herbivory is a major biotic stress to plants and increased biosynthesis of anti-herbivory specialized metabolites, such as phenolics in infested hemlock trees, can reduce damage (Rigsby et al., 2020). Induced changes to specialized metabolite biosynthesis can be highly specific. For example, mechanical and herbivore damage cause similar tissue disruption but can cause different biochemical responses. In Cowpea (*Vigna unguiculata*) the oral secretions of the fall armyworm (*Spodoptera frugiperda*) can induce the production of plant hormones and specialized volatile metabolites. Volatile signals released due to herbivory can result in biochemical changes to neighbouring plant tissue, or other nearby plants, thus signalling herbivore stress to non-damaged leaves. Therefore, giving the neighbouring plant time to alter gene expression and ultimately its phytochemical profile in preparation for anticipated herbivore damage (Wu and Baldwin, 2009). However, strictly mechanical damage does not result in the same response and release of volatile signals (Schmelz et al., 2007). Specialized metabolite concentrations can also be altered in response to abiotic stresses such as high light and drought. For example, a typical plant response to high light is the accumulation of anthocyanins in foliar tissue. Anthocyanins are specialized metabolites that provide to plant resistance in high light environments through visible and ultraviolet light absorption and reactive oxygen species sequestration (Steyn et al., 2002; Yokoyama et al., 2021).

Specialized metabolites are categorized into three major structurally defined groups: the nitrogen containing compounds including alkaloids, isoprene derived terpenoids, and phenylpropanoids synthesized from aromatic amino acids. Alkaloids are nitrogen containing compounds derived from amino acids and nucleotides. Many of the most well recognized plant metabolites such as caffeine and nicotine are alkaloids. The second and largest group of specialized metabolites is terpenoids, which are aliphatic and can form lipid and sterol structures (Yazaki et al., 2017). Terpenes are abundant in the

resin of conifer trees, where they can function to reduce insect damage (Whitehill et al., 2018). The third broad category of specialized metabolites is phenylpropanoids, which begin from tyrosine or phenylalanine. The high light induced anthocyanins are blue to red molecules derived from the phenylpropanoid pathway. Another group of phenylpropanoids known as condensed tannins can improve herbivory tolerance in poplar trees (Madritch and Lindroth, 2015). Of these three groups, much diversity arises from the modification of core structures with functional groups such as acetylation, prenylation, or glycosylation. Additionally, specialized metabolites from separate groups such as terpenoids and phenolics can come together forming terpenophenolics (Kavalier et al., 2011). Further modification can occur with peptides and sulfur containing functional groups. Combining metabolites from separate pathways leads to the combinatorial natural of phytochemical diversity (Jacobwitz and Weng, 2020).

With the remarkable diversity of plant specialized metabolites, there is substantial effort by researchers to understand the biosynthesis of specialized metabolites in many plant species. In nature there exists incredible enzymatic capacity for unique and valuable molecule generation. Much effort in the 20th and early 21st century on specialized metabolite biosynthesis focused on novel drug discovery (Cordell, 2000). Research in specialized metabolite biosynthesis can aid in understanding ecological and evolutionary implications of chemodiversity among distinct plant species. Metabolite diversity has evolved to allow plants to adapt to a hostile world (Hartmann, 2008; Pichersky and Lewinsohn, 2011).

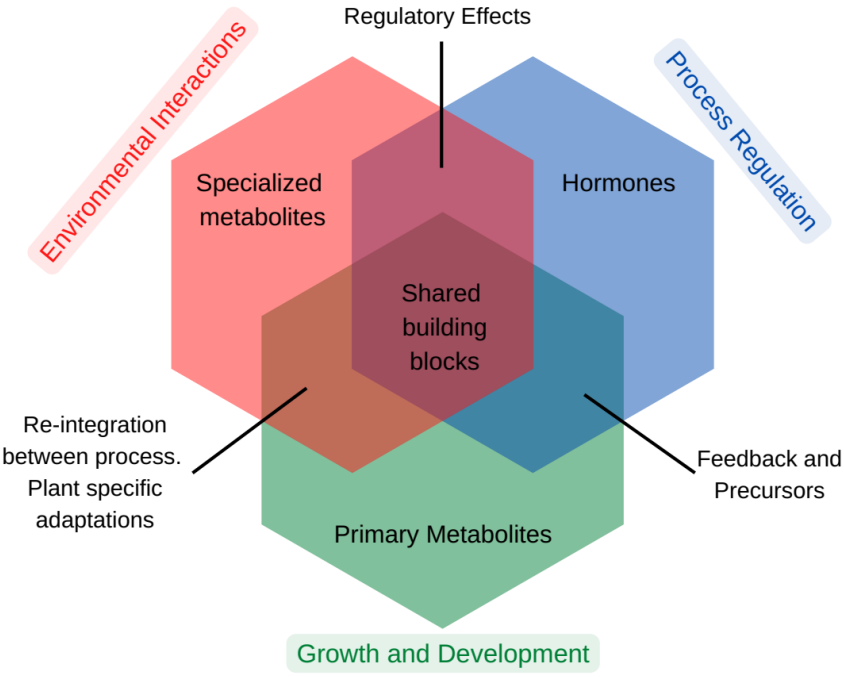


Figure 1.1. The functional small molecule trichotomy highlighting the interconnected nature of plant specialized metabolism. Adapted from Erb and Kleibenstein (2020)

As our understanding of plant biochemistry grows, our understanding of the diverse roles of plant specialized metabolites increases. (Fig. 1.1) (Dare et al., 2017; Erb and Kliebenstein, 2020). With up to a million unique compounds across an estimated 370,000 plant species, research into the biosynthesis and functionality of specialized metabolites has no shortage of questions (Afendi et al., 2012; Christenhusz and Byng, 2016). Changing climates, altered precipitation, and expanded pest distributions underscore the importance of studying plants capacity for biochemical adaptations in a changing world (Nicotra et al., 2010; Gauthier et al., 2014). In this dissertation, I have conducted research on the biosynthesis of phenylpropanoid derived specialized metabolites unique to the plant family Salicaceae. Poplars, aspens, cottonwoods (*Populus*) and willows (*Salix*) are all salicaceous species. My work focuses on poplar trees, where I used a common research hybrid aspen clone of *Populus tremula x Populus alba* INRA 717-1B4 (Mader et al., 2016).

1.1.2 Poplars produce phenolic metabolites called salicinoids

Rich genetic resources, rapid growth, amenability to genetic engineering, and phylogenetic grouping in the rosid clade have cemented poplar as a favoured organism in tree and plant biology (Wullschleger and Jansson, 2002; Jansson and Douglas, 2007). The model poplar species *P. trichocarpa* was the first tree to have its genome sequenced (Tuskan et al., 2006). Across the northern hemisphere the genus *Populus* (poplars and aspens) is ecologically important. Poplars perform ecosystem services such as preventing erosion, improving biodiversity, and remediating soil (Rogers et al., 2020). Poplars are fast growing trees and used as a biomass source for biofuel production (Stanton et al., 2021). Historically, poplar and willow bark has been an important source of pain relieving medicine. Poplars contain a phenylpropanoid derived specialized metabolite called salicin. Salicin is the source of the analgesic properties associated with poplar and willow bark (Desborough and Keeling, 2017). The origin of plant specialized metabolite research is intertwined with the salicinoids from Salicaceae. One of the first isolated plant specialized metabolites, from the bark of a willow tree, was salicin. Salicin was first isolated in the early 19th century by chemist Johann Buchner (Vlot et al., 2009). Understanding the biological properties of salicin ultimately led to the development of the essential medicine acetylsalicylic acid, or Aspirin (Desborough and Keeling, 2017). salicaceous species remain a source of novel phytochemicals with important medicinal properties such as miyabeacin, a metabolite with anti-cancer properties only recently isolated from two willow species (*Salix miyabeana* and *Salix dasyclados*) (Ward et al., 2020). The ecological success of *Populus* spp. and the broader Salicaceae family can be attributed, in part, to the unique chemical properties of poplar specialized metabolites.

Biochemical adaptations in poplar include abundant phenylpropanoid derived specialized metabolites. Poplars accumulate high concentrations of polymeric proanthocyanidins, known as condensed tannins, which function can mitigate plant stress (Gourlay and Constabel, 2019). Poplars also accumulate hydroxycinnamate esters, often in the form of chlorogenic acid isomers and shikimate esters (Babst et al., 2014). Hydroxycinnamate esters likely have multiple functions in poplar leaves, including as antioxidants and high light sunscreens. Hydroxycinnamate glucose esters are also

important acyl donors during the biosynthesis of other metabolites (Tsai et al., 2006). One hydroxycinnamate ester, grandidentatin, appears particularly susceptible to phenolic pathway perturbations in poplar (Chedgy, 2015). *Populus tremuloides* can accumulate very high concentrations (up to 20% dry weight) of glucosylated phenolic compounds known as salicinoids in young leaves (Donaldson et al., 2006). The defining moiety of salicinoids is a central glucosylated phenol ring with an ethanol functional group at the ortho position relative to the glycosidic bond. The central aromatic ring is known as salicyl alcohol, or saligenin. The simplest salicinoid is glucosylated salicyl alcohol, which is the compound salicin (Fig. 1.2). Complex and abundant salicinoids such as salicortin or tremulacin have a glucosylated two-ring structure, which can be described as a dibenzenoid. Salicinoids are often modified with acyl, sulfate, benzenoid, hydroxycinnamate, and 6-hydroxy-2-cyclohexen-on-oyl (HCH) moieties (Keefover et al., 2014).

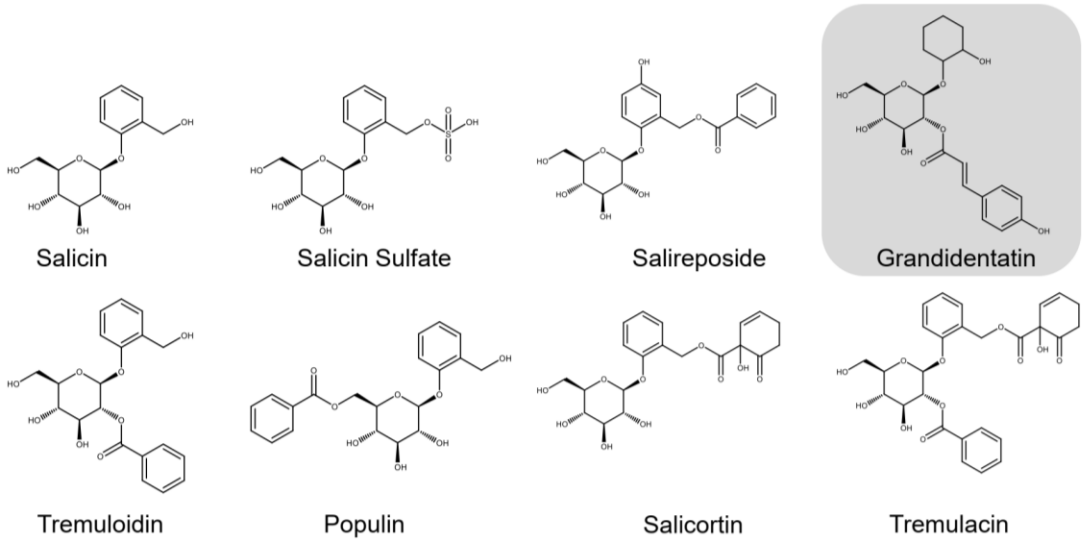


Figure 1.2 Common salicinoids and the coumarate ester grandidentatin found in the leaves of *Populus* spp. Grandidentatin is not a salicinoid and is indicated with a grey background.

There are more than 35 unique salicinoids found in poplar leaves (Keefover et al., 2014). The salicinoid profile varies considerably between species (Boeckler et al., 2011; Fabisch et al., 2019). For example, salicin sulfate and salirepin sulfate are salicinoids that accumulate in the leaves of *P. trichocarpa* and *Populus alba* but are not found in the leaves of the related *Populus nigra* (Lackus et al., 2020). In the leaves of our research hybrid *P. tremula x P. alba* salicin, salicortin, and tremulacin are the most abundant salicinoids (Fellenberg et al., 2020). Tissue localization, ontological variation, and genotypic variation all influence the concentration and chemical profile of salicinoids (Lindroth and Hwang, 1997; Donaldson et al., 2006). An analysis of salicinoid variation in a collection of *P. tremula* genotypes shows that salicinoid concentrations do not correlate with geographic location (Keefover et al., 2014). Rather it is genotype and environmental factors influencing insect herbivores that determine salicinoid concentrations in Salicaceae (Barker et al., 2019).

Salicinoids in poplar are potent anti-herbivory metabolites (Boeckler et al., 2011). The interaction of salicinoids with herbivores has been well studied in *Salix* and *Populus* species. A meta-analysis of insect herbivores while feeding on salicaceous species showed a significant relationship with condensed tannin and salicinoid concentrations, which provide insect resistance. Metabolism of consumed salicortin by the generalist lepidopteran herbivore *Lymantria dispar* shows that salicortin is metabolized into salicin and catechol, both of which reduce insect growth and survival (Boeckler et al., 2016). The salicinoid tremulacin is also a strong inhibitor of insect growth in *L. dispar* (Rubert-Nason et al., 2015). High salicinoid concentrations can reduce leaf consumption by browsing elk (Wooley et al., 2008). Phytochemical analysis of trees felled by beavers (*Castor canadensis*) showed an avoidance of *P. tremuloides* with higher concentrations of phenolic compounds (Basey et al., 1988). Furthermore, mammalian foraging can have landscape level impacts on *P. tremuloides* (quaking aspen) (Wooley et al., 2008). In *Salix* spp. there is a negative correlation between salicinoid concentrations and the feeding preference of mountain hares (*Lepus timidus*) (Tahvanainen et al., 1985). Mammalian and insect herbivory resistance demonstrates the broad-spectrum anti-herbivory activity of salicinoids. Interestingly, there are several minor salicinoids derivatives such as salicin-7-sulfate that do not appear to reduce herbivory (Lackus et al., 2020). Therefore, the

importance of understanding the biosynthesis of salicinoids is critical in improving our understanding of how widespread *Populus* spp. Genotypes with distinct salicinoid profiles will adapt to shifting herbivore pressures.

1.1.3 Salicinoid biosynthesis involves numerous unidentified biosynthetic steps

Despite the ecological and medicinal importance of salicinoids our understanding of the salicinoid biosynthetic pathway remains incomplete. However, technical advances in molecular biology, heterologous protein expression, and mass spectral analysis are reducing barriers in specialized metabolite pathway elucidation. Transcriptomic analysis using next generation sequencing and commercially available library preparation kits allows for novel gene expression analysis across multiple tissues or experiments. With a known or hypothesized gene in a pathway transcriptomic and coexpression analysis can reveal many biosynthetic candidate genes (Stark et al., 2019; Fellenberg et al., 2020). Using commercial oligonucleotide and gene synthesis, entire biochemical pathways can be reconstituted and tested in heterologous expression systems such as *Nicotiana benthamiana* in a matter of weeks (Jacobowitz and Weng., 2020). Increasingly accessible analytical equipment such as liquid chromatography – mass spectroscopy (LC/MS) is routinely used to quantify chemically similar metabolites in crude extracts without the need for extensive purification protocols (Afendi et al., 2012). Tandem mass spectroscopy of crude extracts can provide metabolite structural information which aids in the identification of isomers, functional groups, and novel molecules (Dührkop et al., 2019).

While our understanding of the entire salicinoid biosynthetic network is incomplete, radioisotope and stable isotope labelling of precursor molecules have demonstrated that salicinoids are derived from the phenylpropanoid pathway (Zenk, 1967; Babst et al., 2010). Research into condensed tannin regulatory transcription factors discovered that overexpression of a R2R3-MYB transcriptional repressors designated as MYB134 reduces salicinoid concentrations in poplar leaves (Mellway et al., 2009). Two putative salicinoid biosynthetic genes were identified using gene coexpression analysis of low salicinoid poplars. The two acyltransferase proteins mis-regulated in poplar

overexpressing MYB134 were similar to a benzyl benzoate synthesizing enzyme from *Clarkia breweri* (D'Auria et al., 2002; Mellway et al., 2009). As phenylpropanoid derived dibenzenoid esters such as benzyl benzoate were hypothesized to be involved in salicinoid biosynthesis (Babst et al., 2010), functional characterization of the acyltransferase enzymes was conducted. The two BAHD-type acyl transferase enzymes are capable of in vitro synthesis of two hypothesized salicinoid intermediates, salicyl benzoate and benzyl benzoate (Chedgy et al., 2015), and the acyltransferases were thus designated ACT47 (benzoyl-CoA: salicyl alcohol O-benzoyltransferase; SABT) and ACT49 (benzoyl-CoA: benzyl alcohol O-benzoyltransferase; BEBT). SABT and BEBT are postulated to contribute to the biosynthesis of salicinoid precursors, although direct evidence is still lacking. Coexpression analysis of the acyl-transferase genes was conducted to search for salicinoid biosynthesis candidate genes (Table 1.1). Two uridine diphosphate dependent (UDP) family 1 glycosyltransferase genes, annotated as UGT71L1 and UGT78M1, were highly coexpressed with SABT (Fellenberg et al., 2020). Phylogenetic clustering of UGT71L1 grouped the enzyme with UGTs expressing hydroxybenzoates glucosylation activity. UGT78M1 grouped with UGTs capable of flavonoid glucosylation, although UGT78M1 is not a flavonoid glycosyltransferase (Veljanovski and Constabel, 2013).

Table 1.1. Selected gene identities and SABT coexpression values putatively involved in salicinoid biosynthesis.

Gene ID[1]	Gene name	Function	PCC
Potri.013G074500	SABT	Acyl transferase	1
Potri.016G014500	UGT71L1	Glycosyltransferase	0.71
Potri.003G066400	CYP91	Oxidation	0.71
Potri.006G171100	UGT78M1	Glycosyltransferase	0.69
Potri.001G031200	SCPL	Acyl transferase	0.65
Potri.017G028100[2]	CYP81K1	Oxidation	0.37

[1] SABT coexpression is presented as described by Fellenberg et al. (2020).
[2] Gene selection was conducted through analysis of differentially expressed genes as determined by RNA-Seq of MYB-165 overexpressing plants (Ma et al., 2018)

Coexpression of UGT71L1 and UGT78M1 with SABT suggests the enzymes may exhibit glucosylation activity on the SABT product, salicyl benzoate. Fellenberg et al. (2020) demonstrated that UGT71L1 and UGT78M1 had strong glucosylation activity on salicyl benzoate, with minor activity on salicylaldehyde and *ortho* hydroxylated cinnamic acid. Glucosylation of hydroxylated dibenzenoids can occur in plants outside Salicaceae such as in *Brickella veronicaefolia*: Asteraceae (Rivero-Cruz et al., 2005). The genus *Solidago* (Asteraceae) also generates glucosylated benzyl benzoate derivatives (Shiraiwa et al., 2012). Presence of glucosylated benzyl benzoate derivatives across at least two plant families and several genera demonstrates multiple instances for the adoption of dibenzenoid specialized metabolites in plants. The identification of biosynthetic genes for salicinoid and benzenoid metabolism can be aided through the systematic elimination of functional gene products.

In poplar, the CRISPR/Cas9 genome editing system was used to eliminate functional UGT71L1 in the hairy root cultures of *P. tremula x P. alba*. A near total elimination of the salicinoids salicortin, tremuloidin, and tremulacin was observed after the functional loss of UGT71L1. By contrast, the simple monophenolic salicinoid salicin was significantly reduced, but not eliminated, in the hairy root cultures (Fellenberg et al., 2020). Demonstrating the essential role of UGT71L1 for salicinoid biosynthesis represented the first clear identification of an enzyme in the salicinoid biosynthetic pathway.

1.2 Dissertation Outline

This dissertation aims to address several key aspects of salicinoid biosynthesis and biological roles of salicinoids in poplar. Primarily, I sought to expand the understanding of the in planta roles of the glycosyltransferase enzymes UGT78M1 and UGT71L1. This work also aims to identify intermediate molecules to further elucidate the salicinoid biosynthetic pathway. The current chapter (Chapter 1) has broadly presented plant specialized metabolism. It also provides crucial background information and preceding work that is the basis of the current understanding of salicinoid biosynthetic enzymes in poplar.

Functional knockouts of UGT71L1 in transgenic plants were generated using CRISPR/Cas9 genome editing prior to my work; however, these transformed plants had not been subject to extensive study. Therefore, Chapter 2 presents work to better understand the whole plant functions of UGT71L1. Specifically, the unexpected effects of disrupted UGT71L1 on plant growth and development were explored. High-throughput metabolomics using targeted and non-targeted analysis was used to search for chemical perturbations in UGT71L1 knockout poplar lines. I conducted the non-targeted metabolomic analysis in collaboration with Dr. David Overy at Agriculture and Agri-food Canada in Ottawa. I conducted targeted metabolomic analysis on phytohormones, amino-acids, and phenolic analytes in collaboration with Dr. Tobias Köllner, Dr. Nathalie Lackus, and Dr. Johnathan Gershenzon at the Max-Planck Institute for Chemical Ecology in Jena, Germany. I carried out all other experiments at the University of Victoria. Insect feeding preferences were assessed on the salicinoid deficient UGT71L1 knockout trees. Alterations in gene expression were examined in UGT71L1 knockout lines using RNA-sequencing. Furthermore, I generated a gRNA resistant version of UGT71L1 and retransformed this construct into the UGT71L1 knockout lines to restore gene function and rescue the mutant. This chapter establishes a relationship between salicinoid biosynthesis, poplar growth, and phytohormone homeostasis. UGT71L1 knockout poplars have significantly reduced salicinoid, flavonoid, and proanthocyanin concentrations. Disruption of UGT71L1 also resulted reduced plant height and a significantly altered leaf morphology. A slightly modified version of this chapter has been published in entirety in The Plant Cell (Gordon et al., 2022).

In Chapter 3 I administered deuterium-labelled benzenoids to wild-type and UGT71L1 knockout poplars to directly test the involvement of benzyl benzoate and salicyl benzoate in salicinoid biosynthesis. The deuterated benzenoids were administered to poplar leaf discs and leaf petioles. The metabolic fate of these exogenously applied molecules was monitored through the quantification of deuterium incorporation in downstream metabolites to assess the contribution of benzyl benzoate and salicyl benzoate to salicortin and salicin biosynthesis. I show direct evidence for the dibenzenoid benzyl benzoate as an intermediate in salicortin biosynthesis. The deuterium-labelled

substrates were prepared in the laboratory of Dr. Jeremy Wulff in the chemistry department at the University of Victoria.

In Chapter 4 I explored a potential role of salicinoids as carbon storage compounds using resprouting aspen trees as a test system. Using carbon starvation of wild-type and UGT71L1 knockout aspens, I asked if salicinoids can function as a carbon reserve. This work was conducted in collaboration with Dr. Simon Landhäusser and Dr. Rachel Hillabrand at the University of Alberta. Carbon limited resprouting carbohydrate analysis was conducted at the University of Alberta. I carried out the measurements of plant growth, quantification of salicinoid, and statistical analysis in the Constabel lab at the University of Victoria. My work in this chapter determined that salicinoids are an inaccessible carbon sink during carbon limited growth conditions, and that poplar maintains constitutive production of salicinoids under these conditions. I also identified different salicinoid profiles in shoot and root tissues of poplar plants. This work has been submitted for publication to the journal Tree Physiology (Hillabrand et al., 2022; Chapter 4).

In Chapter 5 I aimed to understand the in planta function of the glycosyltransferase gene UGT78M1 (Veljanovski and Constabel, 2013; Fellenberg et al., 2020). I used CRISPR/Cas9 genome editing to generate multiple independent knockout lines of UGT78M1 and multiple double knockout lines of UGT71L1/UGT78M1. Double knockout lines were created to investigate the possible contribution of UGT78M1 to the limited salicin biosynthesis that persists in UGT71L1 knockout lines. As part of this work, I have demonstrated the capacity for stacked CRISPR/Cas9 construct transformations in poplar. I found that contrary to expectations, UGT78M1 does not contribute significantly to salicinoid biosynthesis. However, disruption of UGT78M1 resulted in the unexpected accumulation of salicyl benzoate glucoside. Chapter 5 is written as a manuscript in progress. With additional experimental replication and mechanistic inquiry, this work will be submitted for publication.

Chapter 6 is a broad summary of my doctoral research. This chapter presents a summary of contributions I have made to the understanding of salicinoid biosynthesis and the unexpected connections that have been identified. Chapter 6 presents future directions for research into salicinoid biosynthesis and poplar chemical ecology. The appendices

provide additional experiments and research I have conducted thar are published or to contribute to future publications from the Constabel lab. Appendix 1 contains enzyme assay results on UGT71L1 activity for the publication of Fellenberg et al. (2020). Appendix 2 presents a plasmid diagram used for the generation of UGT71L1 overexpressing poplar lines. UGT71L1 overexpressing poplar lines were generated concurrently with UGT71L1 knockout rescue lines (Chapter 2). The concentrations of salicinoids in UGT71L1 overexpressing poplar are also presented. Appendix 3 presents a plasmid diagram and molecular biology methods used to create a GUS expression construct under control of the UGT71L1 native promoter. Appendix 4 presents the identification and analysis for the isomer specific elimination of trans-cinnamoyl salicortin in a transgenic poplar line. Appendix 5 presents preliminary data on the melatonin concentration in MYB165 overexpressing poplar lines. Appendix 6 presents preliminary enzyme assay chromatograms for recombinant UGT718M1 on novel hydroxylated dibenzenoid substrates.

Chapter 2 – CRISPR/Cas9 Disruption of UGT71L1 in Poplar Connects Salicinoids and Salicylic Acid Metabolism and Alters Growth and Morphology

This chapter has been published in a modified format for The Plant Cell (https://doi.org/10.1093/plcell/koac135; Gordon et al., 2022) Harley Gordon planned and carried out experiments, analyzed the data, and wrote the manuscript. Christin Fellenberg created transgenics, Harley Gordon and Nathalie Lackus performed the phytohormone, amino acids and targeted metabolite analysis, Finn Archinuk and Harley Gordon genotyped the transgenics and determined mutations, Harley Gordon and Amanda Sproule carried out non-targeted metabolomic analysis, Yoko Nakamura synthesized substrates and wrote the substrate synthesis method. Tobias Köllner helped with phytohormone, amino acids, and targeted phenolic analyses, David Overy planned and supervised non-targeted metabolomics and edited the manuscript. Peter Constabel conceived and planned experiments and co-wrote the manuscript.

2.1 Introduction

Poplars (*Populus* spp.) have been widely adopted as the model species for tree genomics, molecular biology, and physiology. A major advantage of *Populus* as an experimental system is its susceptibility to *Agrobacterium*-mediated genetic transformation (Brunner et al., 2004), and CRISPR/Cas9 generated knockouts have become powerful resources for understanding gene function in these trees (Elorriaga et al., 2018). Poplars contain abundant and diverse phenolic metabolites, which make them attractive plants for studies on ecological biochemistry. The most abundant phenolics found in poplar leaves are the proanthocyanidins (condensed tannins), hydroxycinnamate esters, and the phenolic glycosides known as salicinoids. Salicinoids are unique compounds built from glucosylated salicyl alcohol, typically further esterified with other organic acids including benzoate, *p*-coumarate, acetate, or a 6-hydroxy-2-cyclohexen-on-oyl (HCH) moiety. Salicinoids are found only in the Salicaceae family, composed of

poplars and willows (*Salix*). Approximately 35 salicinoid structures have been described from various *Salix* and *Populus* species (Boeckler et al., 2011; Keefover-Ring et al., 2014). *P. tremula* x *P. alba*, a hybrid commonly used in research for its ease of genetic transformation (Han et al., 2000), accumulates tremulacin, salicortin, salicin and tremuloidin as the most abundant salicinoids (Fig. 2.1). Due to their bioactivity, salicinoids are important in traditional medicine among many indigenous communities (Desborough and Keeling, 2017).

Salicinoids are abundant in vegetative tissues including leaves, where they can reach concentrations of over 15% dry weight (Donaldson et al., 2006). Complex salicinoids such as salicortin and tremulacin that contain the unstable HCH moiety are toxic when ingested and act as anti-herbivory compounds (Boeckler et al, 2011; Lindroth and Clair., 2013). High salicinoid concentrations correlate with poor performance of gypsy moth (*Lymantria dispar*) and forest tent caterpillar larvae (*Malacosoma disstria*) (Hwang and Lindroth, 1997). Salicinoids also have an inhibitory effect on mammalian herbivory (Tahvanainen et al., 1985; Basey et al., 1988). Unlike many other defensive specilaized metabolites, herbivory or tissue damage does not appreciably induce accumulation of salicinoids and they are considered constitutive defense compounds (Boeckler et al., 2011; 2013).

Despite the ecological and potential medical importance of the salicinoids, their biosynthesis remains largely unelucidated. Early radiolabeling experiments by Zenk (1967) in *Salix purpurea* demonstrated that salicin is derived from cinnamic acid or benzoic acid via salicylaldehyde. Similar experiments to probe the biosynthesis of more complex salicortin were reported by Babst et al. (2010), who determined that benzoic acid and benzaldehyde are readily incorporated into both salicyl alcohol and HCH moieties of salicortin, but benzyl alcohol contributes exclusively to the salicyl alcohol moiety. By contrast, the labelling data indicate that free salicyl alcohol or salicylaldehyde are not intermediates in the biosynthesis of salicortin (Babst et al., 2010). Rather, salicyl alcohol and salicylaldehyde are incorporated into salicin, suggesting that the biosynthetic routes for salicin and salicortin are at least partly distinct. However, salicylic acid (SA) is not a precursor for salicinoids. Babst et al. (2010) also proposed benzyl benzoate as a potential intermediate in salicinoid biosynthesis, and enzymes for the condensation of

benzyl alcohol and benzoyl-CoA to benzyl benzoate had been previously identified in *Petunia hybrida* and *Nicotiana benthamiana*. Furthermore, the conversion of cinnamate to benzaldehyde, benzyl alcohol, or benzoic acid via cinnamoyl-CoA-dependent and -independent pathways has been demonstrated (Widhalm and Dudareva, 2015). In poplar, benzyl benzoate and other benzenoid leaf volatiles are induced by herbivory and are synthesized from cinnamoyl-CoA by ß-oxidative pathway enzymes (Lackus et al., 2021); however, the pathway for these volatiles appears to be separate from the biosynthesis of salicinoids. Details of benzenoid biosynthesis relevant for salicinoid biosynthesis remain unknown but are actively under investigation.

The idea of potential involvement of benzyl benzoate or salicyl benzoate in salicinoid synthesis was strengthened by the identification of poplar benzoyl-CoA: benzyl alcohol O-benzoyltransferase (BEBT) and benzoyl-CoA: salicyl alcohol O-benzoyltransferase (SABT), two acyl transferases that specifically synthesize benzyl benzoate and salicyl benzoate, respectively (Chedgy et al., 2015). The involvement of these potential intermediates was further corroborated when we identified two UDP-glucosyl transferases, UGT71L1 and UGT78M1, which are co-expressed with SABT and show high specificity for salicyl benzoate as the glucose acceptor (Fellenberg et al., 2020). BEBT, SABT and the UGTs thus provide the outlines of a pathway to salicortin and other complex salicinoids via benzyl benzoate, salicyl benzoate and its glucoside (Fig. 2.1). We demonstrated the central importance of salicyl benzoate and UGT71L1 using CRISPR/Cas9 knock-out hairy root lines of UGT71L1, which are unable to make salicortin, tremulacin, and tremuloidin. These knockout lines did retain about 50% of salicin content, however, confirming the existence of an alternate biosynthetic route to salicin (Babst et al., 2010). Similar results have been obtained using RNAi suppression of UGT71L1 carried out independently (Kulasekaran et al., 2020). How salicyl benzoate glucoside is further oxidized to give rise to the HCH moiety is not known, but likely begins with hydroxylation of the aromatic ring. To date, no other enzymes or intermediates of the pathway have been identified.

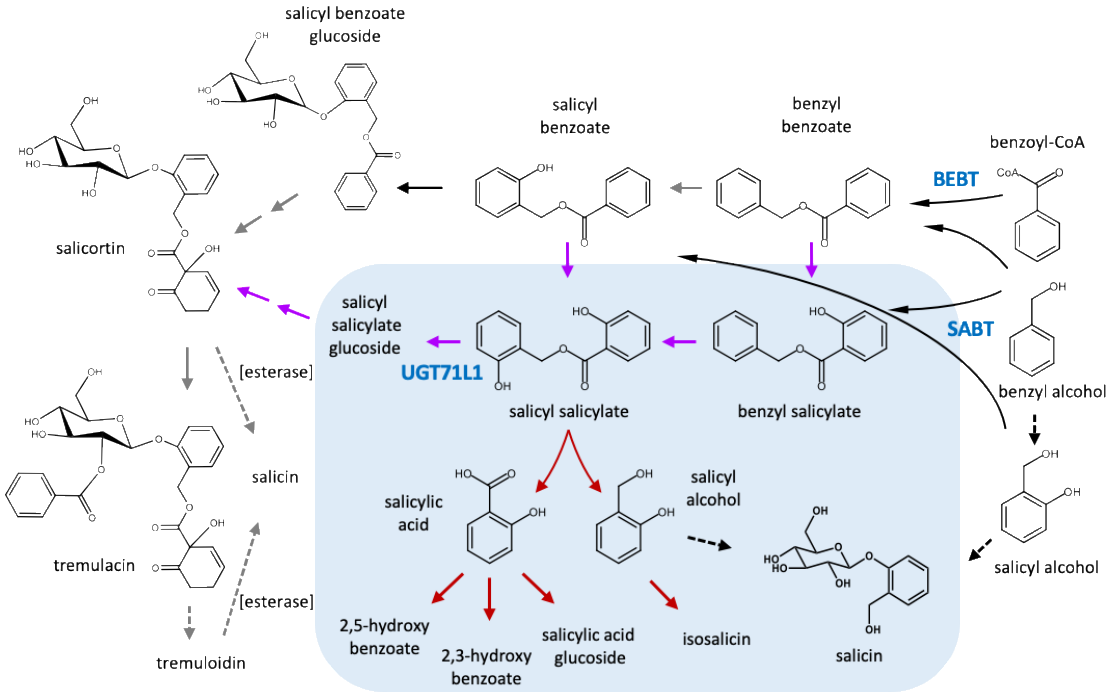


Figure 2.1. Proposed biosynthetic pathway for salicinoids based on prior work in various Populus or Salix species and the current experiments. Alternate entry points starting with BEBT (benzoyl-CoA: benzyl alcohol O-benzoyltransferase) and SABT (benzoyl-CoA: salicyl alcohol O-benzoyltransferase) are shown. Solid black arrows and blue font indicate established enzymatic reactions (Chedgy et al., 2015; Fellenberg et al., 2020; Kulasekaran et al., 2020); dashed black arrows are proposed biosynthetic steps based on published in vivo labeling experiments (Zenk, 1967; Babst et al., 2010); dashed gray arrows represent known non-enzymatic or enzymatic decomposition steps (Julkunen-Tiitto and Sorsa, 2001; Ruuhola et al., 2003); gray arrows indicate hypothesized biosynthetic reactions. Compounds within the blue box and purple arrows indicate proposed biosynthetic steps and products based on the current work. Red arrows represent additional reactions deduced to be enhanced in UGT71L1-KO plants based on metabolomic analysis in this study.

The presence of salicinoids in the Salicaceae family is notable because willow bark was the source of the original identification of salicylic acid (SA) as an active analgesic in the 19th century (Ding and Ding, 2020). The source of SA in willow and poplar extracts is now known to be the salicinoids, which decompose spontaneously or enzymatically (Ruuhola et al., 2003). However, SA itself is not a direct precursor of salicinoids. Although SA has been established as a bioactive plant phenolic for over a century, its function as a stress hormone was not discovered until 1979, when its role in inducing resistance against viruses and as an inducer of plant defense genes was described (Raskin, 1992). How SA acts as a signal for plant defense has been intensively studied since then, leading to detailed current knowledge of NPR1 and NPR3/NPR4 SA receptors and signalling pathways (Ding et al., 2018; Innes, 2018; Liu et al., 2020). Following pathogen detection, SA levels rise and signal a transcriptional response leading to the local expression of defense genes as well as triggering secondary signals that activate a systemic defense response (Ding and Ding, 2020). Treatment of Arabidopsis plants with SA alters the expression of thousands of genes, upregulating defense as well as down-regulating growth-related genes (Van Butselaar and Van den Ackerveken, 2020). SA is involved in regulating senescence, modulating effects of other phytohormones, and controlling plant growth within a defense-growth trade-off (Zhang et al., 2013; Liu et al., 2016; Ding and Ding, 2020).

Detailed investigations in Arabidopsis and *N. benthamiana* have determined that SA is synthesized through distinct plastidic and cytosolic pathways. In the plastidic pathway, SA is derived from isochorismate via chorismate in the shikimate and aromatic amino acid pathway. Isochorismate is exported into the cytosol and converted to isochorismate-9-glutamate, which then spontaneously or enzymatically decomposes to release SA (Rekhter et al., 2019). In the cytosolic SA pathway, SA is derived from *trans*-cinnamic acid produced by phenylalanine ammonia lyase (PAL) of the phenylpropanoid pathway (Ryan et al., 1995; Zhang and Li, 2019). In Arabidopsis, the isochorismate pathway is responsible for the majority of pathogen-induced SA accumulation (Wildermuth et al., 2001). By contrast, in soybean, isochorismate- and phenylpropanoid-derived SA contribute equally to pathogen-induced SA accumulation (Shine et al., 2016). SA biosynthesis and the relative importance of the isochorismate-derived SA and

phenylpropanoid derived SA in poplar has not been investigated. The importance of SA in defense against biotrophic rust fungi in poplar has been recently demonstrated. (Ullah et al., 2019).

To obtain additional insight into salicinoid biosynthesis, we generated transgenic CRISPR/Cas9 UGT71L1 knock-out poplars. While hairy root cultures had provided initial proof of function of UGT71L1, rooted plants grown under greenhouse conditions accumulate higher levels of phenolics than tissue cultures (Fellenberg et al., unpublished; Chapter 4) and are thus more likely to show metabolic connections. Furthermore, greenhouse grown plants expand the capacity for testing the effect of alterations in environmental conditions and provide ample tissue. Our objective was to use non-targeted and targeted metabolomics to identify changes in phenolic profiles caused by loss of UGT71L1 to identify potential intermediates in salicinoid biosynthesis. Our experiments confirmed the central in planta role of UGT71L1 in the salicinoid pathway, but also indicate that additional UGT enzymes may be play minor roles. Surprisingly, loss of UGT71L1 in whole poplar plants manifested novel phenotypic effects including leaf curling and reduced expansion and growth. Additional experiments suggested that the effects on growth could be due to SA release from the disrupted salicinoid pathway, thus connecting a specialized metabolic pathway to growth and development.

2.2 Materials and Methods

2.2.1 Poplar transformation and growth

CRISPR/Cas9 UGT71L1 knockout constructs were created using the p201N vector (Jacobs et al., 2015) containing the previously described guide RNA sequence for UGT71L1 (Fellenberg et al., 2020). P201N:gRNA was moved into *Agrobacterium tumefaciens* strain GV3101 by electroporation. Leaf explants from in vitro *P. tremula* x *P. alba* (INRA 717-1B4) leaves were excised and used for transformation experiments as previously described (James et al., 2017). Positive transformants were selected on media containing kanamycin. UGT71L1 mutations were screened using the TIDE Webtool (Brinkman et al., 2014) with PCR amplification products of the gRNA target site (Fig.

2.2). Mutations were confirmed in both tremula and alba alleles by Sanger sequencing of the PCR product. Three independent empty-vector control lines expressing the Cas9 protein with no gRNA were also generated. Confirmed transgenic and mutant lines were propagated as in vitro cultures, and when rooted were transferred to soil in two-inch pots and grown in a mist chamber for three-weeks. Following 3 weeks in the mist chamber, plants were transferred to one-gallon pots containing peat-moss based soilless mix (Sunshine Mix 4, Sungro, Seba Beach, AB, Canada). Soil was supplemented with slow-release fertilizer as described by Major and Constabel (2006). Plants were grown in a greenhouse under ambient light conditions, with supplemental lighting provided by 600 W high pressure sodium lights with an average intensity of 300 $\mu mol \cdot m^{-2} s^{-1}$ for a 16h photoperiod. Plants were irrigated as required and rotated within the greenhouse weekly. For coronatine and salicylic acid treatments, in vitro plantlets were removed from tissue culture, potted in rooting media, and moved into a greenhouse mist chamber. Following one week of acclimation, coronatine in water was applied via handheld spray bottles. Coronatine solution (10 µM or 25 µM) was applied as a mist until droplets formed. For SA treatments solutions of SA in water were prepared at 1 mM, 100 µM, and 10 µM concentrations. SA solutions were adjusted to pH 6 with 0.2 M KOH and 0.1 % (v/v) Tween-20 (Fisher Scientific, Ottawa, ON, Canada) was added. Plantlets were maintained in the mist chamber for the duration of the experiment. The mist chamber was under ambient daylight with supplemental lighting from 400W high pressure sodium lights with an average intensity of 75 $\mu mol \cdot m^{-2} \cdot s^{-1}$ for a 16 h photoperiod. Plants were rotated periodically, and physical measurements were taken weekly throughout the experiment.

2.2.2 *Orygia leucostigma* feeding assays

Small leaves (< 4 cm long) from 5-month-old *P. tremula x P. alba* saplings were harvested and used in choice feeding assays with *O. leucostigma*. Insect eggs and diet were obtained from Insect Production Services of the Great Lakes Forestry Centre (Natural Resources Canada, Sault Ste. Marie, ON, Canada). Single 3rd or 4th instar larvae were placed in small petri dish arenas for 30 minutes prior to the addition of a pre-weighed leaf from both a UGT71L1-KO plant and an empty vector control plant. Feeding

assays were carried out in a darkened room for 20 hours. Following feeding, individual leaves were re-weighed to determine the weight eaten per leaf. Control dishes with no caterpillars were used to correct for water loss. Only dishes where caterpillars had fed were used for data collection. Data were recorded as weight of foliage consumed.

2.2.3 Non-targeted metabolomic analysis

For all chemical analysis of greenhouse experiments, biological replication consisted of four or five individual copies (clonal greenhouse plants) for each transgenic or control line. Typically, three or more independent transgenic lines were included in the experiment. Leaves of three-month-old greenhouse-grown poplar were numbered, with Leaf #1 designated as the youngest leaf with a lamina breadth of > 1 cm. Expanded leaves #10-13 were harvested, the midvein excised, and the leaf blades frozen in liquid nitrogen. Samples were stored at -80 °C prior to extraction and analysis. Samples were homogenized under liquid nitrogen using a ceramic mortar and pestle. Homogenized samples were freeze-dried and stored at room temperature. For extraction, 50 mg of freeze-dried tissue was weighed into a 2 mL cryo-vial with 4 x 5 mm steel ball bearings and 750 mL of chilled methanol. Samples were homogenized 2 x 45 seconds at 5500 rpm using a Precellys 24 Homogenizer, then sonicated for 2 minutes in a sonicating water bath. Samples were centrifuged at 13,000 *g* for 2 minutes and the supernatant was filtered through 0.22 µm PTFE filters into glass vials. Methanol extractions were repeated twice, and extracts were pooled. Extracts were dried under vacuum at ambient temperature. Final extraction weight was determined, vials were sealed and stored at -20 °C until analysis. Each biological replicate included an extraction blank. For analysis, samples were resuspended to 500 $\mu g \cdot mL^{-1}$ in methanol and stored briefly at 4 °C prior to injection.

Samples were analyzed in a randomized injection order using a Thermo Ultimate 3000 (ultrahigh–performance liquid chromatography (UPLC)) coupled to Thermo LTQ Orbitrap XL (high resolution mass spectrometer (HRMS)) and UltiMate Corona VeoRS charged aerosol detector (Thermo Fisher Scientific Inc, Waltham, MS, USA). Samples were separated using a Phenomenex Kinetex 1.7 µm C18 100 Å 50 x 2.1 mm column

with a SecurityGuard ULTRA C18 guard cartridge. For positive ionization mode samples were run with a continuous flow rate of 0.35 $mL \cdot min^{-1}$. Solvent A was LC/MS grade water with 5 mM ammonium acetate and 0.1% formic acid; solvent B was acetonitrile with 0.1% formic acid (v/v). Separation occurred on a gradient of 2% B to 45% B for 10 minutes; followed by 45% B to 100% B over 5 minutes and then 100% B to 2% B over 3 minutes for a total run time of 18 minutes. The electrospray ionization (ESI) source had a capillary voltage of 35 V. Negative ionization mode samples were run under the same separation gradient, the ESI source had a capillary voltage of -35 V. Both modes collected m/z from 100 – 2000.00. For MS/MS collection alternate scans were fragmented from the previous scan with a m/z range of 300 - 2000.00. MS/MS was conducted on representative samples for the UGT71L1-KO and control lines.

Raw MS1 data was processed using MZMine2 (Pluskal et al., 2010). Peaks were filtered for duplicates and then peak area and height data was exported to Microsoft Excel. Mass features were processed, and any mass features present in extraction and solvent blanks were eliminated from the data set. Peaks were also removed if less than 50% of a sample group (UGT71L1-KO, partial-KO, and control) were below the noise threshold. MS2 data was processed through MZMine2 (Pluskal et al., 2010) as previously described, except that MS1 data was paired with MS2 fragmentation data and exported to SIRIUS 4 utilizing CSI:fingerID (Dührkop et al., 2015; Dührkop et al., 2019) was used to query molecular structure databases assign putative structural identities.

2.2.4 Targeted metabolomic, hormone, and amino acid analyses

For targeted metabolomics, phytohormone and amino acid analysis tissue was prepared and freeze-dried as per the non-targeted metabolomic methodology. Pooled tissue harvested from leaves # 4-8 was analyzed. Ten mg of freeze dried material was weighed into deep 96 well plates and 1 mL of methanol containing internal standards (40 $ng \cdot ml^{-1}$ D_6-abscisic acid (D6-ABA), 40 $ng \cdot ml^{-1}$ D_6-jasmonic acid (D6-JA), D_4-salicylic acid, 8 $ng \cdot ml^{-1}$ D_6-jasmonic-acid isoleucine (D_6-JA-Ile) (Santa Cruz Biotechnology, Dallas, TX, USA), 40 $ng \cdot ml^{-1}$ D_5-indolacetic acid (D5-IAA, OlChemIm s.r.o., Olomouc, Czech Republic)) was added. Samples were shaken vigorously for 30 s and then shaken

at 200 rpm for 30 min. Samples were centrifuged at 1,400 *g* for 2 minutes and 50 µL of supernatant was aliquoted into 96 well plate for amino acid analysis and 400 µL of supernatant was aliquoted into 96 well plates for targeted metabolomics and phytohormone analysis. All samples were stored at -20 °C prior to analysis. For amino acid analysis 450 µL of MilliQ water containing 10 µg amino acids was added to each sample.

Targeted metabolomics - Chromatographic separation was achieved using an Agilent 1260 infinity II LC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a Zorbax Eclipse XDB-C18 column (50 × 4.6 mm, 1.8 µm, Agilent Technologies), using aqueous formic acid (0.05% (v/v)) and acetonitrile as mobile phases A and B, respectively. The mobile phase flow rate was 1.1 $ml \cdot min^{-1}$. The elution profile is shown in Supplemental Table S2.5 as gradient A. The column temperature was maintained at 20 °C. The LC system was coupled to a QTRAP 6500® tandem mass spectrometer (Sciex, Darmstadt, Germany) equipped with a turbospray ion source, operated in negative ionization mode. The ion spray voltage was maintained at -4500 eV and the turbo gas temperature was set at 600 °C. Nebulizing gas was set at 60 psi, curtain gas at 45 psi, heating gas at 60 psi, and collision gas at medium level. Multiple reaction monitoring (MRM) was used to monitor analyte parent ion → product ion formation for each analyte as displayed in Supplemental Table S2.6. Identification of compounds were performed using authentic standards. Data acquisition and processing was performed using Analyst 1.6.3 (Sciex) and MultiQuant 3.0.3 (Sciex) software.

Phytohormone analysis- Analysis was conducted as described previously in Irmisch et al. (2014). In brief, an Agilent 1260 infinity II LC system (Agilent Technologies) coupled to a QTRAP 6500® tandem mass spectrometer (Sciex) was used for the analysis. Chromatographic separation was achieved using a Zorbax Eclipse XDB-C18 column (50 × 4.6 mm, 1.8 µm, Agilent Technologies), and aqueous formic acid (0.05% (v/v)) and acetonitrile as mobile phases A and B, respectively. The mobile phase flow rate was 1.1 $ml \cdot min^{-1}$. The elution profile is listed in Supplemental Table S2.5 as gradient B. The tandem mass spectrometer was equipped with a turbospray ion source, operated in negative ionization mode. The ion spray voltage was maintained at -4500 eV and the turbo gas temperature was set at 650 °C. Nebulizing gas was set at 60 psi, curtain

gas at 40 psi, heating gas at 60 psi, and collision gas at medium level. Indoleacetic acid was quantified using the same LC-MS/MS system with the same chromatographic conditions but using positive mode ionization with an ion spray voltage at 5500 eV. Multiple reaction monitoring (MRM) was used to monitor analyte parent ion → product ion formation for each analyte as displayed in Supplemental Table S2.6. Data acquisition and processing was performed using Analyst 1.6.3 (Sciex) and MultiQuant 3.0.3 (Sciex) software. SA, ABA, JA, and JA-Ile were quantified using their deuterated internal standard. SA-glucoside was quantified using the SA internal standard. Sulfo-JA, OH-JA, JA-glucoside, and OPDA were quantified using the deuterated JA internal standard and an experimental determined response factor (Irmisch et al., 2014). OH-JA-Ile and COOH-JA-Ile were quantified using deuterated JA-Ile as an internal standard.

Amino acid analysis- Chromatography was performed on an Agilent 1200 HPLC system, equipped with a Zorbax Eclipse XDB-C18 column (50 x 4.6 mm, 1.8 µm; Agilent Technologies). Aqueous formic acid (0.05% (v/v)) and acetonitrile were used as mobile phases A and B, respectively, and the elution profile is listed in Supplemental Table S2.5 as gradient C. The mobile phase flow rate was 1.1 $ml \cdot min^{-1}$, and the column temperature was maintained at 25 °C. The liquid chromatograph was coupled to a QTRAP 6500® tandem mass spectrometer (Sciex, Darmstadt, Germany) equipped with a turbospray ion source, operated in positive ionization mode (ion spray voltage, 5500 eV; turbo gas temperature, 650 °C; nebulizing gas, 70 psi.; curtain gas, 40 psi.; heating gas, 70 psi.; collision gas at medium level). MRM were chosen as described by Jander et al. (2004), except for Arg and Lys, and are listed in Supplemental Table S2.6. Analyst 1.5 software (Applied Biosystems) was used for data acquisition and processing. Data acquisition and processing was performed using Analyst 1.6.3 (Sciex) and MultiQuant 3.0.3 software (Sciex).

2.2.5 Lignin Extraction and Acetyl Bromide Quantification

Lignin quantification and extraction was based on the protocols presented by Moreira-Vilar et al. (2014) with some modifications. Leaves were collected from four-month-old greenhouse grown *P. tremula x P. alba* plants. Leaf midveins were excised

and material was dried for 36 hours at 60 °C. Dried leaves were lightly ground in a mortar and pestle prior to extraction. 300 mg of dried leaf was homogenized using a mortar and pestle in 7 mL of potassium phosphate buffer (pH 7, 50 mM), tissue was then centrifuged at 1,700 *g* for six minutes. Two additional washes were conducted using 7 mL of potassium phosphate buffer each wash and included insoluble material resuspension via vortexing. Three washes were conducted using 7 mL of potassium phosphate buffer containing 1% (v/v) of Triton X-100. The insoluble portion was then washed twice with 7 mL of 1 M NaCl. Insoluble portion was then washed twice with 7 mL of Milli-Q grade water, followed by four washes with 7 mL of analytical grade acetone. This protein free cell wall (PFCW) component was then dried overnight at 60 °C.

20 mg of PFCW was then dissolved in 0.5 mL of 1:3 acetyl bromide:glacial acetic acid and incubated at 70 °C for 30 minutes. Samples were cooled and 0.9 mL of 2 M NaOH, 5 mL of glacial acetic acid, and 0.1 mL of 5 M hydroxylamine-HCl were added. Samples were centrifuged at 3,000 *g* for 10 minutes and the supernatant collected. Samples were assessed in a 96 well plate measured at 280 nm, a standard curve was prepared using alkali lignin. Results were expressed as mg Lignin·g^{-1} PFCW and presented as percent lignin in the PFCW.

2.2.6 Generation and analysis of UGT71L1-KO rescue plants

A modified UGT71L1 without the Cas9 gRNA target sites used to create the original UGT71L1-KO transgenics was designed based on the validated UGT71L1 coding sequence of *P. tremula x P. alba* (INRA 717-1B4) genome v2 (http://aspendb.uga.edu/databases/spta-717-genome) and synthesized in a PBK+ vector (BioMatik, Kitchener, ON, Canada). Six nucleotide mismatches were introduced into degenerate codon positions into the UGT71L1 gRNA target site. The modified UGT71L1 sequence (UGT71L1m) was cloned into a pMDC32 vector (Curtis and Grossniklaus, 2003) using added 5’ KpnI and 3’ Spe1 restriction sites. PMDC32:UGT71L1m was moved into *Agrobacterium tumefaciens* GV3103 via electroporation. One UGT71L1-KO transgenic line was re-transformed with this construct using previously described poplar

transformation methods (James et al., 2017). Successfully transformed calli were sequentially selected on woody plant medium (Caisson) containing hygromycin (10 $\mu g \cdot mL^{-1}$). Successful transformants were rooted in vitro and acclimated to greenhouse growth as described above. Wild-type *P. tremula x P. alba* was also transformed with the rescue construct under the expression of a CAM35S promoter. The WT lines containing the rescue construct were labelled as an overexpressor and data regarding their phytochemical screening is presented in Appendix 2. Overexpressor plants were phytochemically analyzed as described below.

For chemical analysis, expanded leaves of two-month-old rescue plants were flash frozen and homogenized in a mortar and pestle in liquid nitrogen. Approximately 100 mg of frozen powder was extracted into 1 mL of methanol, sonicated in a bath sonicator for 5 min and shaken on a benchtop orbital shaker for 3 min. Extracts were centrifuged for 10 minutes at 20,000 *g* and the supernatant was collected and stored at -20 ºC prior to phenolic analysis using a Waters Acquity UPLC system equipped with QDa MS detector (Waters, Mississauga, ON, Canada). 2 µL sample injections were separated on a Waters Acquity UPLC Ethylene Bridged Hybrid column (C18, 2.1 x 50 mm, pore size 1.7 µm) and a binary solvent gradient comprising ultra-pure water containing 0.1% formate (A) and LC/MS grade acetonitrile with 0.1% (v/v) formate (B). The flow rate was 0.5 $mL \cdot min^{-1}$. Separation gradient was 99.9% A, 0-0.5 min; 99.9-80.0 % A, 0.5-6.0 min; 80.0-50.0% A, 6.0-9.0 min; 50.0-10.0% A, 9.0-10.0 min; 10% A, 10.0-11.5 min; 10-99.9% A, 11.5 – 12.0 min; 99.9% A, 12.0-13.0 min. The QDa detector using electrospray ionization in negative mode was used to detect and quantify salicortin, tremulacin, tremuloidin and salicin. Salicin formate adduct at m/z = 331.0 with cone voltage (V) of 10 was detected at 2.9 min. Salicortin (-H) at m/z = 423.0 with V = 10 was detected at 5.6 min. Tremulacin (-H) at m/z = 537.0 with V = 10 was detected at 8.4 min. Tremulacin was quantified using linear regression of authentic standard peak areas provided by Dr. Richard Lindroth. Salicortin was quantified using linear regression of an authentic standard (Biosynth, San Diego, CA, USA). Salicin was quantified using commercially available standards from Sigma-Aldrich (Oakville, ON, Canada). Peak integration was conducted using MassLynx 4.2 (Waters).

2.2.7 RNA-Sequencing and analysis

RNA was extracted as per Yoshida et al. (2015) from four clonal plants (biological replicates) for one UGT71L1-KO, the partial KO, and the wild-type control line. Isolation and enrichment of mRNA was conducted using a MagJET enrichment kit (ThermoScientific, Waltham, MA, USA). mRNA was fragmented at 94 °C for 13 minutes to achieve an approximate fragment size of 350 bp. cDNA libraries were prepared as per Ma et al. (2018) using commercially available kits for Illumina sequencing. Twelve libraries were sequenced by the University of British Columbia Sequencing + Bioinformatics Consortium (Vancouver, BC, Canada) on one NextSeq High Output Lane. Sequencing results were de-multiplexed, adaptor sequences were trimmed, and raw data was aligned to the *P. tremula* transcriptome V.2 (Xue et al., 2015) using HiSat2 (Kim, et al., 2019). Cufflinks (Trapnell et al., 2012) was used to analyze differentially expressed genes and DeSeq2 package in R was used to generate the final differentially expressed gene table. Gene annotation was done using the annotated transcriptome in PopGenie (popgenie.org; Sjödin et al., 2009). Enrichment Analysis was conducted using highly expressed genes in WT and UGT71L1-KO samples and the PoplarGene database (Liu et al., 2016). Metabolic pathway expression of identified transcripts was assessed using POPLARCYC 3.0 (Schläpfer et al., 2017).

2.2.8 Synthesis of salicyl salicylate

2-hydroxy benzoyl chloride was prepared from salicylic acid (Sigma-Aldrich) according to Zhao et al. (2018). To a mixture of salicyl alcohol (248 mg, 2 mmol, Merck KGaA) and 2-hydroxy benzoyl chloride (368 mg, 2 mmol) in anhydrous CH_2Cl_2 (10 mL), dry triethylamine (280 μL, 2 mmol, Sigma-Aldrich) was added at 0 °C. The mixture was stirred for overnight at room temperature under Argon, then CH_2Cl_2 and saturated NH_4Cl (aqueous) were added. The mixture was extracted twice with CH_2Cl_2. The organic phase was washed with brine, dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified twice by silica gel flash column chromatography (*n*-hexane: EtOAc gradient, 96:4- 66:34, v/v) to give salicyl salicylate (2-hydroxybenzyl 2-

hydroxybenzoate) (50 mg, 0.2 mmol, 10%). NMR spectra (^{1}H and ^{13}C NMR, ^{1}H-^{1}H COSY, ^{1}H-^{13}C HSQC and ^{1}H-^{13}C HMBC) of the product were measured on a Bruker Avance III HD 500 NMR spectrometer, operating at 500.13 MHz for ^{1}H and 125.75 MHz for ^{13}C. The spectrometer was equipped with a TCI cryoprobe (5 mm for Avance III HD 500). Spectrometer control and data processing was accomplished using Bruker Topspin ver. 3.6.1. All NMR measurements were recorded at 25 °C in $CDCl_3$. Chemical shifts were referenced to the residual solvent signals for measurements in $CDCl_3$ (δ_H 7.26/δ_C 77.16). HRMS was carried out on a Bruker Compact ESI-OTOF spectrometer (Bruker Daltonics GmbH, Bremen, Germany) using ESI ionization in the negative mode. For ^{1}H and ^{13}C NMR data, see Supplemental Table S2.7; HRMS (ESI, negative mode) *m/z* $[M-H]^-$: calculated for $C_{14}H_{11}O_4$, 243.0663; found, 243.0668.

2.2.9 UGT enzyme assays

Recombinant UGT71L1 and UGT78M1 were produced and purified as per Fellenberg et al. (2020). Enzyme assays were incubated for 3 hours at 30 °C, in a 200 µL reaction volume in 50 mM sodium phosphate (pH 7.4). Reactions contained UDP-glucose (5 mM), substrates at a final concentration of 250 µM and 5 µg of purified recombinant protein. Reactions were halted after 3 h with the addition of 20 µL 30% trichloroacetic acid (w/v), and 6 min of centrifugation at 16,100 *g*. Supernatant was collected and stored at -20 °C prior to analysis. Analysis and product identification of enzyme reaction products was conducted on a Waters Acquity UPLC LC/MS system. For analysis, an 8 µL sample injection was separated using a Waters Acquity Ethylene Bridged Hybrid (C18, pore size: 1.7 µm, dimensions: 2.1 x 50 mm) column, and glucoside products were identified based on mass and retention time. Product peak areas were determined using UV absorbance at 280 nm. Compounds were separated on a gradient binary consisting of ultra-pure water containing 0.1% formate (solvent A) and LC-MS grade acetonitrile with 0.1% formate (solvent B). The flow rate was 0.5 $mL \cdot min^{-1}$. The separation gradient was 90.0% A, 0-0.5 min; 90.0-10.0% A, 0.5-6.0 min; 10.0-0.5% A, 6.0-7.0 min; 0.5-10.0% A, 7.0-8.0 min; 10.0-90.0% A, 8.0-8.5 min. A single quadrupole MS in positive ionization mode (ES+) continually scanned masses 50.0 to

1000.0 m/z from 0.2 - 8.5 min. Cone voltage was set at 35 V. Sodium adducts of benzyl salicylate glucoside and salicyl benzoate glucoside were identified at 413 m/z. The sodium adduct of salicyl salicylate glucoside was identified at 429 m/z.

2.2.10 Statistical analysis

Multivariate and univariate statistics for metabolomic samples were processed using a combination of Rstudio (RStudio Team, 2018) and the MUMA package (Gaude et al., 2012), as well as MetaboAnalyst (Chong et al., 2019). Statistical processing of growth data, insect feeding assays was conducted using RStudio (RStudio Team, 2018).

2.3 Results

2.3.1 CRISPR/Cas9 generates diverse biallelic mutations in the *UGT71L1* gene in transgenic *P. tremula x P. alba* plants

We generated independent *P. tremula x P. alba* UGT71L1 knock-out lines using CRISPR/Cas9 and the previously utilized gRNA sequence (Fellenberg et al. 2020). Preliminary analysis of regenerating *in vitro* UGT71L1 plantlets confirmed very low concentrations of tremulacin, tremuloidin, and salicortin compared to wild-type plants. Transgenic lines with low salicinoid accumulation were sequenced at the UGT71L1 gRNA site, and four lines carrying mutations in both UGT71L1 alleles were studied further. Mutations in these lines generally consisted of small indels with a mix of single base-pair (bp) insertions, single bp deletions, and slightly larger deletions (Fig. 2.2). We also observed a large deletion in one allele of line UGT71-3. Large deletions have previously been reported in CRISPR/Cas9 knock-outs in plants utilizing non-homologous end joining (Shen et al., 2017). In addition, one allele in transgenic line UGT71-4 contained an in-frame 6 bp deletion and thus encoded a functional gene product. This transgenic line was considered a partial knock-out and could be expected to show an intermediate phenotype. To guard against non-target gene mutations, we PCR-amplified

and sequenced the five genomic regions most similar to the target gRNA. No off-target mutations were detected (Supplemental Table T2.1).

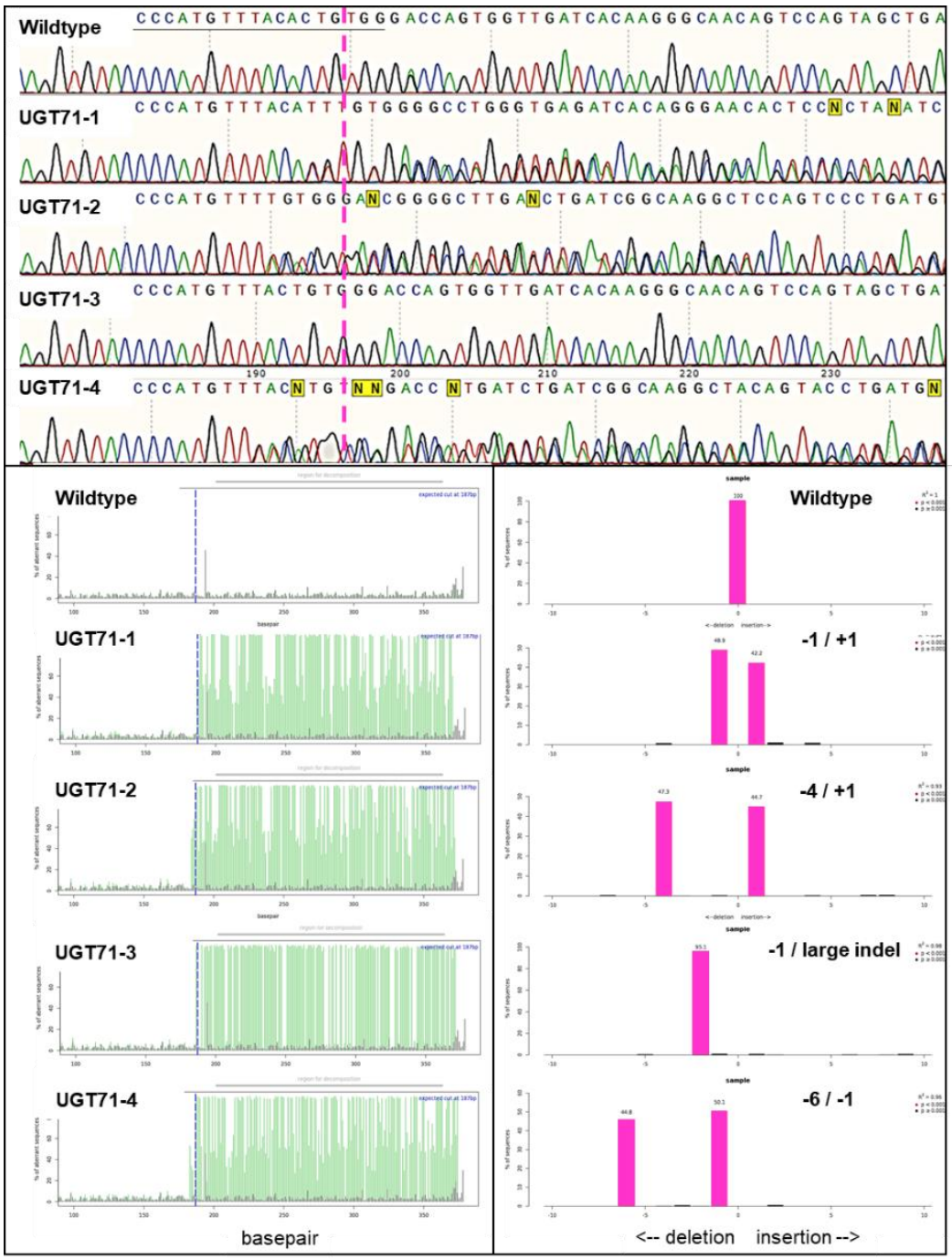


Figure 2.2. Detection of biallelic mutations and indels four independent UGT71L1-knock-out lines. Upper panel shows Sanger sequencing traces of the *UGT71L1* gRNA target site. The gRNA is underlined, and the magenta dashed line marks the Protospacer Adjacent Motif (PAM) necessary for Cas9 double-strand cleavage. N indicates ambiguous bases in the sequence readout. The lower panels show read-outs of the TIDE (shinyapps.datacurators.nl/tide/) webtool used to detect mutations in the target sequence. The vertical blue dashed line indicates the position of the gRNA PAM site, while the vertical green traces indicate variant sequences in chromatograms indicating a mutation in the downstream sequence. Lower right panels show the number of nucleotides inserted or deleted along the horizontal axis for each mutated line. Lines UGT71-1 to UGT71-3 are biallelic knock-outs with frameshifts in both alleles. Line UGT71-4 has mutations in both alleles but only one results in a frameshift. This line is designated as partial-KO.

2.3.2 Metabolomic analysis of UGT71L1 knockout poplars demonstrates a substantial reduction in salicinoid content and a major shift in phenolic metabolites

Both non-targeted and targeted metabolomic approaches were used to investigate the impact of eliminating UGT71L1 activity on poplar secondary metabolism. Methanolic leaf extracts from greenhouse-grown saplings were first profiled by ultra-high pressure liquid chromatography – high resolution mass spectrometry (UPLC-HRMS). The experiment consisted of the replicated UGT71L1 biallelic and monoallelic knock-out plants along with three independent empty vector lines and wild-type controls. Unsupervised multivariate data analyses revealed trends within this data set. Principal component analysis of non-targeted metabolomic data clearly separated the UGT71L1-KO mutant from the UGT71L1 partial-KO (monoallelic) and control lines (Fig. 2.3) along principal component 1 explaining 93.4% of the variance in the data model. Results from hierarchical cluster analysis indicated no difference between the empty vector and wild-type lines, and supported combining both into a single control group (Fig. 2.3B). These untargeted analyses demonstrated the strong effect of disrupting UGT71L1; it also

showed the quantitative effects of the gene on the metabolome, since the monoallelic mutant line was grouped by principal component analysis as an intermediate phenotype. Overall, the non-targeted analysis identified 3,019 features that differed significantly between UGT71L1-KO and control plants. From these, a list of 68 tentatively identified differentially expressed metabolites in the UGT71L1-KO lines was generated by manual inspection of spectra, annotation, accurate masses, and MS2 fragmentation analysis (Supplemental Table T2.2).

We also probed the effect of UGT71L1 disruption in our transgenic plants by targeted metabolomic analysis (Lackus et al., 2020, Lackus et al., 2021). This method targeted compounds known to occur in poplar and was validated with authentic standards, providing a positive identification for many of the features and compounds seen in the non-targeted analyses. We again carried out a multivariate analysis of the targeted metabolomic dataset, which consisted of 40 identified poplar salicinoids and phenolics (Supplemental Table T2.3). Principal component analysis and hierarchical clustering again showed a clear separation of the transgenic types into three chemically distinct groups as in the non-targeted metabolomic analysis (Supplemental Figure S2.1).

Loss of UGT71L1 in whole plant transgenics led to a pronounced reduction in the concentrations of major salicinoids in leaves (Table 2.1). Tremulacin and salicortin exhibited reductions of more than 90%, whereas salicin concentration was reduced by 80% compared to control lines. Therefore, unlike our previous observations in hairy root cultures where disruption of UGT71L1 leads to the complete loss of salicortin and tremulacin, the transgenic plants were observed to retain a basal level of salicinoid biosynthesis. The biological impact of these phytochemical changes on herbivore preference was tested with a generalist tree-feeding lepidopteran, *O. leucostigma*. Choice assays with 3rd and 4th instar larvae demonstrated that UGT71L1-KO leaf discs were consumed significantly more than control discs (Fig. 2.3C).

Table 2.1. Salicinoid content in leaves of transgenic UGT71L1-KO (biallelic), partial-KO (monoallelic) and control *P. tremula x P. alba* plants as determined by targeted metabolomic analysis[1]

Compound	UGT71L1-KO	Partial-KO	Control
Salicortin	178.5 (±83.16)[a]	2813.45 (±76.07)[c]	3185.16 (±96.53)[c]
Tremulacin	96.23 (±52.76)[a]	1289.85 (±20.87)[b]	1647.52 (±43.00)[c]
Salicin	33.25 (±4.61)[a]	149.04 (±4.88)[c]	157.91 (±4.41)[c]
Salicin 7-sulfate	13876.9 (±807.39)[ns]	14517.65 (±654.9)[ns]	12418.06 (±377.08)[ns]

[1] Values represent mean peak area·mg $DW^{-1}\cdot10^{-3}$ ± SE. UGT71L1-KO data are from 3 independent transgenic lines (4 biological replicates/line); partial-KO data are from a single line (5 biological replicates); control data are from a pool of 3 independently generated empty vector lines (4 or 5 biological replicates/line) and a wild-type line (3 biological replicates). Significant differences ($P < 0.05$) between genotypes for individual compounds are designated by different letters as determined by Tukey's post hoc HSD test. ns = non-significant.

The metabolomic analysis indicated that salicin-7-sulfate, a recently discovered poplar salicinoid (Noleto-Dias, et al., 2018; Lackus et al., 2020), was only slightly affected by the loss of UGT71L1. Our targeted data showed equivalent concentrations across genotypes whereas the non-targeted data showed a relative increase in UGT71L1-KO lines. The discrepancy between both data sets is likely due to a difference in leaf age, as the non-targeted data were derived from mature fully expanded leaf samples whereas targeted analysis was conducted on younger expanding leaves. In addition, we observed an increase in the minor salicinoids such as salireposide and trichocarpin in UGT71L1-KO lines (Supplemental Table T2.3), suggesting parallel pathways and redundant UGTs. Grandidentatin, a related phenolic glycoside, also increased in concentration in UGT71L1 knockouts.

Both targeted and non-targeted metabolomic analyses showed that many non-salicinoid phenolic compounds are severely impacted in UGT71L1-KO plants. Reductions were observed for some flavonoids including kaempferol, dihydrokaempferol, apigenin and naringenin, and their glycosides (Supplemental Table T2.2 and Supplemental Table T2.3). Lignin content was also reduced, as determined through acetyl-bromide analysis (Supplemental Figure S2.2). We observed reduced levels of several caffeoyl quinic and coumaroyl quinic acids, which are abundant constituents of

poplar leaves (Ma et al., 2018). In general, most identified phenolic compounds decreased in UGT71L1-KO plants (Table 2.2). However, some phenolics became more abundant, and we observed significant increases in concentration of small benzoic acid-derived compounds. For example, gentisic acid (2,5-dihydroxy benzoic acid) and 2,3-dihydroxy benzoic acid concentrations increased substantially in UGT71L1-KO plants. Benzoic acid and benzyl alcohol glucosides were elevated in UGT71L1-KO plants, though at overall low concentrations (Supplemental Table T2.2 and Supplemental Table T2.3). A compound tentatively identified via MS2 fragmentation analysis as benzoic acid glucoside-2,3-hydroxy-2-ethoxy-4,6-cyclohexene, this compound was significantly elevated in concentration in UGT71L1-KO plants, shares structural similarities with benzoic acid glucoside, and may be an intermediate molecule in the biosynthesis of the HCH moiety. Altogether, we observed broad impacts of UGT71L1 knock-out on phenolic pathways, suggesting direct and indirect metabolic effects of UGT71L1.

Table 2.2. Phenolic content in leaves of transgenic UGT71L1-KO (biallelic), partial-KO (monoallelic) and control *P. tremula x P. alba* plants as determined by non-targeted metabolomic analysis

Compound type[1]	UGT71L1-KO	Partial-KO	Control
Salicinoids	5196.27 (±285.35)[a]	43089.75 (±2430.75)[b]	58145.9 (±2127.27)[c]
Flavonols and flavanones	369.08 (±25.16)[a]	10770.07 (±655.44)[b]	12380.45 (±433.35)[c]
Procyanidins	583.94 (±36.78)[a]	3922.67 (±208.93)[b]	4658.04 (±174.13)[c]
Chlorogenic acids	676.73 (±43.95)[a]	1515.57 (±138.87)[c]	1441.4 (±83.29)[c]
Lignin (% in PFCW)[2]	7.50 % (±0.14)[a]	11.10 % (±0.23)[c]	11.05 % (±0.13)[c]

[1] Values indicate total combined average peak areas·10^{-4} (± SE) of all identified compounds for each type of phenolic compound. Each biological replicate is an average value of 3 technical replicates. UGT71L1-KO data are from a pool of 5 biological replicates for each of 3 independent lines; Partial-KO data represent 5 biological replicates of a single line; control data represents a pool of 3 independent empty vector lines each with 4-5 biological replicates and 5 biological replicates of a wild-type line.

[2] Lignin concentrations are expressed as % by weight of the Protein Free Cell Wall (PFCW), calculated as the mean of 2 technical replicates for each biological replicate. For UGT71L1-KO, 3 independent lines were used each with 2 biological replicates. Partial-KO concentrations were determined from a single independent line with 3 biological replicates, and in controls from a pool of 3 independent empty vector lines each with 3 biological replicates and one wild-type line with 3 biological replicates. Significant differences ($P < 0.05$) between genotype for individual compounds are designated by different letters as determined by a Tukey's post hoc HSD test.

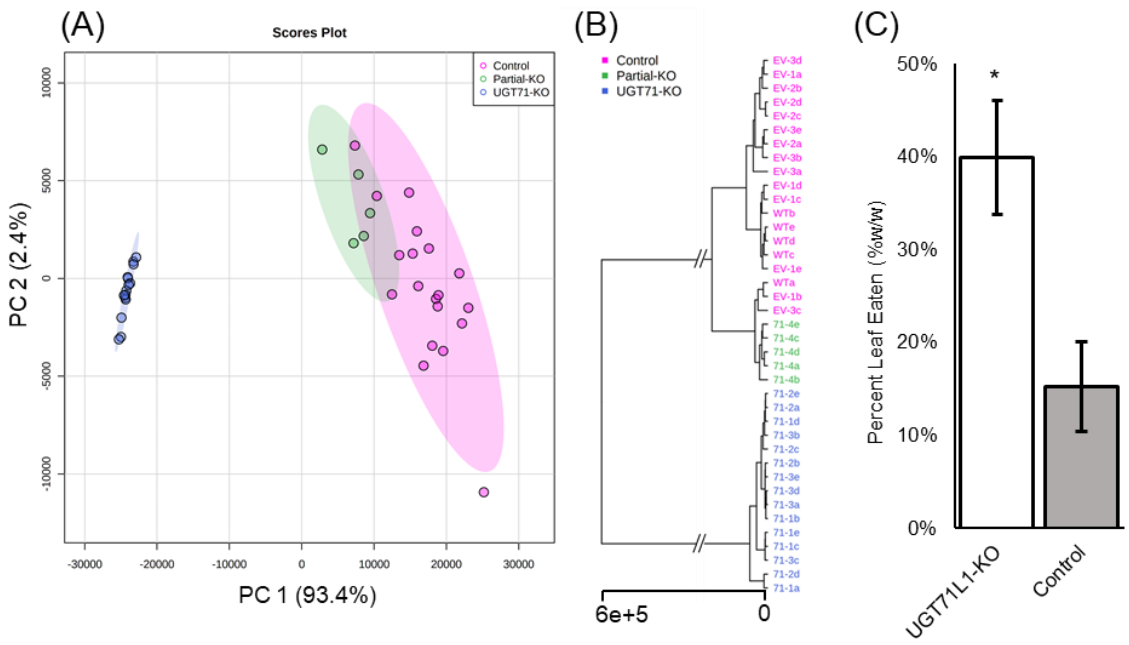


Figure 2.3 Non-targeted metabolomic analysis of UGT71L1-KO plants. (A) Principal component (PC) score plots of non-targeted metabolomic analysis of 3 independent UGT71L1-KO lines (5 biological replicates each), one partial-KO line (4 biological replicates), and controls consisting of 3 empty vector lines and a wild-type (4 or 5 biological replicates each). (B) Unweighted pair group method with arithmetic mean hierarchical cluster analysis of non-targeted metabolomic samples. Multivariate analysis was conducted using Metaboanalyst 5.0 (Chong et al., 2019). Blue colour indicates UGT71L1-KO lines, green indicates partial-KO lines, and magenta indicates empty vector or wild-type control lines. (C) *O. leucostigma* feeding preference (3rd and 4th instar larvae) expressed as percent leaf mass consumed. Leaf discs from 2 independent UGT71L1-KO lines were pooled for choice bioassays with discs from 2 independent empty vector lines. The entire choice test was conducted twice for a total of 25 arena assays (error bars indicate ± SE). * indicates significant difference (two-tailed *t*-test, P < 0.005).

2.3.3 Disruption of UGT71L1 leads to significant growth alterations that can be reversed by retransformation with a functional enzyme

While UGT71L1-KO plantlets in tissue culture showed minor and non-significant growth alterations, these became progressively more apparent after plants were acclimated and moved to the greenhouse. This was seen in all independently generated transgenic UGT71L1-KO poplar lines. We first observed that leaves from UGT71L1-KO plants had a serrated or irregular margin, with a narrower leaf shape than wild-type or empty vector controls and were smaller overall (Fig. 2.4; Supplemental Figure S2.3). UGT71L1-KO plants also showed reduced internode length and grew more slowly. This was manifested as a reduction in stem diameter, plant height, internodal space, and leaf size for all biallelic knock-out lines. By contrast, the monoallelic partial-KO lines were physically indistinguishable from control lines.

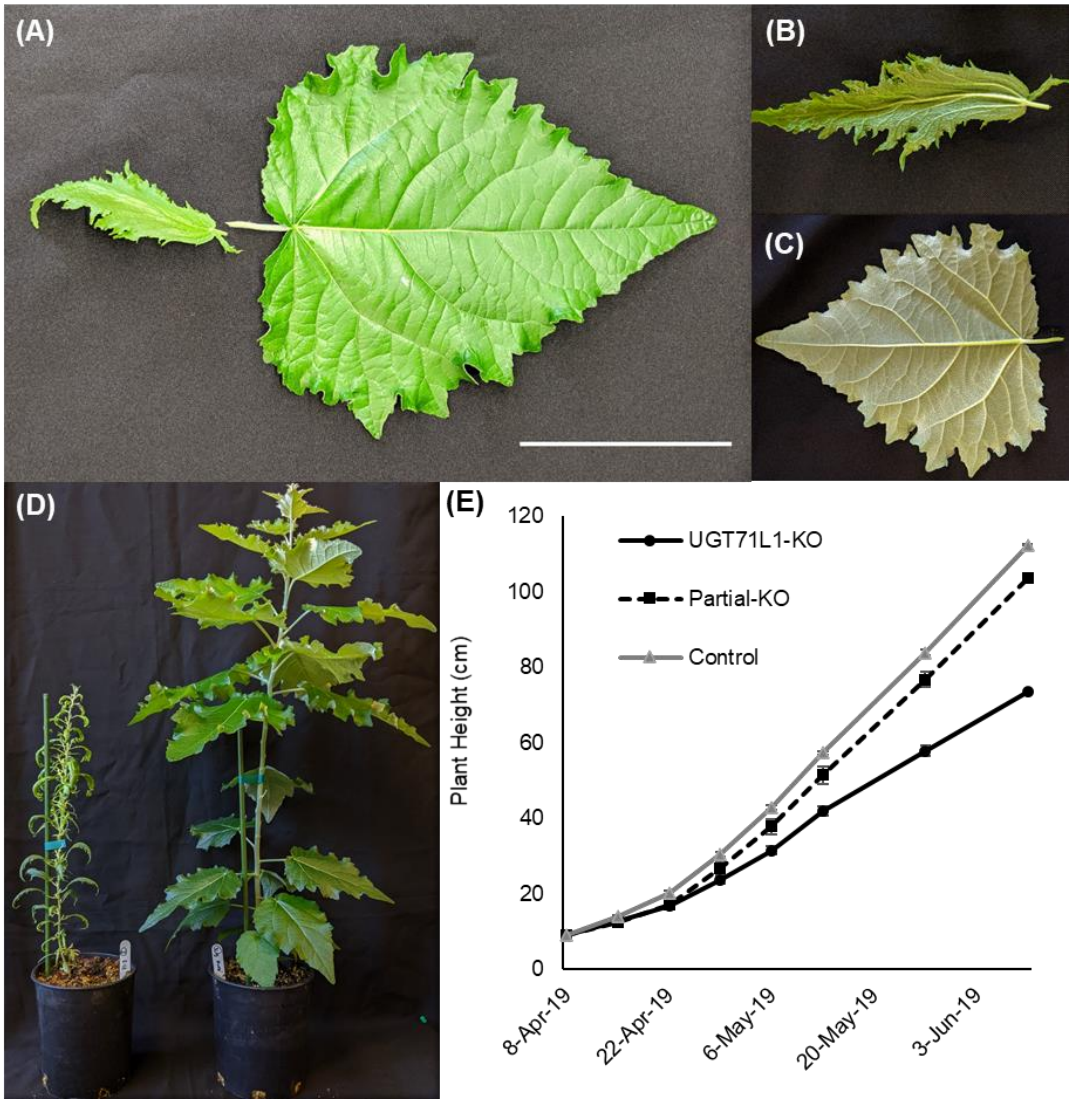


Figure 2.4. Phenotypic effects of UGT71L1 disruption in transgenic poplar. (A) Fully expanded leaves from 3-month-old greenhouse grown UGT71L1-KO plants (left) and empty vector control plants (right). Scale bar represents 10 cm. (B) Abaxial view of fully expanded 3-month-old greenhouse grown UGT71L1-KO leaf. (C) Abaxial view of corresponding control leaf. (D) Whole plant image of UGT71L1-KO (left) and control (right) plants after 3 months of greenhouse growth. (E) Change in plant height for UGT71L1-KO, partial-KO line, and control plants grown in greenhouse for 9 weeks. Data for 7-8 biological replicates each of 3 independent UGT71L1-KO lines, 9 replicates of the partial-KO line, and pooled control samples from 3 independent empty vector lines and a wild-type plant (4-5 replicates per line) are shown. Data are shown as means ± SE.

The observed leaf and growth alterations were surprising, since physiological functions or growth effects for salicinoids have not been previously described. To directly test whether the severe growth phenotypes of UGT71L1-KO were due specifically to *UGT71L1* disruption and not off-target or other effects, we carried out a transgenic rescue experiment. One of the biallelic UGT71L1-KO transgenic lines was retransformed with binary vector to insert a synthetic *UGT71L1* sequence back into the mutant. The coding sequence of the synthetic gene had been modified to eliminate the gRNA binding site, making it immune to editing by CRISPR/Cas9. Transgenic rescued plants were regenerated, moved into soil and acclimated to the greenhouse in parallel with the original UGT71L1-KO line and the wild-type control. All transgenic rescue lines developed normally without any visible growth abnormalities, and rescued plantlets were physically indistinguishable from wild-type plants (Fig. 2.5). UPLC quantification of the major salicinoids confirmed that salicinoid biosynthesis was also restored (Fig. 2.5, right panel). Concentrations of SA-glucoside, which typically accumulates proportionally with SA and is a useful proxy, also decreased back to wild-type levels. This finding suggested that the rescue plants had lost the high-SA phenotype and that elevated SA is directly linked to the loss of UGT71L1. These data also established that the growth phenotype observed in UGT71L1-KO is caused by the disruption of UGT71L1 and is not due to other effects of the CRISPR/Cas9 construct.

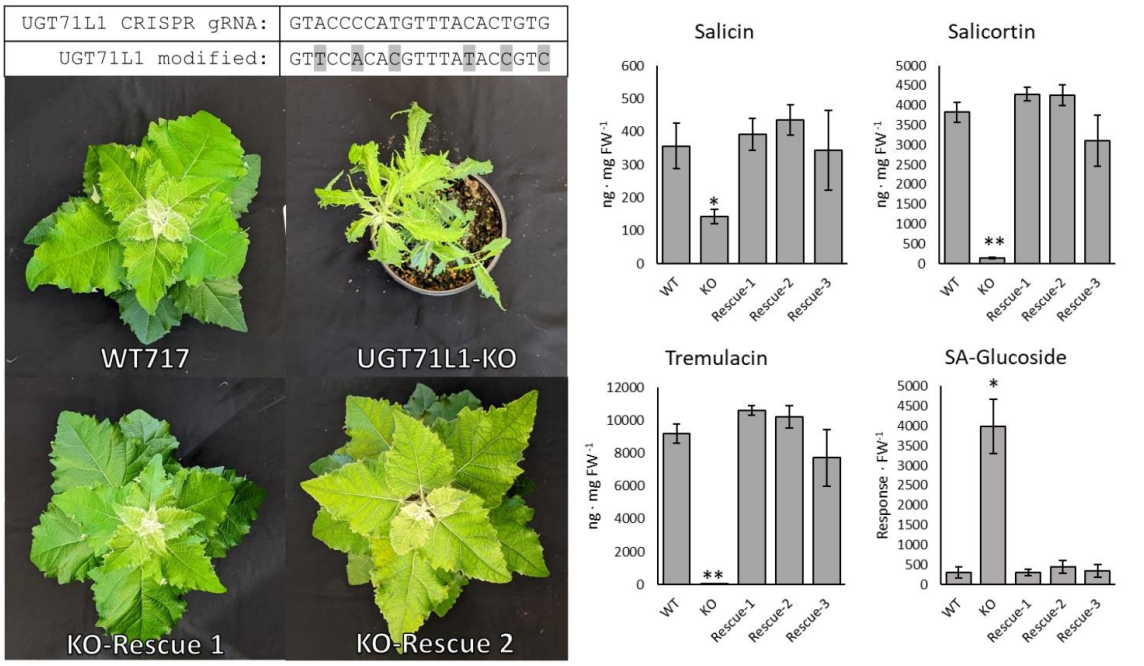


Figure 2.5. Rescue of UGT71L1-KO plant phenotype by retransformation with a synthetic UGT71L1 coding sequence. Upper left panel (top) shows the gRNA target sequence for native UGT71L1 and the corresponding sequence modified to be resistant to the CRISPR/Cas9 construct. Degenerate codons were modified as indicated in gray highlights in the lower nucleotide sequence. Lower left images show two independently transformed 2-month-old rescue plant lines, together with a wild-type and a UGT71L1-KO line. Right panels show concentrations of major salicinoids and salicylic acid glucoside in expanded leaves of 2-month-old greenhouse-grown poplars. Values shown are means ± SE. WT, wild-type with 5 biological replicates; KO, UGT71L1-KO with 4 biological replicates. Rescue-1, -2, and -3 represent independently transformed rescue lines, with 4-5 biological replicates per line. Statistical significance was determined using a *t*-test. * Indicates significant ($P < 0.05$) differences from WT plants; ** indicates $P < 0.01$.

We next considered that plant growth regulators and hormones may be affected and analyzed these using a previously established method (Irmisch et al., 2014). This analysis revealed that both SA and SA-glucoside were 10-fold higher in UGT71L1-KO plants compared to control lines (Table 2.3), an intriguing observation given the structural similarity of SA to salicinoids. High SA content also correlated with elevated levels of 2,5- and 2,3- hydroxybenzoic acids, both SA breakdown products (Zhang et al., 2017). Abscisic acid and indole-3-acetic acid were quantified but were not consistently altered in transgenic lines. However, concentrations of JA and several other jasmonates were strongly enhanced. Importantly, the most active jasmonates, JA and JA-Ile, were elevated by 27- and 17-fold, respectively. 12-hydroxy-jasmonic acid sulfate production exhibited an opposite trend and was detected at very high concentrations in control lines and was less abundant in the UGT71L1-KO lines. The biological role of 12-hydroxy-jasmonic acid sulfate is unclear, but its presence at such high concentrations suggest it may not be active as a regulatory molecule (Fernández-Milmanda et al., 2020). We also measured soluble amino acid concentrations in leaves, which overall increased approximately 3-fold in UGT71L1-KO plants. Most amino acids, apart from threonine, showed an increased concentration in UGT71L1-KO lines. Asp, Glu, and Trp were among the most abundant amino acids and exhibited the greatest changes, with 26-fold, 5-fold, and 2.5-fold concentration increases, respectively (Table 2.4).

Table 2.3. Phytohormone content of young UGT71L1-KO, partial-KO and control *P. tremula x P. alba* leaves[1]

Compound[2]	UGT71L1-KO	Partial-KO	Control
Salicylic acid (SA)	30.27 (±6.04)[a]	7.95 (±0.75)[c]	3.81 (±1.6)[c]
SA-glucoside	328.87 (±29.78)[a]	95.69 (±35.99)[c]	32 .97 (±1.83)[c]
Jasmonic acid (JA)	17.27 (±1.95)[a]	0.99 (±0.14)[c]	0.63 (±0.06)[c]
JA-glucoside	168.66 (±7.41)[a]	640.63 (±23.1)[c]	588.65 (±17.44)[c]
12-OH-JA	8.92 (±0.55)[a]	2.42 (±0.34)[c]	3.06 (±0.19)[c]
JA-Ile	0.16 (±0.02)[a]	0.01 (±0.001)[c]	0.01 (±0.002)[c]
OH-JA-Ile	0.22 (±0.01)[a]	0.07 (±0.01)[c]	0.06 (±0.005)[c]
COOH-JA-Ile	0.02 (±0.008)[a]	0.007 (±0.002)[ac]	0.007 (±0.001)[c]
OPDA	9.01 (±0.92)[a]	7.48 (±0.25)[a]	4.46 (±0.18)[c]
Sulfo-JA	4873.8 (±194.73)[a]	20903.2 (±648.3)[c]	20957.39 (±680.02)[c]
Abscisic acid (ABA)	6.86 (±0.82)[a]	20.37 (±4.47)[b]	12.11 (±1.17)[c]
Indole acetic acid[3]	297.14 (±10.9)[a]	193.15 (±11.56)[b]	246.43 (±9.03)[c]

[1] Content in ng·mg DW $^{-1}$ (± SE). UGT71L1-KO data represents a pool of 5 biological replicates for each of 3 independent lines; partial-KO data represent 5 biological replicates of a single line; control data are from a pool of 3 independent empty vector lines each with 5 biological replicates and a wild-type line with 5 biological replicates. Significant differences ($P < 0.05$) between genotype for individual compounds are designated by different letters as determined by Tukey's post hoc HSD test.

[2] Abbreviations: SA, salicylic acid; JA, jasmonic acid; OH, hydroxy; OPDA, 12-oxo-phytodienoic acid; Ile, isoleucine; ABA, abscisic acid; Sulfo-JA: 12-hydroxy jasmonic acid sulfate.

[3] Indole acetic acid values represent response·g DW^{-1} (± SE).

Table 2.4. Soluble amino acid content in young leaves of UGT71L1-KO, partial-KO, and control poplar[1]

Amino acid	UGT71L1-KO	Partial-KO	Control
Alanine	8.19 (±0.39)^{a}	3.63 (±0.29)^{c}	4.60 (±0.25)^{c}
Serine	17.13 (±1.59)^{a}	4.04 (±0.28)^{c}	3.15 (±0.23)^{c}
Proline	0.83 (±0.05)^{a}	0.36 (±0.02)^{c}	0.50 (±0.02)^{c}
Valine	0.59 (±0.04)^{a}	0.20 (±0.01)^{c}	0.18 (±0.007)^{c}
Threonine	5.14 (±0.34)^{ns}	4.54 (±0.2)^{ns}	5.45 (±0.22)^{ns}
Isoleucine	0.56 (±0.04)^{a}	0.16 (±0.009)^{c}	0.14 (±0.004)^{c}
Aspartate	9.03 (±0.41)^{a}	4.97 (±0.35)^{b}	7.45 (±0.23)^{c}
Glutamate	25.59 (±1.4)^{a}	15.62 (±1.07)^{b}	21.22 (±0.89)^{c}
Methionine	0.31 (±0.02)^{a}	0.12 (±0.008)^{c}	0.12 (±0.004)^{c}
Histidine	2.72 (±0.31)^{a}	0.38 (±0.06)^{c}	0.32 (±0.01)^{c}
Phenylalanine	0.93 (±0.09)^{a}	0.54 (±0.05)^{c}	0.29 (±0.01)^{c}
Arginine	4.26 (±0.55)^{a}	0.11 (±0.02)^{c}	0.11 (±0.02)^{c}
Tyrosine	0.10 (±0.007)^{a}	0.05 (±0.002)^{c}	0.05 (±0.001)^{c}
Asparagine	43.24 (±5.78)^{a}	2.32 (±0.33)^{c}	2.71 (±0.23)^{c}
Glutamine	23.73 (±2.27)^{a}	4.14 (±0.74)^{c}	4.14 (±0.54)^{c}
Lysine	7.00 (±0.6)^{a}	1.46 (±0.28)^{c}	1.45 (±0.16)^{c}
Tryptophan	53.67 (±4.91)^{a}	25.09 (±0.7)^{c}	19.23 (±0.69)^{c}
Leucine	0.34 (±0.02)^{a}	0.07 (±0.005)^{c}	0.07 (±0.001)^{c}
S-methylmethionine	0.07 (±0.005)^{a}	0.005 (±0.0004)^{c}	0.004 (±0.0002)^{c}
Totals	**203.53 (±18.89)^{a}**	**67.88 (±4.49)^{c}**	**71.29 (±3.56)^{c}**

[1] Values are concentrations in nmol·mg DW^{-1} (± SE). UGT71L1-KO data represents 5 biological replicates each of 3 independent lines, Partial-KO are 5 biological replicates of a single line, and controls are a pool of 3 independent empty vector lines each with 5 biological replicates and 5 biological replicates of a wild-type line. Significant differences ($P < 0.05$) between genotype for individual compounds are designated by different letters as determined by a Tukey's post hoc HSD test. ns = non-significant.

2.3.4 Transcriptome analysis indicates a general upregulation of defense metabolism and downregulation of photosynthetic processes in UGT71L1-KO plants

To further investigate the effect of UGT71L1 disruption on poplar growth and metabolism an RNA-seq experiment was conducted. We generated RNA-seq libraries for three biological replicates each of an individual UGT71L1-KO, wild-type, empty vector control, and partial-KO line, which were subjected to high-throughput Illumina sequencing. Mapping of reads to the *P. tremula* transcriptome using AspenDB (Xue et al., 2015) identified 16,343 unique transcripts, which were then used to identify

differentially expressed genes. Comparison of the biallelic UGT71L1-KO plants with wild-type (WT) identified a set of 5,982 differentially expressed genes at $P < 0.05$. Enrichment analysis using the PoplarGene webtool (Liu et al., 2016) identified striking trends in the UGT71L1-KO transcriptome. First, the most highly enriched GO categories were related to pathogen defense including chitinase activity, cell wall catabolism, phenylpropanoid biosynthesis, and salicylic acid glucoside biosynthetic genes (Fig. 2.6). Second, most primary metabolic GO categories including carbon fixation were under-represented in the UGT71L1-KO transcriptome. Both shifts in gene expression suggest the activation of plant defense in UGT71L1-KO plants. Such activation is consistent with the elevated SA levels we observed (Table 2.3).

We next inspected the list of differentially regulated genes from key pathways related to salicinoid, shikimate, and phenylpropanoid biosynthesis and defense (Table 2.5). Several enzyme-encoding genes of the shikimic acid and aromatic amino acid pathway were downregulated in UGT71L1-KO plants, in particular those encoding 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase, shikimate dehydrogenase, chorismate mutase (CM) as well as prephenate aminotransferase (PAT). By contrast, general phenylpropanoid and flavonoid genes were not significantly affected by UGT71L1 disruption, except for a gene for 4-coumaroyl CoA ligase (4CL) and one for flavonoid-3',5'-hydroxylase and flavonol synthase. Phenylalanine ammonia lyase and chalcone synthase are all encoded by multiple genes in poplar, none of which were significantly differentially expressed.

JA biosynthesis and metabolic genes were significantly upregulated in the UGT71L1-KO plants (Table 2.5). Using the Plant Metabolic Network POPLARCYC 3.0 metabolic network database we explored genome-wide metabolic pathways and gene clusters in UGT7L1-KO and WT lines (Schläpfer et al., 2017). JA biosynthetic genes 13-lipoxygenase 2 (LOX2), 3-oxo-2-(2'-[Z]-pentenyl)cyclopentane-1-octanoic acid (OPC – 8:0) CoA ligase 1 (OPCL1), jasmonate-amido synthetase (JAR1), and JA carboxyl methyltransferase (JMT) were all significantly upregulated in UGT71L1-KO lines. JA modifying and degradation genes including those encoding putative JA-Ile 12 hydroxylases, JA-Ile amino hydrolase (ILL6), IAA-alanine hydrolase (IAR3), JA-ZIM domain (JAZ), coronatine insensitive 1 (COI1), and JA-sulfotransferases (JA-SOT) were

also significantly upregulated alongside the biosynthetic genes (Liechti and Farmer, 2002; Mugford et al., 2009; Huang et al., 2017). It thus appears both JA synthesis and inactivation genes were upregulated in these plants; together with our observed enhancement of JA and JA-Ile levels, this suggests that JA metabolism was stimulated. No cohesive trend in GA-related gene expression was identified; GA biosynthetic genes exhibited both significant up- and down-regulation (Supplemental Table S2.4). Transcripts encoding salicylic acid glycosyltransferase (SAGT), an enzyme that converts SA to its inactive glucoside (Dean and Delaney, 2008), accumulated more than 4-fold higher levels in UGT71L1-KO plants. This gene was likely upregulated in response to the high SA content and contributed to the 10-fold increase in SA-glucoside we observed (Table 2.3).

The correspondence of higher SA and JA content as well as greater abundance of SA- and JA- related transcripts in the UGT71L1-KO plants motivated us to look closely at the growth impacts of these phytohormones. SA and JA both modulate growth-defense trade-offs and cause growth inhibition in many plants (Zhang and Turner, 2008; Huang et al., 2017; Guo et al., 2018). The elevated levels of these stress hormones could therefore underlie the unique physical UGT71L1-KO plant phenotype and reduced growth and expansion. As a preliminary test of this hypothesis, we treated plantlets with exogenous SA or the potent and highly stable JA analogue coronatine (Feys et al., 1994; Xiao et al., 2004). No effect was observed after SA treatment, perhaps because SA can be readily inactivated by plants (Supplemental Figure S2.4). However, after five weeks of treating plantlets with 25 μM coronatine, reduced growth and smaller leaves relative to water treated plants was observed (Supplemental Figure S2.4). These effects confirmed the interaction of JA with growth in poplar.

Table 2.5. Differentially expressed genes (P-value ≤ 0.05) in leaves of three-month-old UGT71L1-KO plants compared to wild-type controls[1]

Gene	Poplar Gene ID[2]	Fold-change	p-Value	Adj. p-value
Shikimate Pathway				
DAHPS[3]	Potri.002G099200	0.62	<0.01	0.02
Shikimate Dehydrogenase (SDH)	Potri.013G029800	0.24	0.01	0.05
Chorismate Mutase (CM)	Potri.010G162300	0.49	0.02	0.07
Arogenate dehydratase (ADT/PAT)	Potri.009G148800	1.57	<0.01	<0.01
Salicinoid and Phenylpropanoid Pathway				
UGT71L1	Potri.016G014500	0.24	<0.01	<0.01
4-coumarate CoA ligase (4CL)	Potri.003G188500	0.47	0.04	0.12
Flavonoid 3'5'-hydroxylase (F3'5'H)	Potri.009G069100	0.23	<0.01	<0.01
Flavonol synthase (FLS)	Potri.004G139600	0.18	<0.01	<0.01
	Potri.004G139700	0.32	<0.01	<0.01
Salicylic acid glycosyltransferase (SAGT)	Potri.007G140500	4.52	<0.01	<0.01
Jasmonate (JA) Metabolism				
13-Lipoxygenase 2 (LOX2)	Potri.001G015600	3.86	<0.01	<0.01
OPC CoA ligase 1 (OPCL1)	Potri.017G033600	4.82	<0.01	<0.01
	Potri.015G092300	1.65	0.02	0.07
	Potri.012G095000	2.97	<0.01	<0.01
Jasmonate-amido synthetase (JAR1)	Potri.002G168200	2.48	<0.01	0.01
JA carboxyl methyltransferase (JMT)	Potri.005G230100	1.83	0.04	0.13
CYP94B - JA-Ile 12 hydroxylase	Potri.005G220700	4.35	<0.01	<0.01
	Potri.001G277300	0.58	<0.01	<0.01
CYP94C1- JA-Ile 12 hydroxylase	Potri.004G185300	63.56	0.01	0.03
IAA-alanine hydrolase (IAR3)	Potri.010G035900	2.83	<0.01	<0.01
	Potri.003G045200	2.03	<0.01	<0.01
Jasmonate-ile amino hydrolase (ILL6)	Potri.002G082400	7.89	<0.01	<0.01
Jasmonate ZIM domain (JAZ) (JAZ6)	Potri.003G068900	24.59	<0.01	<0.01
Jasmonate ZIM domain (JAZ3)	Potri.015G035800	1.68	<0.01	<0.01
Coronatine Insensitive 1 (COI1)	Potri.008G064400	1.32	<0.01	0.01
JA-Sulfotransferase (JA-SOT)	Potri.003G188800	4.72	<0.01	0.02
	Potri.003G193500	9.78	<0.01	0.01
	Potri.003G005200	11.88	<0.01	<0.01

[1] RNA-seq data from 3 biological replicates of one UGT71L1-KO line and 3 biological replicates of the wild-type control line were analyzed analyzed by HiSat2 (Kim, et al., 2019) and Cufflinks (Trapnell et al., 2012). differentially expressed genes were identified using the DeSeq2 (Love et al., 2014) R package.

[2] More than one poplar gene ID may have the same Arabidopsis gene correspondence and annotation

[3] Abbreviations: DAHPS, 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase; OPC, 3-oxo-2-(cis-z-pentenyl)-cyclopentane-1-octanoic acid (OPC); JAR1, jasmonate resistant 1; IAA, indole-acetic acid; ILE, isoleucine; ZIM, zinc finger expressed in inflorescence meristem.

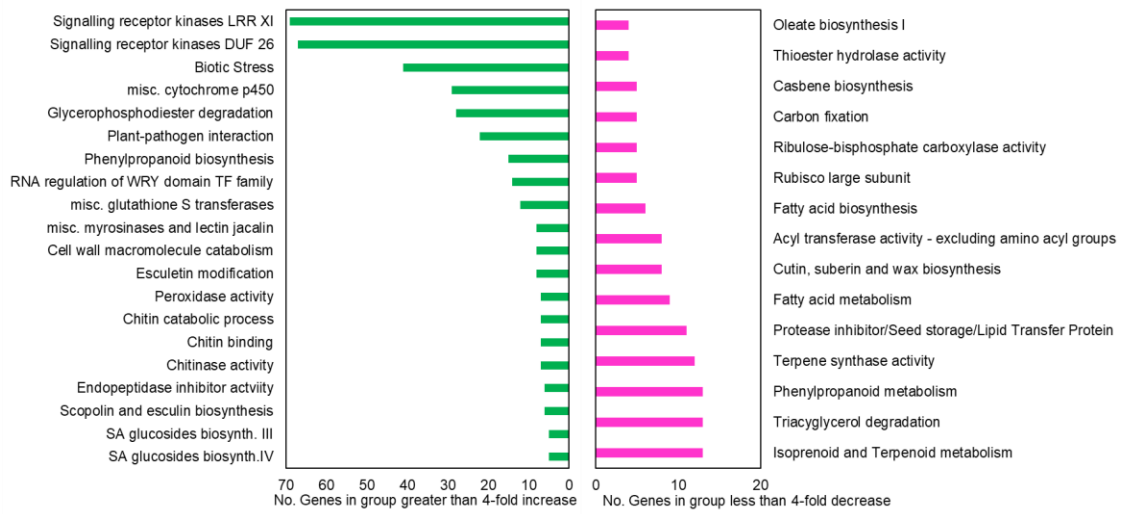


Figure 2.6. Gene ontology (GO) terms significantly enriched in UGT71L1-KO transcriptome compared to wild-type analyzed by the PoplarGene webtool. Green bars indicate on left significant enrichment of GO terms in UGT71L1-KO lines (log_2FC >2); magenta bars on right show significant enrichment of GO terms in wild-type plants (log_2FC >2).

2.3.5 SA is a potential product of disrupted salicinoid biosynthesis and possible driver of the UGT71L1-KO phenotype

Although SA is not a precursor or otherwise involved in salicinoid biosynthesis, we hypothesized that it could be released from a salicinoid intermediate when UGT71L1 is blocked. Specifically, SA could be generated during salicinoid biosynthesis if the benzoate ring moiety of salicyl benzoate is first hydroxylated to salicyl salicylate prior to the UGT71L1-catalyzed glucosylation step (Fig. 2.1). In the absence of UGT71L1, salicyl salicylate could not continue into salicinoid biosynthesis. As an unstable molecule, it could be readily hydrolyzed into salicyl alcohol and free SA. Postulating that salicyl benzoate is hydroxylated prior to glucosylation is a variation of our previously proposed pathway (Fellenberg et al., 2020). It would be a direct mechanism for the

substantial accumulation of SA observed in UGT71L1-KO plants, however (Fig. 2.1). It also represents a previously unrecognized pathway for the generation of SA in plants (see Discussion).

This hypothesis predicts that salicyl salicylate should also be a substrate for UGT71L1. Salicyl salicylate is not commercially available, but it was synthesized for enzyme assays with recombinant UGT71L1, together with salicyl benzoate and benzyl salicylate. The enzyme was able to efficiently convert salicyl salicylate to the corresponding glucoside (Fig.2.7; Supplemental Figure S2.5), although we did not determine kinetic parameters due to a limited supply of this compound. We also repeated assays using salicyl salicylate with UGT78M1, a second UGT hypothesized to be involved in biosynthesis of other minor salicinoids, but only a very small or no role in biosynthesis of salicortin or tremulacin (Fellenberg et al., 2020). The profile of activity with all three substrates with UGT78M1 was similar to UGT71L1. The acceptance of salicyl salicylate as a glucose acceptor for both enzymes is notable, since both enzymes take a narrow range of substrates (Fellenberg et al. 2020). The ability of these UGTs to glucosylate salicyl salicylate provided indirect experimental evidence for the hypothesized sequence of reactions and potential source of SA suggested above.

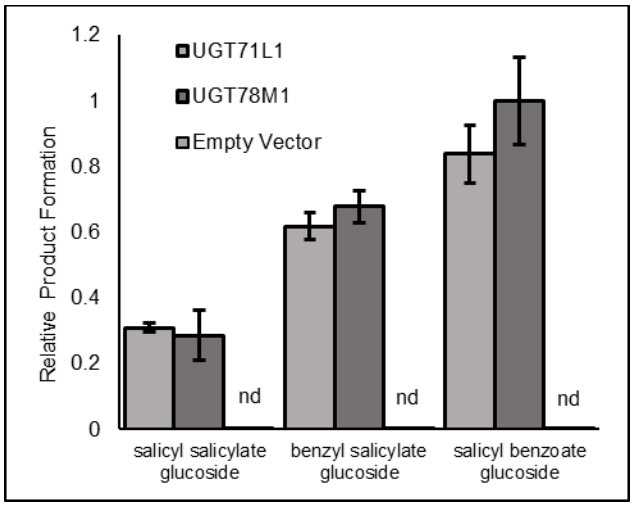


Figure 2.7. Enzymatic activity of recombinant UGT71L1 and UGT78M1 with potential salicinoid biosynthetic intermediates. Bars show means (n = 3) of relative product formation relative to UGT78M1 glucosylation of salicyl benzoate under standard conditions. Empty vector controls were carried out in duplicate and consisted of extracts from bacterial host strains without the recombinant enzyme. Each assay contained 5 µg protein and 250 µM substrate and was conducted as described in materials and methods. Product formation was quantified by UPLC-UV and absorbance at 280 nm. Data shown are means ± SE.

2.4 Discussion

The demonstration that UGT71L1 is required for the biosynthesis of salicortin, tremulacin, and tremuloidin in poplar tissue cultures (Fellenberg et al., 2020; Kulasekaran et al., 2020) was a key advance in our understanding of salicinoid biochemistry. The present work with UGT71L1 whole-plant knockouts broadens these findings and suggests the presence of additional enzymes and routes in a biosynthetic network. Metabolomic and transcriptomic data indicated the unexpected impact of disruption of a key step in salicinoid biosynthesis on other phenolic pathways as well as primary

metabolism and growth, potentially via feedbacks to the shikimate pathway and highly elevated levels of SA and JA. Together, our experiments suggest that excess SA is released in UGT71L1-KO poplars from salicinoid metabolites, thus driving the chemical and physical phenotype of these plants and revealing a novel mechanism for generating SA in poplar.

2.4.1 UGT71L1 is a central enzyme in the pathway to salicortin and tremulacin, but additional UGTs are likely involved in the biosynthesis of other salicinoids

Our knock-out plants confirm that UGT71L1 is a core enzyme in the pathway to salicortin and tremulacin, the major salicinoids in *P. tremula x P. alba*. Unlike UGT71L1-KO hairy root cultures where tremulacin and salicortin could not be detected, in whole plants UGT71L1 disruption by CRISPR/Cas9 reduced these compounds by 90% - 95% but did not eliminate them. This suggests redundancy in the pathway and implies the presence of other UGT enzymes and genes with parallel functions (see below). Salicin content in UGT71L1-KO plants was less affected than salicortin or tremulacin and only reduced to 20% of wild-type concentrations. Prior in vivo labeling experiments had already demonstrated that salicin biosynthesis proceeds by a route that is distinct from the pathway to larger salicinoids, since salicyl alcohol and salicyl aldehyde are incorporated efficiently into salicin but not salicortin (Zenk, 1967; Babst et al., 2010). For example, salicin may also be derived via *o*-coumarate via salicylaldehyde as suggested by Babst et al. (2010). Such alternate routes to salicin would require additional UGT enzymes. In addition, salicin may also be derived via hydrolysis of the ester bond in salicortin or salicyl benzoate glucoside (Fig. 2.1), which could explain the partial reduction by loss of UGT71L1 we observed.

Interestingly, UGT71L1 disruption did not reduce salicin-7-sulfate concentrations, and in older leaves enhanced its accumulation. This compound was recently found to be present in some, but not all, willow and poplar species (Noleto-Dias et al, 2018; Lackus et al., 2020) and is highly abundant in the *P. tremula x P. alba* hybrid. The function of salicin-7-sulfate is not known, but it does not affect *L. dispar* feeding preference and thus

appear not to be an important compound for defense against lepidopteran herbivores (Lackus et al., 2020). Our metabolomic data show that while young leaves (targeted analysis) contain equivalent concentrations of salicin-7-sulfate in UGT71L1-KO lines and controls, older UGT71L1-KO leaves (non-targeted analysis) showed elevated concentrations. This pattern correlated with a 12-fold increase in expression of the corresponding sulfotransferase gene *PtSOT1* (Lackus et al., 2020) in UGT71L1-KO plants relative to wild-type seen in our RNA-Seq data. It suggests that synthesis and accumulation of salicin-7-sulfate is not disrupted and perhaps even favored in UGT71L1 knock-outs. How the biosynthesis of salicin-7-sulfate and other sulfated salicinoids is related to that of the non-sulfated versions, and which UGT enzymes are involved, requires further study. Since salicin-7-sulfate is not found in some species of *Populus* that do synthesize other salicinoids, this compound may not have a central role in salicinoid metabolism.

Likewise, we observed that the minor salicinoids trichocarpin and salireposide were not impacted by UGT71L1 disruption, but rather increased in concentration. Their presence implies that they are synthesized by distinct UGTs. These salicinoids contain benzyl alcohol or 2,5-hydroxybenzyl alcohol, respectively, rather than salicyl alcohol (2-hydroxybenzyl alcohol) at their core, which suggests there may be UGTs with substrate preferences distinct from UGT71L1. These enhanced levels of trichocarpin and salireposide, the persistence of low concentrations of salicortin, tremulacin and salicin, and the presence of large concentrations of salicin-7-sulfate all indicated that additional UGT71L1-like enzymes are active in UGT71L1-KO plants and likely contribute to overall salicinoid accumulation. For example, we previously proposed that UGT78M1 could be involved in salicinoid biosynthesis based on its substrate preferences for salicyl benzoate (Fellenberg et al., 2020), but direct evidence for its function using knock-out plants is still lacking. Surprisingly Kulasekaran et al. (2020) report that *Salix* does not have a UGT78M1 homologue. However, genes phylogenetically similar to *UGT71L1* are present in the poplar genome and are promising targets for future study.

2.4.2 Metabolomic analysis of UGT71L1 knock-out plants provides insight into salicinoid biosynthesis

The comparative metabolomic analyses provided a strategy for identifying salicinoid pathway intermediates, since these are expected to accumulate in UGT71L1-KO plants where flux into the salicinoids is blocked. According to our earlier model (Fellenberg et al., 2020), a block at the UGT71L1 step would be expected to result in the accumulation of salicyl benzoate. Alternatively, if as suggested above, the 2-hydroxylation of salicyl benzoate occurs prior the glycosylation step, salicyl salicylate can be expected to accumulate in UGT71L1-KO plants instead (Fig. 2.1). However, we were not able to detect salicyl salicylate or salicyl benzoate in our metabolomic data. Kulasekaran et al. (2020) also noted that in similar experiments using tissue cultured UGT71L1 RNAi plants, neither salicyl salicylate nor salicyl benzoate was detected. These esters may not be stable and could hydrolyze spontaneously or enzymatically. The inability of cells to glucosylate salicyl salicylate or salicyl benzoate due to UGT71L1 disruption would therefore generate a release of benzoic acid or SA, as we observed in UGT71L1-KO leaf extracts.

Since SA concentration in tissues is strictly controlled, free SA is also expected to be rapidly conjugated or modified by metabolic reactions involved in SA homeostasis. This explains the increase in gentisic acid (2,5-dihydroxybenzoate) and 2,3-dihydroxybenzoic acid glucosides in UGT71L1-KO lines, compounds known from work in Arabidopsis to constitute a key pathway for inactivating SA (Zhang et al., 2013; Zhang et al., 2017; Zhang and Li, 2019). Therefore, we propose that the increase in these hydroxybenzoates reflects a metabolic response to the buildup and degradation of salicyl salicylate and/or salicyl benzoate and release of SA. Support for this idea also comes from work in *Nicotiana benthamiana*, where the application of benzyl salicylate to leaves enhances expression of systemic acquired resistance genes and reduces the number and size of tobacco mosaic virus lesions (Kamatham et al., 2017), suggesting that SA is released from this compound. Overall, our metabolomic data are consistent with a proposed salicinoid pathway in which the salicyl benzoate pool is 2-hydroxylated to salicyl salicylate prior to glucosylation by UGT71L1. The salicylic acid moiety of salicyl

salicylate glucoside could then further oxidized to the HCH moiety by as yet unknown enzymes.

An alternate pathway proposes that the HCH moiety is synthesized from benzyl glucose and subsequently transferred to salicinoid acceptor molecules. This could also include oxidation of benzoyl-CoA to HCH-CoA followed by transesterification to other salicinoids, as suggested by Kulasekaran et al. (2020). There is no direct evidence for these alternate routes to date. However, the presence of HCH-salicortin in *P. tremula x P. alba* and many other poplar species (Feistel et al., 2015) suggests a mechanism to transfer HCH via acyltransferase reactions from a donor molecule. Alternatively, HCH may be synthesized from salicyl benzoate glucoside or salicyl salicylate glucoside, but subsequently transferred to other salicinoids. While the ultimate source of HCH is benzoic acid (Babst et al., 2010), how this moiety is derived remains a central question for salicinoid biosynthesis. In vivo-labelling experiments may help to provide an answer.

2.4.3 UGT71L1-KO plants manifest broad perturbations in shikimate and phenylpropanoid metabolism

UGT71L1-KO plants accumulated reduced concentrations of flavonoids, procyanidins and lignin, all abundant phenylpropanoid-derived compounds. The plants also showed a reduction in caffeoyl quinate (chlorogenic acid) and the corresponding coumaroyl- and shikimoyl-esters (Ma et al., 2018). These compounds are directly linked to the shikimate pathway (Clifford et al., 2017). Such broad effects of UGT71L1 disruption were unexpected based on the specific substrate preference of UGT71L1, and they indicate that the phenylpropanoid and/or shikimate pathways were perturbed in UGT71L1-KO plants by additional feedbacks. Consistent with this, our transcriptomic analysis indicated that expression of shikimic acid pathway and aromatic amino acid biosynthesis genes was downregulated: DAHP synthase, shikimate dehydrogenase, chorismate mutase, and prephenate aminotransferase all showed reduced transcript levels. However, few phenylpropanoid pathway genes were found to be down-regulated. Transcriptional mechanisms for regulating the shikimate pathway in plants are poorly

understood but are assumed to involve feedbacks from aromatic amino acids and other downstream products (Maeda and Dudareva, 2012).

By contrast, at the post-transcriptional level complex mechanisms for regulating the shikimate pathway in plants have been described. In particular, feedback regulation of DAHPS, the enzyme controlling pathway influx, appears to play a central role (Maeda and Dudareva, 2012). Negative feedback mechanisms for DAHPS in Arabidopsis were recently characterized in detail (Yokoyama et al., 2021). Whereas the three DAHPS isoforms can be regulated differently by several allosteric regulators, all are inhibited by chorismate and caffeate (Yokoyama et al., 2021). These are both key intermediates in the shikimate and phenylpropanoid pathways, respectively. Interestingly, both caffeate and phenylethanol caffeate levels were enhanced in the UGT71L1-KO plants, as were the aromatic amino acids. We therefore speculate that the broad reduction in flavonoids, lignin, hydroxycinnamate esters and other phenylpropanoids in UGT71L1-KO plants is caused by allosteric downregulation of the shikimate pathway due to accumulation of a downstream product or shikimate pathway intermediate in UGT71L1-KO plants. The downregulation of shikimate and phenylpropanoid products is reminiscent of the broad impact of overexpression of the flavonoid repressors MYB165 and MYB194 (Ma et al., 2018). Although the mechanisms are likely different, in both transgenic experiments it is apparent that the shikimate pathway in poplar is sensitive to perturbation in phenylpropanoid and phenolic metabolism.

The increase in non-salicinoid phenylpropanoids, specifically grandidentatin, in UGT71L1-KO plants remains an area requiring further investigation (Supplemental Table S2.2). Grandidentatin is a glucoside of *p*-coumarate and dihydroxycylcohexane first isolated in *Populus grandidentata* (Pearl et al., 1962; Zhang et al., 2006; Kim and Bae, 2009) but its biosynthesis has not been studied. We previously noted a small increase of grandidentatin in RNAi knockdown plants of BEBT and SABT (Chedgy, 2015), and Coleman et al. (2008) found an increase in grandidentatin after downregulation of the shikimoyl 3-coumarate hydroxylase. Grandidentatin thus appears to be particularly impacted by diverse metabolic disruptions.

2.4.4 Elevated SA and JA levels may connect disrupted salicinoid biosynthesis to plant growth phenotypes

The most striking result from our study is the altered phenotype of UGT71L1-KO plants, which includes reduced leaf and plant size, and which is correlated with strong upregulation in SA-regulated defense genes. The unexpected effect of the UGT71L1 disruption on growth and morphology motivated us to carry out a rescue experiment, retransforming knockouts with a functional *UGT71L1* coding sequence. The recovery of normal leaf shape and growth of these plants, together with wild-type levels of salicinoids and SA-glucoside confirmed that the strong phenotype is due to the loss of UGT71L1 and not caused by off-target effects. The phenotypic abnormalities observed in these poplar lines are particularly striking, since our earlier experiments with MYB repressor-overexpressing poplars showed that such plants also produce very low concentrations of many phenolics and phenylpropanoids, but their growth or form was not affected to an appreciable extent (Yoshida et al., 2015; Ma et al., 2019). This indicates that reduced levels of phenylpropanoids *per se* are unlikely to cause the physical phenotype observed. However, analysis of phytohormones altered in UGT71L1-KO plants suggested that elevated SA and JA may be potential causes of the phenotype. Hyperaccumulation of either SA or JA is known to reduce growth and lead to dwarfed growth forms. For example, smaller plant size is observed in Arabidopsis mutants that lack the SA 5-hydroxylase gene which contributes to SA inactivation, leading to elevated SA content (Zhang et al., 2017).

By contrast, overexpression of this hydroxylase gene leads to larger leaves and rosettes compared to wild type plants due to reduced SA levels. Similar effects were observed in *Medicago*, where transgenic downregulation of monolignols and the phenylpropanoid pathway caused dwarfism; since SA is also derived from the phenylpropanoid pathway, a reversal of the phenotype was achieved by applying exogenous SA (Gallego-Giraldo et al., 2011). These negative effects of SA on growth are manifested as the apparent tradeoff between growth and defense metabolism, which allows the plant to prioritize resources depending on environmental conditions (Zhang et al., 2017; Van Butselaar and Van den Ackerweken, 2020). SA homeostasis is thus critical

to normal growth and development in Arabidopsis and other species, and SA metabolism is tightly regulated at both transcriptional and post-transcriptional levels (Ding and Ding, 2020).

How SA metabolism is organized and regulated in poplar is not known. SA content of wild type poplar leaves in our study was approximately three-fold greater than what is reported for healthy three-week-old Arabidopsis leaves, prior to elicitation after which SA increases sharply (Liu et al., 2016). Previous work in poplar had attempted to manipulate SA concentrations by overexpressing a bacterial IrP9 gene, part of the microbial pathway to synthesize SA from chorismate, or by expressing a bacterial *nahG* salicylate hydroxylase gene in transgenics in order to reduce SA content (Xue et al., 2013). Transgenic poplar expressing IrP9 in plastids showed a 2-3-fold enhanced SA content, while SA-glucoside content increased up to 800-fold, suggesting that the plant is conjugating excess SA. By contrast, overexpressing *nahG* in transgenic poplar caused no change in free SA, emphasizing that SA accumulation is carefully regulated (Morse et al., 2007; Xue et al., 2013). No growth phenotypes were observed in either the IrP6 or NahG transgenic poplars. This could reflect the much smaller changes in SA concentrations observed in these transgenics, compared to a 10-fold increase in free SA measured for UGT71L1-KO plants; it may also reflect localization of SA biosynthesis. In Arabidopsis, 90% of SA is synthesized via the ICS pathway from plastidic isochorismate, which is moved across the plastid envelope into the cytosol where it is converted to SA (Rekhter et al., 2019). This may limit the potential of plastidic overexpression of IrP6. SA can also be produced in the cytosol from phenylpropanoid metabolism and cinnamate, though in most species this pathway produces much less SA. By contrast, we expect that in the UGT71L1-KO plants SA derived from salicyl salicylate is generated entirely in the cytosol based on the localization of the enzyme. This source of SA is separate from the standard cytosolic and plastidic SA biosynthetic pathways and may thus be outside the normal SA homeostatic mechanism that regulate SA concentrations in cells. This could explain the high degree of SA overaccumulation, in turn leading to strong effects on growth.

In addition to SA, the UGT71L1-KO plants showed very high JA and JA-Ile content and also overaccumulated several inactivate forms of JA including 12-hydroxy JA and sulfo-JA. Expression of both JA biosynthetic and catabolic genes was enriched in the UGT71L1-KO transcriptome, further suggesting that JA metabolism was stimulated. An increase in both SA and JA/JA-Ile was unexpected, since these signaling molecules often act antagonistically (Rivas-San Vicente and Plasencia, 2011; Wasternack and Hause, 2013). However, an induction of JA due to SA release has also been observed, for example during effector-triggered immunity in Arabidopsis (Liu et al., 2016). We therefore speculate that in our transgenic plants JA overaccumulation could have been stimulated by SA. It is important to note that the antagonism between SA and JA seen in other species appears not to be present in poplar. Both hormones are upregulated in parallel during pathogen stress (Ullah et al., 2019), and treatment with methyl jasmonate enhances SA content of poplar leaves in vitro (Park et al., 2017). Conversely, treatment of *P. trichocarpa* with SA induces multiple JAZ transcripts (Xia et al., 2017). Interestingly, *JAZ6* was strongly upregulated in UGT71L1-KO in our data (Potri.003G068900, Table 2.5), and was also the most strongly SA-induced JAZ identified in Xia et al. (2017). Together with the high JA and JA-Ile content of our UGT71L1-KO plants, these data suggest that JA metabolism is activated by UTG71L1 disruption, and that SA could explain the upregulation of JAZ and JA-related gene expression.

Like SA, JA inhibits growth as part of a defense-growth tradeoff, including inhibition of root elongation, leaf expansion and hypocotyl elongation (Huang et al., 2017; Guo et al., 2018). In our experiments, the effect of coronatine in inhibiting growth and partially mimicking the CRISPR/Cas9-generated phenotype confirmed that JA growth inhibition is also operational in poplar. Our transcriptomic dataset showed that a gene for GA20 oxidase is significantly downregulated, and some GA2 and GA3 oxidase genes are upregulated in UGT71L1-KO lines. This pattern suggested a reduction in active gibberellins (Supplemental Table S2.4). However, we did not measure gibberellins as part of our experiments. Growth inhibition by JA may be in part mediated by downregulation of gibberellin signaling (Yang et al., 2012; Howe et al, 2018; Guo et al., 2018). Although both JA and SA have different effects on gene expression overall,

common expression modules associated with growth inhibition have been identified for poplars treated with SA or JA (Luo et al., 2019).

Alternative explanations for the UGT71L1-KO phenotype include disruptions of primary metabolism. For example, analysis of amino acids indicated significantly enhanced accumulation of several amino acids, specifically Asn, Gln and Trp. Whether high amino acid content could be a direct cause of growth inhibition is not clear. Recent work in Arabidopsis with the *jazD* mutant, which is deficient in ten JAZ genes and constitutively expresses JA-mediated defense and shows extreme growth inhibition, also reported elevated Trp content and increased expression of Trp biosynthesis genes (Major et al., 2020). This suggests that high JA and growth inhibition can lead to significant adjustment of primary metabolism including elevated Trp, although it may not be the primary driver of the phenotype (Guo et al., 2018). As an auxin biosynthetic precursor, Trp could influence auxin function or metabolism and thus effect growth. However, Arabidopsis differs from poplar in that it requires significant biosynthesis of Trp for the indole glucosinolate group of secondary metabolites. By contrast, most of the poplar defense molecules including salicinoids are phenylpropanoids and thus derived from phenylalanine. Overall, based on the multiple lines of evidence, we favor the hypothesis that the combined effect of salicinoid-derived SA and JA on growth and development is most likely the primary cause for the phenotypic effects observed. We summarize this hypothesis in figure 2.8.

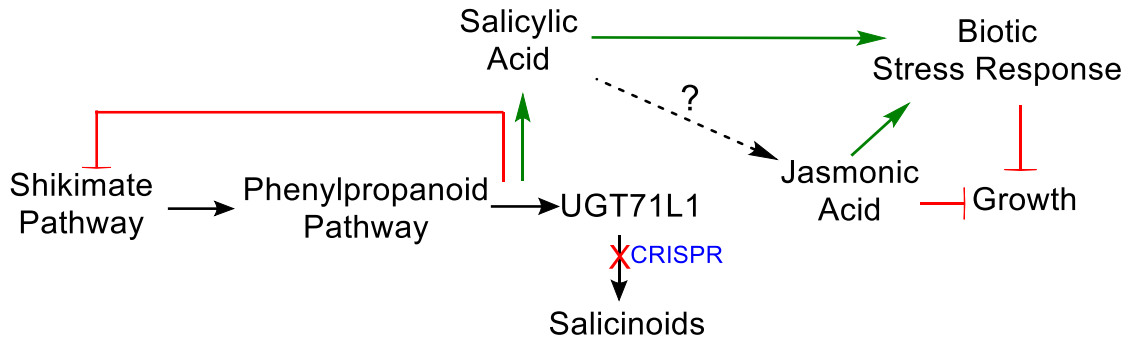


Figure 2.8. Proposed metabolic and hormonal interactions leading to growth and metabolic phenotypes in UGT71L1-KO plants. Flat magenta arrows represent inhibited responses, and pointed green arrows represent stimulated processes. The interruption of salicinoid biosynthesis by CRISPR/Cas9 is shown with a magenta X, suggested to cause the hyperaccumulation of salicylic acid (SA) in these plants. The dashed arrow represents a hypothesized but not yet demonstrated stimulation of jasmonate by SA in poplar. Stimulation of biotic stress responses and inhibition of growth reflect the GO categories of enhanced and repressed genes observed in RNA-seq data (Fig. 2.5), as well as the direct effects of coronatine on poplar (Fig. 2.4). Feedback inhibition of the shikimate pathway is hypothesized based on RNA-seq and metabolomics data (Table 2.5 and Supplementary Tables T2.3 and T2.4), including accumulation of possible allosteric regulator caffeic acid.

2.4.5 Broader implications of the interaction of salicinoid metabolism with growth and development

The growth phenotype of UGT71L1-KO plants points to previously unknown connections of poplar specialized metabolites to primary processes. Interactions of growth regulators and defensive secondary metabolism have been observed in other species and may be more widespread than previously realized (Erb and Kliebenstein,

2020, Duran-Medina et al., 2021). Several studies in Arabidopsis have demonstrated that glucosinolates and their breakdown products can influence growth. Katz et al. (2020) showed that allyl glucosinolate catabolites interact with components of the auxin signaling pathway, leading an inhibition of root growth and development. Similar effects on auxin signaling or growth were observed with indole and aliphatic glucosinolates (Katz et al., 2015; Malinovsky et al., 2017). In *P. trichocarpa*, the herbivore-induced volatile, benzyl cyanide, can be converted to the auxin phenyl acetic acid (Günther et al., 2019). Flavonoids have been observed to affect auxin transport in some experimental systems such as Arabidopsis seedling roots (Peer and Murphy, 2007). In apple, flavonoids and dihydochalcones have been suggested to interact with auxin transport or uptake (Dare et al., 2017). Interestingly, disruption of biosynthesis of the dihydrochalcone glycoside phloridzin by RNAi-inhibition of *MdUGT88F1* in apple causes stunted growth and narrow leaves (Dare et al., 2017), reminiscent of the poplar UGT71L1-KO phenotype described herein. Although we cannot rule out hormonal interactions by other phenolics, our data are most consistent with the hypothesis that SA accumulation is the ultimate cause of the phenotype in the UGT71L1-KO plants. Regardless of the mechanism, our work implies a potential interaction of the salicinoid pathway with growth regulatory processes and adds to a growing list of secondary metabolic pathways impacting plant development.

The linkage of salicinoid metabolism to growth processes could have important implications for the evolution of this pathway. Secondary chemical profiles are generally considered strongly heritable but under weak evolutionary constraints, compared to morphological and phenological traits (Kessler and Kalske, 2018). In *P. tremuloides*, leaf salicinoid content is highly heritable and under strong genetic control, and usually varies several-fold among individual genotypes (Hwang and Lindroth, 1997; Cole et al., 2021). Salicinoids are effective defensive chemicals against both insect and mammalian herbivores (Boeckler et al., 2011), and thus expected to be maintained in populations by herbivore pressure. In a common garden experiment using a suite of *P. tremuloides* genotypes, Cope et al. (2021) used salicinoid content as a proxy for defensive capacity to demonstrate a growth – defense tradeoff. However, based on the apparent persistence of

such tradeoffs across many environments, the authors suggest that although salicinoid content may have evolved as a result of herbivore pressure,

"covariance with other ecologically important traits renders the evolution of defense traits sensitive to selection pressures other than herbivory" (Cope et al., 2021).

Such selective pressures could include additional, metabolic constraints on the control and plasticity of salicinoid biosynthesis, such as the effects of salicinoid-derived SA on growth and development shown here. Tighter metabolic constraints may also explain why salicinoids are generally constitutively expressed and developmentally regulated (Boeckler et al., 2011). This contrasts with the strong stress-induction of the other major phenolic pathways in poplar, including flavonoids and condensed tannins, hydroxycinnamic acid esters, and benzenoid volatiles (Mellway et al., 2009; Lackus et al., 2021). Investigations of salicinoid regulatory genes could shed further light on potential developmental linkages with salicinoid biosynthesis. Such genes still await identification, however.

2.5 Supplementary Material

Supplemental Table T2.1 Off-target Sanger sequencing analysis utilized to screen off-target CRISPR mutations as per Fellenberg et al. (2020). Bases flanked by brackets represent single nucleotide polymorphisms between *P. alba* and *P. tremula* alleles.

Off Target Site Sequence	**Poplar Gene ID**	**Mutations**			
		ugt71l1-1	*ugt71l1*-2	*ugt71l1*-3	*ugt71l1*-4
AGACCCAAAGTTTACACT(C/T)TGGG	Potri.001G119600	0	0	0	0
G(C/T)AGGCCATGTTTACAATATAGG	Potri.001G383500	0	0	0	0
CCCACA(A/G)AGTAAAGATGCTGGAC	Potri.004G120700	0	0	0	0
ACACCCCCAGTTTACACAGTGGG	Potri.016G015700	0	0	0	0
GTGGGCCATGTTTACAATGTTGG	Potri.011G103700	0	0	0	0

Supplemental Table T2.2. List of tentatively identified compounds detected as significantly different (P < 0.05) in UGT71L1-KO versus control plant lines[1].

Compounds (response area·10^{-4})	UGT71L1-KO	Partial-KO	Control
Salicin[s]	384.5 (±68.79)	1336.74 (±116.91)	1157.64 (±55.65)
Isosalicin[**]	65.42 (±4.21)	2.91 (±0.19)	0.71 (±0.03)
Salicoylsalicin[*]	15.51 (±2.18)	16.27 (±1.59)	11.76 (±1.46)
Salireposide[*]	45.81 (±5.4)	1.46 (±0.09)	1.38 (±0.09)
Salicin 7-sulfate[**]	1832.93 (±64.58)	452.2 (±10.28)	326.05 (±12.67)
2-O'Acetylsalicin[*]	15.66 (±0.71)	82.27 (±11.97)	118.75 (±10.81)
Salicortin[s]	1283.66 (±53)	12801.26 (±521.96)	15411.59 (±443.11)
HCH-Salicortin[**]	53.95 (±4.14)	8057.17 (±303.48)	9490.54 (±274.25)
Reduced salicortin (CID: 45783083, C20H26O10)[**]	414.18 (±8.49)	300.74 (±9.9)	164.77 (±2.02)
Tremulacin[s]	171.17 (±18.5)	17652.24 (±1266.28)	28012.5 (±1210.1)
Tremulacin isomer[**]	7.15 (±0.74)	651.48 (±66.46)	1143.56 (±40.83)
Hydroxytremulacin[**] or Homaloside D[**]	43.68 (±1.92)	235.39 (±12.83)	233.27 (±6.83)
Tremuloidin[s]	16.19 (±2.06)	171.85 (±19.26)	250.62 (±15.15)
Virgaureoside A[*]	51.05 (±2.11)	76.23 (±0.88)	104.46 (±3.43)
Populoside A[*]	136.92 (±8.63)	43.58 (±3.04)	35.92 (±1.04)
Populoside B or Tricharposide[*]	7.28 (±0.76)	34.18 (±4.37)	54.33 (±5.04)
Benzoate-glucose-2,3-hydroxy-2-ethoxy-4,6-cyclohexene[**]	234.28 (±17.89)	23.61 (±2.25)	13.1 (±0.59)
Benzyl 2-methoxy, 6-glucose-benzoate	416.86 (±21.14)	1150.1 (±78.92)	1614.86 (±44.1)
Flavonol bisglycoside (C34H26O11) [**]	129.7 (±8.37)	331.88 (±31.93)	265.05 (±33.31)
Luteolin 7-rutinoside[**]	61.51 (±2.4)	282.65 (±14.96)	223.7 (±13.82)
Kaempferol 3-glucuronide[*]	17.78 (±0.83)	117.26 (±8.97)	139.7 (±9.89)
Isorhamnetin glucose-rhamnoside[*]	90.72 (±5.5)	93.97 (±16.11)	86.54 (±9.09)
6,8-Dirhamnosylapigenin[*]	0.25 (±0.03)	92.37 (±3.85)	147.7 (±4.9)
Flavonol diglucoside[*]	0.16 (±0.02)	43.84 (±2.87)	57.44 (±3.3)
Apigenin 7-glucoside-4-*p*-coumarate[**]	41.93 (±2.9)	422.98 (±24.73)	568.41 (±19.4)
p-coumaroyl glucose flavone (CID: 101672254, C34H34O14) [**]	20.61 (±3.75)	7168.33 (±428.85)	8295.8 (±261.8)
Naringin 6-malonate[*]	6.39 (±1.32)	2216.75 (±123.14)	2596.06 (±77.8)

cis-caffeoylquinic acid**	16.9 (±0.95)	112.36 (±11.14)	193.04 (±9.93)
cis-coumaroylquinic acid**	4.76 (±0.27)	64.85 (±3.56)	65.75 (±5.56)
trans-caffeoylquinic acid**	136.86 (±7.37)	194.17 (±17.19)	204.8 (±8.34)
Chlorogenic acid isomer 1*	372.3 (±25.59)	564.52 (±50.6)	585.33 (±25.86)
Chlorogenic acid isomer 2*	13.2 (±1.01)	23.85 (±3.35)	29.64 (±3.22)
Chlorogenic acid isomer 3*	36.01 (±2.81)	26.79 (±10.87)	64.09 (±11.46)
Coumaroylquinic acid isomer 1**	8.62 (±0.78)	21.31 (±3.83)	22.4 (±2.42)
Coumaroylquinic acid isomer 2*	30.82 (±2.03)	60.1 (±14.11)	60.69 (±7.7)
Coumaroylquinic acid isomer 3*	30.28 (±1)	71.5 (±6.59)	33.08 (±1.46)
Coumaroylquinic acid isomer 4*	19.88 (±1.65)	65.74 (±3.01)	31.89 (±1.32)
Diferuloylquinic acid*	7.04 (±0.45)	310.32 (±14.57)	150.66 (±5.97)
Procyanidin B1 isomer 1*	1.38 (±0.09)	7.26 (±0.45)	10.31 (±1.23)
Procyanidin B1 isomer 2*	246.07 (±14.7)	1881.27 (±82.75)	2166.64 (±61.29)
Procyanidin B1 isomer 3*	40.09 (±4.42)	326.65 (±11.52)	351.02 (±10.81)
Procyanidin B1 isomer 4*	296.39 (±17.56)	1707.47 (±114.19)	2130.05 (±100.79)
Gallic Acid glucoside*	141.84 (±4.22)	65.15 (±7.8)	14.53 (±0.42)
Phenethyl caffeate*	217.54 (±11.74)	59.15 (±1.73)	42.71 (±1.34)
3-methoxy-4-hydroxy-2-glucose-trans-cinnamyl alcohol**	610.13 (±11.86)	582.76 (±12.59)	623.39 (±8.59)
Citroside isomer**	308.32 (±10.85)	383.13 (±18.75)	439.44 (±8.99)
2-phenylethyl xylose-glucoside**	82.84 (±5.11)	36.97 (±2.96)	43.22 (±2.8)
Modified 3,4 methoxycoumaroyl (C26H34O12, CID: 10052956) **	178.76 (±9.49)	270.28 (±26.1)	210.79 (±13.27)
Tanegool**	150.5 (±8.48)	903.95 (±24.03)	965.46 (±14.85)
Monophenolic biglucoside (C22H32O14) **	386.33 (±13.35)	408.07 (±24.79)	562.41 (±15.11)
Piperchabaoside A*	409.14 (±14.62)	471.37 (±35.72)	698.32 (±21.35)
Gentioluteol**	89.68 (±4.15)	433.72 (±14.32)	516.55 (±8.19)
Resveratrol 3,4'-diglucoside (C26H32O13)*	82.36 (±3.77)	103.77 (±1.87)	132 (±4.59)
Asebotin*	16.82 (±0.6)	46.36 (±5.58)	84.88 (±8.22)
Terniflorin**	85.03 (±3.75)	671.6 (±45.65)	876.09 (±40.7)

Methyl-galactose-2,3,6-benzoyl-4-galactoside**	7.41 (±0.57)	501.21 (±64.83)	378.58 (±23.89)
Hymenoside F (CID: 10723381, C35H38O15) **	0.47 (±0.13)	298.74 (±18.79)	407.68 (±12.12)
Salicylic acid glucoside*	238.66 (±10.09)	54.09 (±3.33)	39.07 (±1.52)
Grandidentatin**	1962.88 (±37.46)	1367.42 (±21.42)	1020.22 (±16.54)
Grandidentatin isomer 1**	1381.16 (±26.91)	801.82 (±16.15)	509.49 (±11.04)
Grandidentatin isomer 2**	247.46 (±20)	148.75 (±12.18)	85.11 (±6.05)
Reduced grandidentatin (CID: 102053308, C12H26O9) **	743.3 (±20.51)	400.97 (±7.6)	215.17 (±6.93)
Dihydroxycyclohexane glucoside**	3472.51 (±134.25)	2558.06 (±64.54)	2238.66 (±54.69)

[1] Features were detected through non-targeted metabolomic analysis using high-resolution mass spectrometry. Values indicate total combined average peak areas·10^{-4} (±SE) of all identified compounds for each type of phenolic compound. Each biological replicate is an average value of 3 technical replicates. UGT71L1-KO data are from a pool of 5 biological replicates for each of 3 independent lines; Partial-KO data represent 5 biological replicates of a single line; control data represents a pool of 3 independent empty vector lines each with 4-5 biological replicates and 5 biological replicates of a wild-type line. For compound identification [s] identifies compounds confirmed using an analytical standard, ** identifies compounds determined using MS2 fragmentation spectra and *in silico* analysis, and * identifies compounds tentatively determined with high resolution mass spectroscopy and database query.

Supplemental Table T2.3. Compounds[1] detected in young leaves using targeted metabolomic analysis (LC-MS/MS) using established methods (Lackus et al., 2020).

Compounds (Peak Area · 10^{-3})	UGT71L1-KO	Partial-KO	Control
Salicin-7-sulfate[s]	13876.9 (±807.39)	14517.65 (±654.9)	12418.06 (±377.08)
Salicin[s]	33.25 (±4.61)	149.04 (±4.88)	157.91 (±4.41)
Helicin[s]	4.39 (±0.66)	2.64 (±0.13)	1.91 (±0.1)
Salirepin[s]	18.83 (±1.44)	48.98 (±3.15)	52.55 (±1.59)
Salirepin-7-Sulfate[s]	789.54 (±46.39)	2025.11 (±110.38)	1895.55 (±76.89)
Salireposide[s]	17.52 (±1.52)	3.04 (±0.95)	3.14 (±0.19)
Salicortin[s]	178.5 (±83.16)	2813.45 (±76.07)	3185.16 (±96.53)
Tremulacin[s]	96.23 (±52.76)	1289.85 (±20.87)	1647.52 (±43)
Populin/ Tremuloidin#	1.86 (±0.78)	22.15 (±1.41)	30.29 (±1.14)
Trichocarpin[s]	1.55 (±0.12)	0.22 (±0.06)	0.29 (±0.01)
Apigenin[s]	21.84 (±2.18)	40.13 (±4.41)	51.5 (±1.06)
Catechin[s]	3.06 (±0.49)	2.5 (±0.46)	1.76 (±0.17)
Dihydrokaempferol[s]	0.18 (±0.03)	0.83 (±0.07)	1.37 (±0.07)
Kaempferol[s]	0.14 (±0.01)	0.22 (±0.03)	0.27 (±0.02)
Naringenin[s]	0.95 (±0.12)	3.76 (±0.26)	8.47 (±0.96)
Chlorogenic acid isomer 1#	570.4 (±107.27)	5140.92 (±199.66)	6652.46 (±171.94)
Chlorogenic acid isomer 2#	20 (±3.85)	194.65 (±7.65)	252.9 (±7.17)
Chlorogenic acid isomer 3#	17.84 (±0.66)	3435.14 (±861.96)	5352.69 (±145.92)
Coumaroylquinic acid isomer 1#	35.34 (±4.95)	285.2 (±11.44)	288.8 (±11.24)
Coumaroylquinic acid isomer 2#	347.18 (±27.79)	428.7 (±29.9)	642.81 (±37.49)
Coumaroylquinic acid isomer 3#	193.12 (±27.38)	1483.47 (±59.72)	1562.62 (±65.71)
Coumaroylquinic acid isomer 4#	7.86 (±1.21)	64.87 (±3.05)	67.01 (±3.01)
Cinnamic acid glucoside#	0.28 (±0.04)	0.16 (±0.04)	0.88 (±0.1)
o-coumaric acid glucoside#	128.56 (±9.77)	204.7 (±17.49)	199.06 (±12.16)
o-Coumaric acid[s]	9.46 (±1.36)	10.25 (±0.44)	12.16 (±1.42)
3,4-Dimethoxycinnamic acid[s]	5.25 (±0.41)	4.38 (±0.76)	3.71 (±0.49)
Gentisic acid glucoside#	1353.18 (±96.39)	229.09 (±26.8)	186.25 (±5.71)

2,3-Dihydroxybenzoic acid glucoside#	581.74 (±200.75)	50.92 (±3.35)	51.37 (±1.71)
3-Hydroxybenzoic acid[s]	3.7 (±0.47)	1.01 (±0.35)	0.72 (±0.07)
Benzoic acid glucoside[s]	6.26 (±0.57)	0.37 (±0.06)	0.36 (±0.03)
Benzyl alcohol glucoside[s]	0.69 (±0.04)	0.06 (±0.01)	0.07 (±0)
Benzoic acid[s]	5 (±1.16)	21.6 (±1.21)	14.33 (±0.87)
Salicyl alcohol[s]	7.71 (±0.68)	6.88 (±0.43)	4.26 (±0.4)
Caffeic acid[s]	31.23 (±2.01)	23.53 (±1.56)	21.48 (±1.43)
Ferulic acid[s]	13.25 (±0.84)	14.83 (±1.81)	16.86 (±0.52)
Gallic acid[s]	2.61 (±1.03)	13.00 (±4.12)	1.76 (±0.4)
Arbutin[s]	1.89 (±0.13)	3.33 (±0.19)	2.79 (±0.09)
Phenylpyruvic acid[s]	14.64 (±1.04)	31.25 (±3.92)	30.17 (±2.15)
Umbelliferone[s]	6.03 (±0.47)	1.47 (±0.37)	0.65 (±0.04)
Esculetin[s]	2.37 (±0.42)	3.18 (±0.24)	4.48 (±0.59)

[1] Values present mean peak area·mg $DW^{-1}\cdot10^{-3}$ (± SE). UGT71L1-KO data are from 3 independent transgenic lines (4 biological replicates per line); partial-KO data are from a single line (5 biological replicates); control data are from a pool of 3 independently generated empty vector lines (4 or 5 biological replicates per line) and a wild-type line (3 biological replicates). [s] Identity confirmed by authentic standards; # tentative compounds based on mass and fragmentation pattern

Supplemental Table T2.4. Differentially expressed genes (P-value ≤ 0.05) in leaves of three-month-old UGT71L1-KO plants compared to wild-type controls putatively identified to be involved with gibberellin metabolism.

Gibberellin Biosynthesis		UGT71L1-KO vs Wild-type		
Gene	Poplar Gene ID	Fold-change[1]	p-value	adj p-val
GA20 oxidase	Potri.005G184200	0.53	0.04	0.12
	Potri.007G103800	0.14	<0.01	<0.01
GA13 oxidase	Potri.019G064200	3.10	<0.01	<0.01
	Potri.008G026200	0.11	<0.01	0.01
	Potri.013G160800	4.89	0.02	0.08
GA3 oxidase	Potri.003G128100	2.38	<0.01	<0.01
	Potri.001G176600	0.38	<0.01	<0.01
	Potri.003G057400	11.88	<0.01	<0.01
GA2 oxidase	Potri.004G065000	1.87	0.02	0.08
	Potri.011G095600	2.36	0.04	0.12
	Potri.007G103800	0.14	<0.01	<0.01
	Potri.005G184200	0.53	0.04	0.12

[1] Values are an average of 3 biological replicates per line. Abbreviations: GA, gibberellic acid.

Supplemental Table T2.5. HPLC gradients used for separation and analysis of targeted metabolites.

Gradient	time (min)	solvent A (%)	solvent B (%)
gradient A	0.00	95.0	5.0
	0.50	95.0	5.0
	6.00	62.6	37.4
	6.01	20.0	80.0
	7.50	0.0	100.0
	9.50	0.0	100.0
	9.52	95.0	5.0
	12.00	95.0	5.0
gradient B	0.00	90.0	10.0
	0.50	90.0	10.0
	4.00	10.0	90.0
	4.02	0.0	100.0
	4.50	0.0	100.0
	4.51	90.0	10.0
	7.00	90.0	10.0
gradient C	0.00	97.0	3.0
	1.00	97.0	3.0
	2.70	0.00	100
	3.00	0.00	100
	3.10	97.0	3.0
	6.00	97.0	3.0

Supplemental Table 2.6. Parameters used for targeted metabolomic analysis (LC-MS/MS)[1].

Target compound	Q1 (m/z)	Q3 (m/z)	DP (V)	CE (V)	mode	HPLC gradient
2-Hydroxycinnamic acid	163	118.9	-30	-20	neg	A
4-Hydroxycinnamic acid	163	118.9	-30	-20	neg	A
Caffeic acid	179	134.9	-30	-22	neg	A
Ferulic acid	193.1	133.9	-30	-22	neg	A
Gallic acid	169	125	-30	-18	neg	A
Benzoic acid	121	121	-30	-5	neg	A
Salicylic acid	137	93	-30	-22	neg	A
3,4-Dimethoxycinnamic acid	207	103	-30	-20	neg	A
3-Hydroxybenzoic acid	137	93	-30	-16	neg	A
4-Hydroxybenzoic acid	137	93	-30	-20	neg	A
2,3-Dihydroxybenzoic acid	153	108/109	-30	-28	neg	A
2,5-Dihydroxybenzoic acid	153	108	-30	-28	neg	A
Salicyl alcohol	123	93	-30	-22	neg	A
Salicin	331	122.8	-30	-18	neg	A
Salicin-7-sulfate	365	97	-30	-25	neg	A
Salirepin	301	139	-30	-18	neg	A
Salirepin-7-sulfate	381	97	-30	-55	neg	A
Salicortin	422.8	123.1	-30	-7,5	neg	A
Helicin	329.1	121.1	-30	-18	neg	A
Salireposide	405	121	-30	-18	neg	A
Salicortin	422.8	123.1	-30	-30	neg	A
Tremulacin	527	123.2	-30	-34	neg	A
Apigenin	269	117	-30	-44	neg	A
Naringenin	271	151	-30	-46	neg	A
Catechin	288.8	109.1	-30	-32	neg	A
Kaempferol	285	93	-30	-46	neg	A
Dihydrokaempferol	287	125	-30	-28	neg	A
Chlorogenic acid – isomer	353	190.9	-30	-22	neg	A
Chlorogenic acid – isomer	353	173	-30	-22	neg	A
Chlorogenic acid – isomer	353	179	-30	-22	neg	A
Coumaroyl quinic acid – isomer	337	190.9	-30	-22	neg	A
Coumaroyl quinic acid – isomer	337	163	-30	-22	neg	A
Coumaroyl quinic acid – isomer	337	173	-30	-22	neg	A
Cinnamic acid glucoside	309	147	-30	-18	neg	A

Supplemental Table S2.6. continued

Hydroxycinnamic acid glucoside	325	163	-30	-10	neg	A
Dihydroxybenzoic acid glucoside	315	108	-30	-28	neg	A
Dihydroxybenzoic acid glucoside	315	109	-30	-28	neg	A
Benzoic acid glucose ester	283	121	-30	-25	neg	A
Benzyl alcohol glucoside	269	161	-30	-10	neg	A
Arbutin	271	108	-30	-32	neg	A
Phenylpyruvic acid	163	91	-30	-10	neg	A
Umbelliferone	161	133	-30	-26	neg	A
Esculetin	177	133	-30	-24	neg	A
SA	136.93	93	-20	-24	neg	B
SA-glucoside	299.128	136.9	-20	-18	neg	B
ABA	263	153.2	-20	-22	neg	B
JA	209.07	59	-20	-24	neg	B
JA-Ile	322.19	130.1	-50	-30	neg	B
OPDA	290.9	165.1	-20	-24	neg	B
OH-JA-Ile	338.1	130.1	-50	-30	neg	B
12-OH-JA	225.1	59	-20	-24	neg	B
COOH-JA-Ile	352.1	130.1	-50	-30	neg	B
Sulfo-JA	305	97	-20	-10	neg	B
D_4-SA	140.93	97	-20	-24	neg	B
D_6-ABA	269	159.2	-20	-22	neg	B
D_6-JA	215	59	-20	-24	neg	B
D_5-JA	214	59	-20	-24	neg	B
D_6-JA-Ile	328.19	130.1	-50	-30	neg	B
D_5-JA-Ile	327.19	130.1	-50	-30	neg	B
IAA	176	130	40	19	pos	B
D_5-IAA	181	135	40	19	pos	B
Alanine	90.1	44.1	20	17	pos	C
Serine	106	60.1	20	15	pos	C
Proline	116.1	70	20	19	pos	C
Valine	118.1	72.2	20	13	pos	C
Threonine	120.1	74.2	20	13	pos	C
Isoleucine	132.2	86.1	20	13	pos	C
Leucine	132.2	86.1	20	13	pos	C
Aspartate	134.1	74.1	20	19	pos	C
Glutamate	148.1	102.1	20	15	pos	C

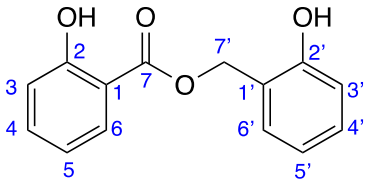
Supplemental Table 2.6. continued

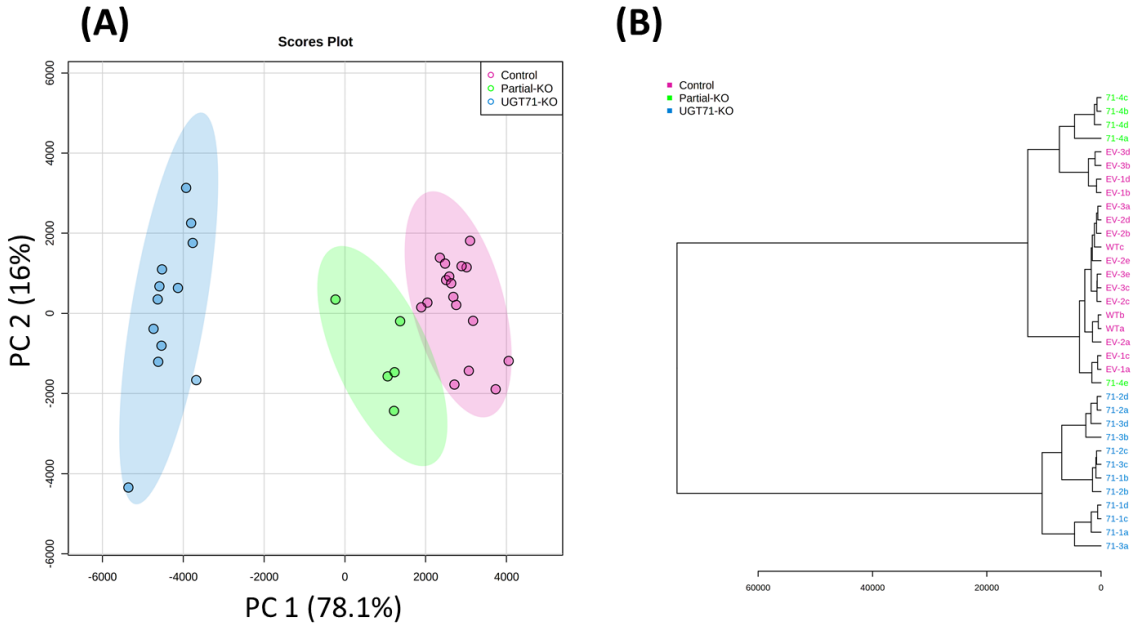
Methionine	150.2	104.1	20	13	pos	C
Histidine	156.2	110.1	20	17	pos	C
Phenylalanine	166.2	120.2	20	17	pos	C
Arginine	175.1	70.1	20	31	pos	C
Tyrosine	182.1	136.2	20	17	pos	C
Asparagine	133.1	74.1	20	19	pos	C
Glutamine	147.1	130	20	13	pos	C
Tryptophan	205.2	188.1	20	17	pos	C
Lysine	147.1	84.1	20	23	pos	C
S-methylmethionine	164.05	102.1	20	13	pos	C
$^{13}C,^{15}N$-alanine	94.1	47.1	20	17	pos	C
$^{13}C,^{15}N$-Serine	110	63.1	20	15	pos	C
$^{13}C,^{15}N$-proline	122.1	75	20	19	pos	C
$^{13}C,^{15}N$-valine	124.1	77.2	20	13	pos	C
$^{13}C,^{15}N$-threonine	125.1	78.2	20	13	pos	C
$^{13}C,^{15}N$-isoleucine	139.2	92.1	20	13	pos	C
$^{13}C,^{15}N$-leucine	139.2	92.1	20	13	pos	C
$^{13}C,^{15}N$-asparagine	139.1	77.1	20	19	pos	C
$^{13}C,^{15}N$-glutamate	154.1	107.1	20	15	pos	C
$^{13}C,^{15}N$-methionine	156.2	109.1	20	13	pos	C
$^{13}C,^{15}N$-histidine	165.2	118.1	20	17	pos	C
$^{13}C,^{15}N$-phenylalanine	176.2	129.2	20	17	pos	C
$^{13}C,^{15}N$-arginine	185.1	75.1	20	31	pos	C
$^{13}C,^{15}N$-tyrosine	192.1	145.2	20	17	pos	C
$^{13}C,^{15}N$-glutamine	154.1	136	20	13	pos	C
$^{13}C,^{15}N$-lysine	155.1	90.1	20	23	pos	C

[1] Details of the HPLC gradients are given in Supplemental Table T2.5 CE, collision energy; DP, declustering potential; Q1, quadrupole 1; Q3, quadrupole 3. SA, salicylic acid; JA, jasmonic acid; OH, hydroxy; OPDA, 12-oxo-phytodienoic acid; Ile, isoleucine; ABA, abscisic acid; Sulfo-JA: 12-hydroxy jasmonic acid sulfate, IAA: indole acetic acid

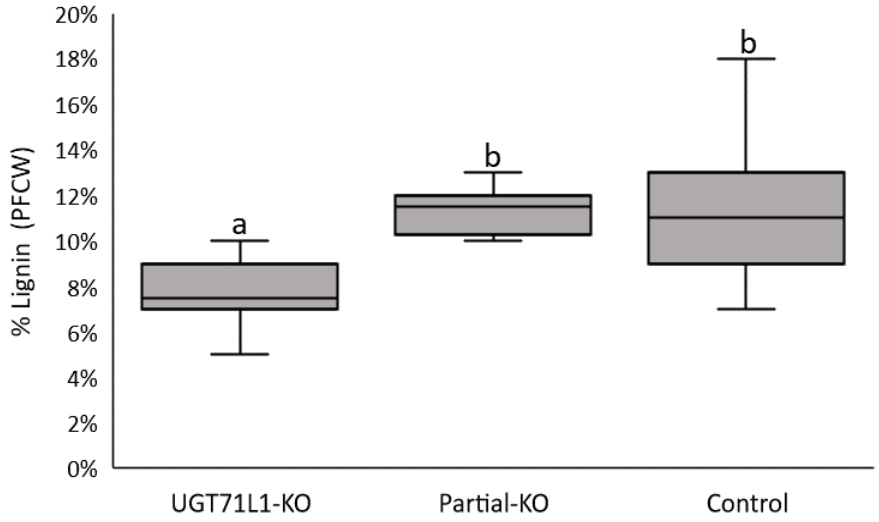
Supplemental Table 2.7. NMR analysis of salicyl salicylate showing chemical shift data (in $CDCl_3$)

position	1		
	δ_C	**δ_H**	***mult.* (J in Hz)**
1	112.0	-	-
2	161.9	-	-
2-OH	-	10.41	*s*
3	118.0	6.98	*m*
4	136.6	7.47	*ddd* (8.3, 7.3, 1.7)
5	119.5	6.88	*ddd* (8.0, 7.3, 1.0)
6	130.4	7.88	*dd* (8.0, 1.7)
7	171.6	-	-
1’	121.3	-	-
2’	155.4	3.59	*m*
2’-OH	-	7.18	*bs*
3’	117.8	6.95	*m*
4’	131.5	7.30	*ddd* (7.9,7.8,1.7)
5’	121.0	6.96	*m*
6’	132.3	7.37	*dd* (7.8, 1.7)
7’	63.9	5.42	*s*

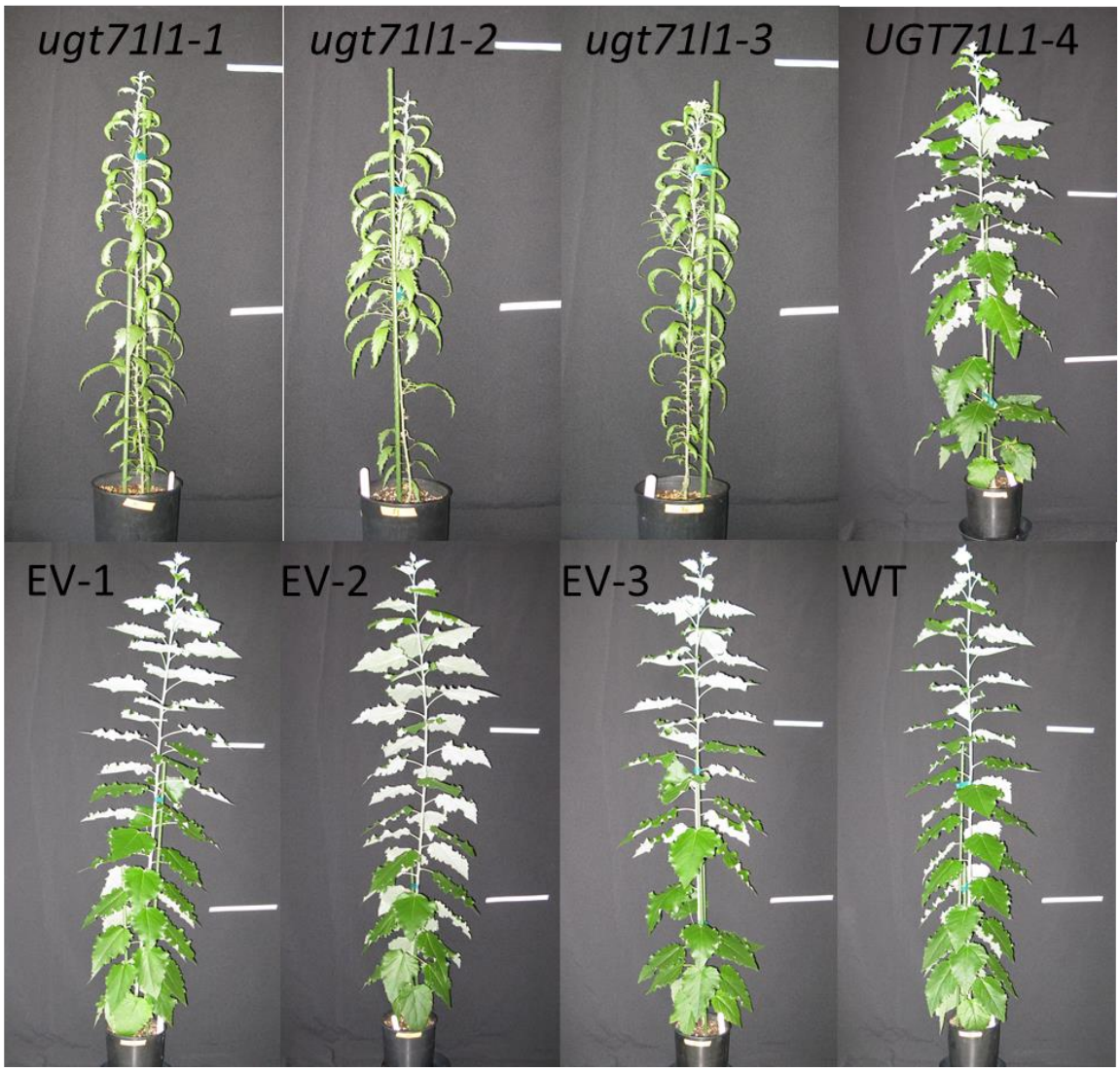




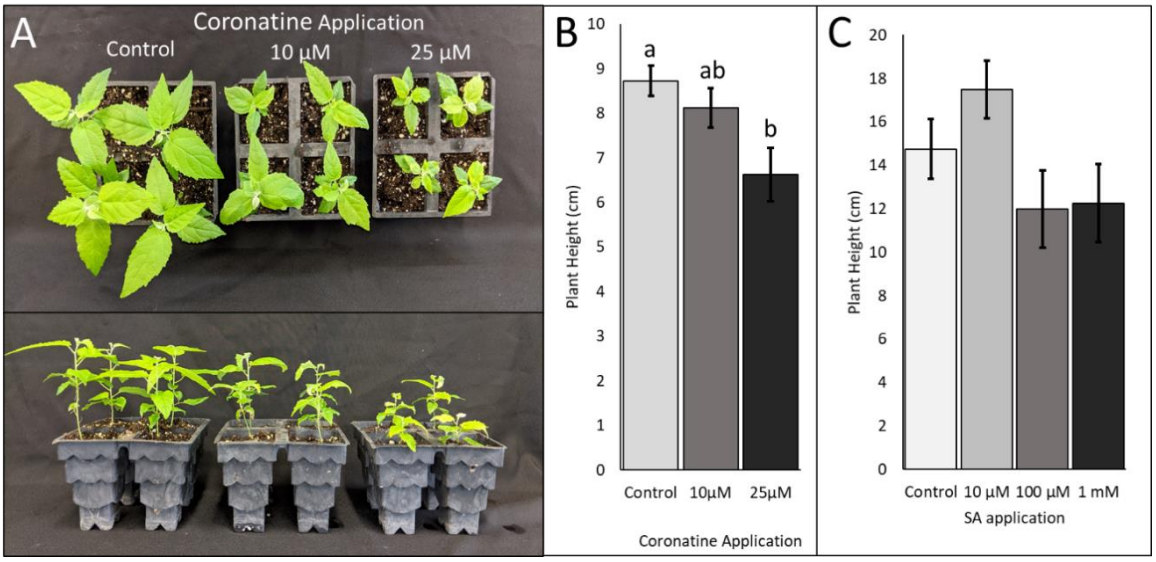
Supplemental Figure S2.1. Multivariate statistical analysis of-targeted metabolomic dataset. (A) Principal component score plots of non-targeted metabolomic analysis of UGT71L1-KO, partial-KO and control poplars. (B) Unweighted pair group method with arithmetic mean hierarchical cluster analysis of non-targeted metabolomic samples. Multivariate analysis was conducted using Metaboanalyst 5.0 (Chong et al., 2019). Blue colour indicates UGT71L1-KO lines, green indicates partial-KO lines, and magenta indicates control lines.



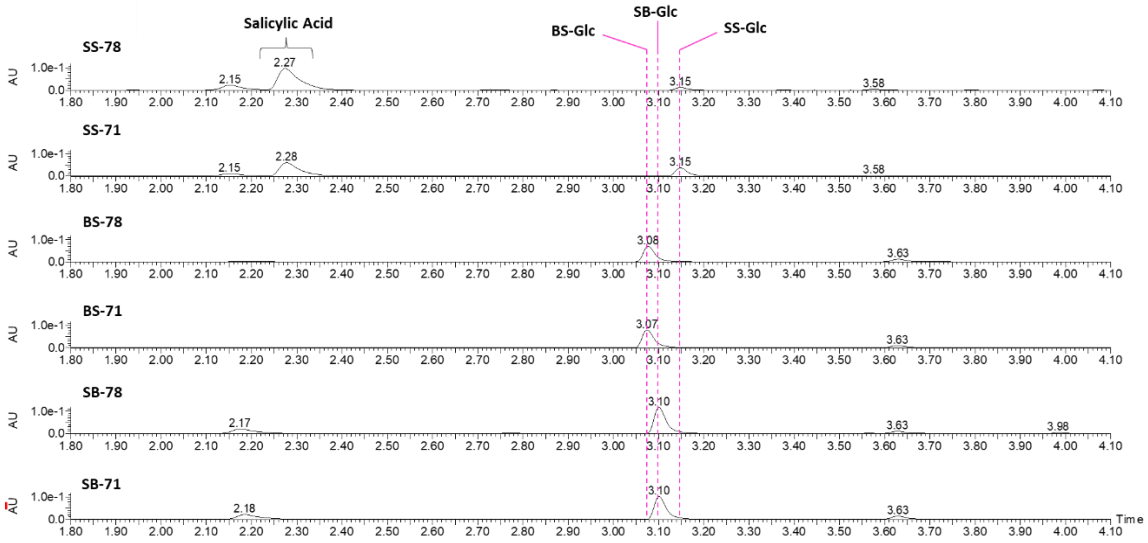
Supplemental Figure S2.2. Acetyl bromide lignin analysis comparing lignin concentrations (mg·g^{-1}) in the protein free cell wall (PFCW) extract of UGT71L1-KO and wild-type *P. tremula x P. alba* leaves. Control group consists of three independent empty vector control lines and one wild-type control line, n = 2-3 for each line with 1-2 technical replicates per line. The intermediate group represents one transformed line, n = 6, with 2 technical replicates each. UGT71L1-KO is data from three independent UGT71L1-KO line each with a biological replicate of n = 5-6 and 1-2 technical replicates per plant. Tukey's HSD test was performed, and the different letters represents significantly different ($P < 0.05$) concentrations.



Supplemental Figure S2.3. Whole plant photographs of 3.5-month old greenhouse grown knock-out, control, and wild-type poplars. Lines UGT71L1-1, -2, and -3 are biallelic knockouts. Line UGT71L1-4 is a monoallelic knockout (partial-KO). EV lines are independent empty vector transformants. WT is wild-type *P. tremula x P. alba* INRA 717-1B4. White markings indicate plant height markers (50 cm and 1 m).



Supplemental Figure S2.4. Effects of coronatine and SA on the growth of UGT71L1-KO and wild-type poplar. (A) Images of 5 week-old plantlets during mist-chamber acclimation and treated with water or coronatine solution. From left to right, groups of four plantlets are shown 5 weeks after weekly spray application of distilled water, 10 µM coronatine, or 25 µM coronatine. (B) Plant height (± SE) of mist chamber plantlets following five weeks of biweekly coronatine application. Data were analyzed by two-tailed t-test. Letters represent significance values from Tukey's HSD test. n = 8 for treatment plantlets and n = 11 for control plantlets. (C) Plant height (± SE) of mist chamber grown plantlets after five weeks of SA application.



Supplemental Figure S2.5. Chromatogram of UGT71L1 and UGT78M1 enzyme assay products separated by UPLC and observed at 280 nm. Glucosides were confirmed with ES(-) ionization m/z. Dashed magenta lines correspond to different elution times of BS-Glc, SB-Glc, and SS-Glc. Abbreviations: Glc, *o*-glucoside; SS, salicyl salicylate; BS, benzyl salicylate; SB, salicyl benzoate; 78, UGT78M1 recombinant enzyme; 71, UGT71L1 recombinant enzyme.

Chapter 3 – Benzenoids are intermediates in salicinoid and salicylic acid biosynthesis in *Populus*

This chapter has been written with the intent of publication in a plant biochemical journal. Harley Gordon designed and conducted experiments and wrote the manuscript. C. Peter Constabel contributed to experimental design and edited the manuscript. The laboratory of Jeremy Wulff at the University of Victoria prepared all deuterium labelled compounds used in this study. A full publication draft will include methodology of deuterated substrate synthesis and NMR substrate confirmation.

3.1 Introduction

Salicyl alcohol glucosides, known as salicinoids or phenolic glycosides, are specialized metabolites produced by poplar and willow trees (Salicaceae). Salicinoids are biologically active and abundant compounds that function as potent herbivore defense molecules (Lindroth and St. Clair, 2013; Barker et al., 2019). As climactic and anthropogenic disturbances alter the habitat and distribution of poplar trees, the production of salicinoids will continue to have a significant influence on poplar adaptation (Ackerman and Breen, 2016; Rogers et al., 2020; Marchais et al., 2022). The traditional medicinal practice of using poplar and willow bark as a pain reliever has been occurring for over 3,500 years (Montinari et al., 2019). The analgesic properties of poplar and willow bark is caused by the salicinoid salicin. The salicyl alcohol component of salicin can chemically convert into salicylic acid upon human consumption (Vlot et al., 2009). Salicylic acid is a pain reliving molecule and was the biological precursor to the synthetic medicine acetylsalicylic acid (Aspirin) (Desborough and Keeling, 2017; Zhang and Li, 2019). Salicinoids are important molecules ecologically and medicinally; however, our understanding of the salicinoid biosynthetic pathway remains incomplete.

Radioisotopically labelled substrates have been used to study plant specialized metabolite biosynthesis for over 50 years (Zenk, 1967; Messiano et al., 2009; Jacobowitz and Weng, 2020). Using radioisotope labelled salicylic acid (SA) in *Salix purpurea* it was

found that salicin and salicyl alcohol are not synthesized from salicylic acid, but rather coumaric or cinnamic acid (Zenk 1967). It was also found that BA and benzyl alcohol can act as biosynthetic precursors to salicyl alcohol (Babst et al., 2010). Salicin biosynthesis is therefore derived from the central phenylpropanoid pathway. Salicin is glucosylated on the hydroxy group and contains an unmodified ethoxy group on the orthologous position. The complex and abundant salicinoid salicortin contains a salicin moiety (Fig. 3.1); however, salicortin and salicin have distinct biosynthetic routes (Gordon et al., 2022; Chapter 4). For example, isotopically labelled salicyl alcohol can efficiently incorporate into salicin, but does not incorporate into salicortin (Babst et al., 2010). However, administration of salicyl alcohol results in the formation of isosalicin, which is a salicyl alcohol glucosylated on the ethoxy rather than the hydroxy group (Fig. 3.1). Isosalicin is normally only a minor component in willow and poplar extracts; however, the exogenously supplied salicyl alcohol increases the isosalicin to salicin ratio (Palo 1984; Payyavula et al., 2009; Babst et al., 2010). Introduction of labelled benzyl alcohol results in 99% salicin compared to isosalicin, which is the natural ratio in poplar (Zenk, 1967; Babst et al., 2010; Gordon et al., 2022). For the conversion of benzyl alcohol to salicin, hydroxylation must occur prior to glucosylation, as the glucosylation benzyl alcohol followed by hydroxylation would result in isosalicin. Thus, salicyl alcohol is an intermediate in salicin biosynthesis, but the conversion of benzyl alcohol into salicyl alcohol may be closely linked to glycosylation, as free salicyl alcohol results in an abundance of isosalicin, which is not seen with benzyl alcohol application (Babst et al., 2010). The altered accumulation of artificial isomers after precursor administration can provide insight into the sequence and connections of salicin biosynthesis.

Exogenous benzyl alcohol and salicyl alcohol can both be converted into salicin in planta; however, the direct biosynthetic steps and enzymes are unestablished (Zenk, 1967; Babst et al., 2010). Salicortin is a widespread complex salicinoid (Fig. 3.1). Salicortin has a two-ring dibenzenoid structure with a 6-hydroxy-2-cyclohexen-onoyl (HCH) moiety ester linked to salicyl alcohol. C^{13} labelled benzoate, benzaldehyde, and cinnamic acid can all be incorporated into the salicyl alcohol and the HCH moiety of salicortin (Fig 3.1B). Benzyl alcohol; however, appears to only incorporate into the salicyl alcohol moiety of salicortin.

Gene coexpression analysis of two poplar acyltransferase enzyme benzoyl-CoA: salicyl alcohol O-benzoyltransferase (SABT)and Benzoyl-CoA: benzyl alcohol O-benzoyltransferase (BEBT) led to identification of the salicinoid specific glycosyltransferase gene designated UGT71L1 (Chedgy et al., 2015; Fellenberg et al., 2020). UGT71L1 can glucosylate salicyl benzoate (SB), which is a product of SABT. Functional knockouts of UGT71L1 nearly eliminate salicortin biosynthesis in whole plants (Fellenberg et al., 2020; Gordon et al., 2022). Benzyl benzoate (BB) and SB are both hypothesized dibenzenoid intermediates involved in salicinoid biosynthesis (Babst et al., 2010; Fellenberg et al., 2020; Gordon et al., 2022). Furthermore, the glucosylated dibenzenoid salicyl benzoate glucoside (SB-Glc) is a natural metabolite in poplars and willows and may be a salicortin biosynthetic intermediate (Kulasekaran et al., 2020). Despite compelling in vitro activity, to date there is no direct evidence for the involvement of the dibenzenoids benzyl benzoate or salicyl benzoate in salicinoid biosynthesis. Benzyl benzoate and salicyl benzoate may both be starting points for salicortin biosynthesis whereas salicyl benzoate may also be a biosynthetic intermediate (Babst et al., 2010; Chedgy et al., 2015; Fellenberg et al., 2020).

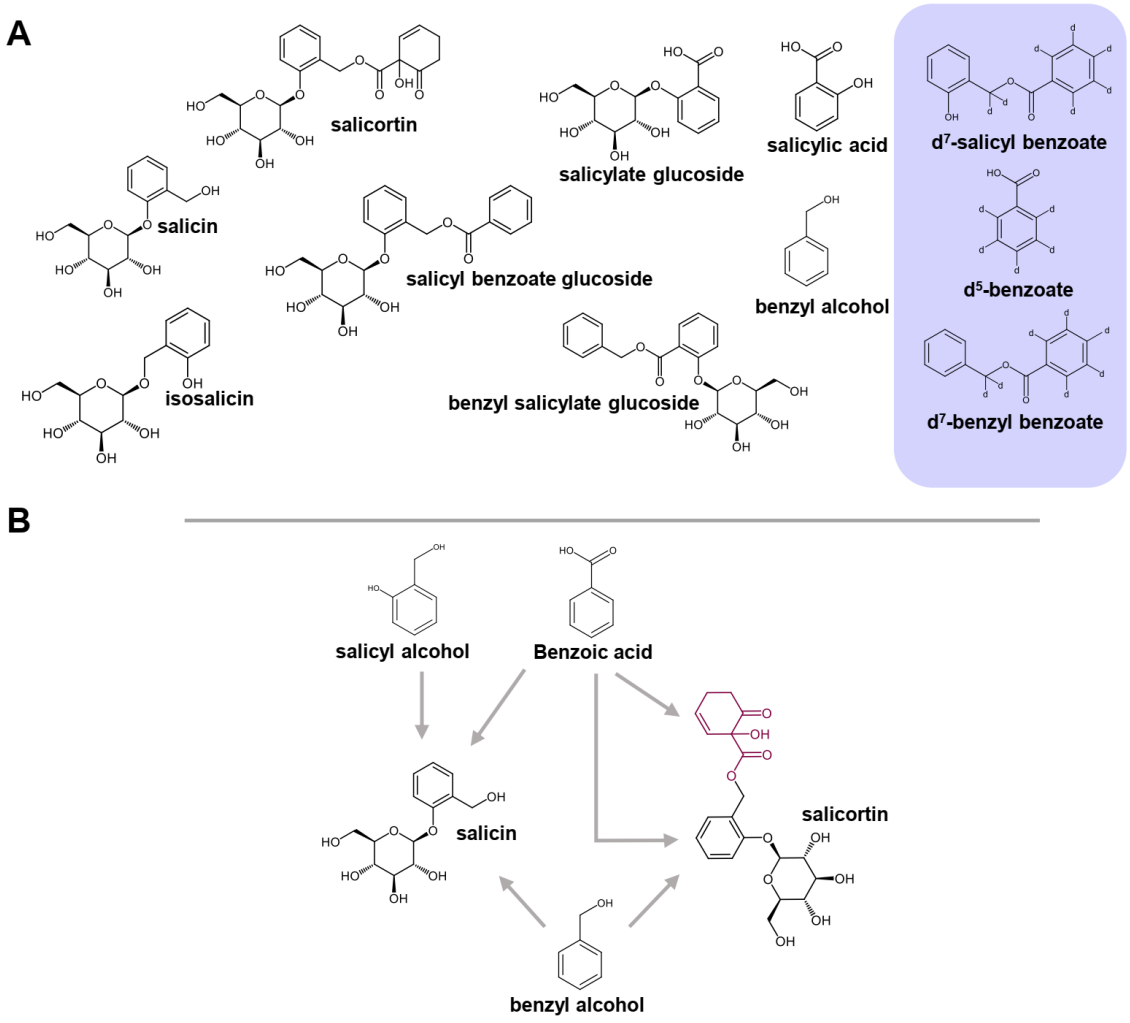


Figure 3.1. Molecular diagrams of metabolites referenced within this study and a schematic of phenolic metabolite incorporation into salicortin. A, benzenoid and salicinoid compounds discussed in this study. Compounds with a shaded background are deuterated substrates used in this study. B, model of the biosynthetic connections of simple phenols to salicortin and salicin biosynthesis based on isotopic labelling data (Zenk 1967; Babst et al., 2010). Grey arrows represent isotopic connections inferred from prior experimental results. Arrow position indicates which ring of salicortin the monophenolic precursor can contribute to. The HCH ring of salicortin is designated with a purple structure.

The analysis of deuterated benzenoid molecules applied to poplar tissue and incorporated into salicinoids was used to investigate the role of salicyl benzoate or benzyl benzoate as salicortin biosynthetic intermediates. Selectively deuterated benzyl benzoate (d^7-BB), salicyl benzoate (d^7-SB), and benzoic acid (d^5-BA) were obtained and used to study the pattern of label incorporation following administration in *P. tremula x P. alba* wild-type (WT) and UGT71L1-knockout (KO) plants (Fig. 3.1A). The poplar UGT71L1-KO lines have strongly reduced phenylpropanoid concentrations, a significant growth reduction, and a 10-fold increase in salicylic acid (SA) concentrations (Gordon et al., 2022). In Chapter 2 we hypothesized that overaccumulation of the phytohormone SA may result from an unstable dibenzenoid intermediate such as salicyl salicylate building up due to interrupted salicinoid biosynthesis. However, in the standard biosynthetic pathways SA is synthesized from isochorismate as a branch of the shikimate pathway or directly from the phenylpropanoid pathway (Zhang and Li., 2019). The contribution of each SA biosynthetic pathway, and the specific biosynthetic route of phenylpropanoid derived SA in poplar remains an ongoing area of study (Xue et al., 2013). Our dibenzenoid labelling experiments aim to demonstrate a dibenzenoid salicinoid intermediate contribution to SA accumulation in poplar.

Using exogenous application of deuterated benzoate (BA), salicyl benzoate (SB), and benzyl benzoate (BB) we aimed to establish the incorporation patterns of applied benzenoids into salicinoid molecules. We aimed to test the contribution of BB and SB to salicortin biosynthesis. Through UPLC-MS analysis of salicinoid products within labelled tissue, we found that BB, but not SB, was able to provide the dibenzenoid core of salicortin. Furthermore, SA overaccumulation in UGT71L1-KO lines is likely from BA, rather than from the accumulation of an unstable dibenzenoid intermediate.

3.2 Materials and Methods

3.2.1 Plant Growth and substrate application

For leaf disc substrate administration three leaf discs 1 cm in diameter and without prevalent midveins were collected from the expanding leaves of one-month old wild-type *P. tremula x P. alba* (INRA 717-1B4) clones. All plants were grown in a greenhouse with supplemental lighting providing a 16 h photoperiod with an average light intensity of 300 $\mu mol \cdot m^{-2} \cdot s^{-1}$ at an average temperature of 23 ºC. Plants were watered as required and the one-gallon pots contained slow release fertilizer in Sunshine Mix 4 soil (Sungro, Seba Beach, AB, Canada) as previously described (Gordon et al., 2022). All deuterium labelled substrates (d^5-BA, d^7-BB, d^7-SB) were provided by Dr. Jeremy Wulff. Labelled substrate solutions were prepared from a 100 mM DMSO stock solution. Water containing no substrate but an equivalent volume of DMSO was used as the control solution. Two different substrate administration methods were used, one for leaf disc infiltration and one for leaf petiole immersion. For leaf disc infiltration, the leaf discs were aspirated with 0.5 mM of d^7-SB or d^7-BB using a syringe until they were visibly saturated. Aspirated discs were placed in a plastic screw cap vial containing 1 mL of 0.5 mM labelled substrate to ensure complete leaf disc immersion. Sealed vials were placed on a greenhouse table with an average light intensity of 75 $\mu mol \cdot m^{-2} \cdot s^{-1}$ for 4 h. This experiment was repeated in triplicate. For leaf petiole substrate administration, fully expanded leaves from three-month old *P. tremula x P. alba* wild-type (WT) and UGT71L1-KO lines were used. Deuterated substrates were administered to petioles by immersing a fully expanded leaf petiole from WT or UGT71L1-KO poplar lines in water containing 1 mM of d^7-BB, d^7-SB, or d^5-BA. The leaf with the immersed petiole was left under ambient bench-top light (approx. 6 $\mu mol \cdot m^{-2} \cdot s^{-1}$) for 24 h. The 1 mM substrate solution was replenished as required throughout the experiment. The experiment was repeated on two separate days with a total of five biological replicates per treatment.

3.2.2 Metabolite analysis by Liquid Chromatography – Mass Spectroscopy

Following substrate administration, leaf discs were blotted dry, and metabolites were extracted. Leaf discs were homogenized in a 1.5 mL microcentrifuge tube using a pestle and 80% methanol in water (v/v) using 10 $\mu L \cdot mg^{-1}$ fresh weight (FW). Homogenized discs were sonicated for 10 min in a bath sonicator and shaken gently for 20 min. Samples were centrifuged for 2 min at 12,000 *g* and the supernatant was applied to a preconditioned 33u Phenomenex 200mg/3mL Strata-X column (Phenomenex, Torrance, CA, USA). The Strata-X column was preconditioned with 1 mL methanol followed by 1 mL water. The extract was washed with 2 mL of water and eluted with 2.5 mL of a 50% methanol-water solution (v/v) containing 0.1% formic acid. The eluant was dried under vacuum at ambient temperature and resuspended in 100% methanol at a ratio of 6 $\mu L \cdot mg^{-1}$ leaf disc FW.

Petioles were separated from the leaf lamina following 24 h of immersion in the substrate solution. Methanol was applied at a rate of 10 $\mu L \cdot mg^{-1}$ tissue FW to petioles and the tissue was homogenized in a microcentrifuge tube. The methanol-petiole solution was briefly vortexed and sonicated in a bath sonicator for 5 min and shaken gently for 30 min. Extracts were centrifuged at 10,000 *g* for 5 min and the supernatant was stored at -20 ºC prior to analysis.

Compounds were detected and analyzed using previously described methods on an Acquity UPLC-MS (Waters, Mississauga, ON, Canada) (Chapter 4). In brief, a 5 µL injection was separated using a Waters Acquity Ethylene Bridged Hybrid (C18, pore size: 1.7 µm, dimensions: 2.1 x 50 mm) column at 40 ºC. Binary solvent gradient separation was used with water containing 0.1% formate (A) and LC/MS grade acetonitrile with 0.1% formate (B). The flow rate was 0.5 $mL \cdot min^{-1}$. The separation gradient was 99.9% A, 0-0.5 min; 99.9-80.0 % A, 0.5-6.0 min; 80.0-50.0% A, 6.0-9.0 min; 50.0-10.0% A, 9.0-10.0 min; 10% A, 10.0-11.5 min; 10-99.9% A, 11.5 – 12.0 min; 99.9% A, 12.0-13.0 min. A single quadrupole MS in electrospray ionization negative mode was used to detect and quantify all analytes. The MS conducted a full scan from 70.00 to 850.00 m/z with a continuous cone voltage of 10 V. Compound identities were assigned using an in-house library based on standards (salicin, salicortin, SA, BB), enzyme assay reactions (SB-Glc,

BS-Glc, SS-Glc, BA-Glc, SA-Glc), m/z values, and retention times. A table of m/z values and retention times used in this study is presented in Supplemental Table T3.1. Peak integration was conducted using MassLynx 4.2 (Waters) and data is presented as MS peak area·10^{-3}.

3.2.3 Metabolite analysis by GC/MS

For gas chromatography – mass spectroscopy (GC/MS) analysis leaf disc infiltration and substrate administration was conducted as per above, except that the leaf discs were blotted dry and macerated in *n*-hexanes applied at 5 µL·mg^{-1} leaf disc FW within glass vials. Tissue was sonicated in a bath sonicator for 5 min and briefly vortexed. Extracts were then stored at 4 ºC overnight followed by centrifugation at 3,300 *g* for 5 min. The separated organic layer was dispensed into glass vials and stored at -20 ºC prior to analysis.

Hexane extracts were analyzed using a TRACE 1300 GC/MS (Agilent, Mississauga, ON, Canada). A 2 µL sample was injected onto a glass wool liner using a split-less injection and applied to a TraceGOLD TG-5SilMS GC column (0.25 mm x 0.25 µm x 30 m) (Thermo Scientific, Waltham, MA, USA) with a helium carrier gas flow rate of 1 ml·min^{-1}. The front inlet was 250 ºC, and oven gradient conditions were as follows: 0.0 – 2.0 min, 60 ºC; 2.0-7.0 min, 60-120 ºC; 7.0-21.5 min, 120 -240 ºC; 21.5 – 22.5 min, 240 -300 ºC; 22.5 – 24.5 min, 300 ºC. The transfer line was 150 ºC and the MS ion source was 270 ºC, with a continuous scan from 50-600 atomic mass units. Data was analyzed using Xcalibur software (Thermo Scientific).

3.3 Results

3.3.1 Exogenously administered dibenzenoid molecules can enter plant metabolic pathways

The labelled dibenzenoids salicyl benzoate (SB) and benzyl benzoate (BB) (Fig. 3.1A) were synthesized with differentially deuterated rings, which allowed for detection of the incorporation of both rings as separate entities. The alcohol component of SB or BB was labelled with two deuterium atoms on the ethoxy group (Fig 3.1A). Incorporation of the alcohol component results in a +2 mass addition on analytes and is not impacted by aromatic ring substitution. The BA moiety is deuterated on the aromatic ring; therefore, incorporation of a fully labelled BA results in a +5 mass addition. Deuterium labels on the aromatic ring allows for the inference of ring substitution. Addition of a hydroxyl group on BA; for example, results in a +4 mass addition. The HCH moiety of salicortin would result in the loss of at least one deuterium and therefore a +4 mass addition if derived from a labelled BA. Deuterium labelled SB and BB both have a +7 mass addition to the monoisotopic mass of their unlabelled form.

To confirm uptake of labelled substrates we analyzed leaf discs using GC/MS to directly detect benzyl benzoate. d^7-BB was detected in leaf extracts following 4 h of d^7-BB administration. Therefore, d^7-BB was successfully taken up into leaf discs (Supplemental Figure S3.1). Administered salicyl benzoate was not detected using GC/MS analysis of leaf disc samples; however, the presence of abundant salicyl benzoate glucoside (SB-Glc) with a +7 mass was detected using LC/MS. Detection of SB-Glc with a +7 mass in d^7-SB administered petioles or leaf discs demonstrated that d^7-SB is taken up by plant tissue and enters metabolism (Table 3.1). While the presence of deuterated d^7-BB and +7 SB-Glc was observed, we also noted that a portion of d^7-BB and d^7-SB were hydrolyzed at the ester bond within plant tissues based on the detection of +2 and +5 mass incorporation in salicinoid products. The ester bond cleavage during metabolic processing of administered substrates may represent a natural enzymatic process or an artifact of exogenous substrate application (Table 3.1). The release of labelled benzoate, benzyl alcohol, and salicyl alcohol provides insight into benzenoid metabolic network.

Table 3.1. Isotope concentration fold-change in petioles from wild-type poplar supplied with deuterated substrate[1]

	Salicortin			**Salicin**			**Isosalicin**		
Isotope	*d^7-SB*	*d^7-BB*	*d^5-BA*	*d^7-SB*	*d^7-BB*	*d^5-BA*	*d^7-SB*	*d^7-BB*	*d^5-BA*
0	1.0	0.8	1.1	1.1	**1.6***	**1.7***	1.6	2.4	1.3
+2	1.0	1.0	1.0	**7.5***	**2.9****	1.7	**149.5*****	**4.0***	3.4
+4	**5.6***	**3.2***	**7.5****	**74.2***	23.2	**114.6****	**16.6*****	4.3	10.2
+5	15.3	**49.5***	**28.7****	-	-	-	-	-	-
+6	10.2	**28.6***	9.5	-	-	-	-	-	-
	SB-Glc			**BS-Glc**			**SA-Glc**		
Isotope	*d^7-SB*	*d^7-BB*	*d^5-BA*	*d^7-SB*	*d^7-BB*	*d^5-BA*	*d^7-SB*	*d^7-BB*	*d^5-BA*
0	0.6	0.2	1.4	0.9	0.4	1.3	1.1	0.6	0.7
+2	**10.2***	0.3	1.5	1.2	6.1	1.4	2.1	1.1	2.4
+4	-	-	-	1.2	0.8	0.9	32.1	12.7	18.3
+5	**97.9****	**17.6***	**83.5***	**13.9****	**18.0***	7.4	-	-	-
+6	-	-	-	3.5	**12.1***	1.6	-	-	-
+7	**997.9*****	9.5	**4.9***	-	-	-	-	-	-

1. Fold-change calculated from the average response area (n = 5) of petioles with labelled substrates over the average isotope response area of control samples (n = 5). Bolded values represent significant differences (α = 0.1) as determined by Welch's *t*-test. Asterisks represent * P < 0.1, ** P < 0.05, ***P < 0.01. Data with no asterisks is non-significant.

3.3.2 Exogenously supplied benzyl benzoate but not salicyl benzoate is incorporated into salicortin

In leaf discs administered with d^7-BB, a +6 label was significantly incorporated into salicortin (P = 0.025) (Fig. 3.2A). The +6 mass indicates that both the alcohol and acid phenolic rings of d^7-BB were incorporated into salicortin. Comparatively, infiltration of d^7-SB resulted in no significant detection of +6 salicortin (Fig. 3.2B). Therefore, label administration into leaf discs suggested that BB is likely an intermediate or entry point into salicortin biosynthesis, while SB does not appear to enter the pathway. By contrast, enrichment of +4 salicortin after d^7-BB or d^7-SB administration indicated the incorporation of the d^5-BA moiety of these dibenzenoids into either the salicyl alcohol or

HCH moiety of salicortin. Both d^7-BB and d^7-SB contributed to the formation of +4 salicortin at similar efficiencies in leaf discs (Fig 3.2C and Fig 3.2D). This indicated that the d^5-BA component is released from d^7-SB and d^7-BB and can contribute to salicortin biosynthesis. These data also demonstrated the ($P = 0.08$) incorporation of benzyl alcohol into +2 salicortin, whereas the salicyl alcohol released from SB did not lead to enriched +2 salicortin (Fig. 3.2E and Fig. 3.2F). This further emphasizes the central role of benzyl benzoate in salicortin biosynthesis.

We examined a second tissue type to provide a comparison for label incorporation patterns in two distinct tissues. Leaves with petioles were placed in deuterated substrate solutions for 24 h, and the petioles were analyzed for deuterium label incorporation patterns. Salicortin label incorporation exhibited a similar pattern to leaf discs. d^7-BB contributed to the accumulation of +6 salicortin but d^{7-}SB did not (Table 3.1, Fig. 3.3). Administration of d^5-BA and d^7-BB both appeared to generate +5 salicortin. The +5 salicortin that resulted from d^7-BB administration is unexpected but may be explained by the loss or replacement of two deuterium atoms during formation of the HCH moiety. The administration of d^5-BA, d^7-SB, or d^7-BB all resulted in significant +4 salicortin. Salicin appeared to incorporate a +2 or a +4 mass addition following d^7-SB or d^7-BB administration representing incorporation of either ring. d^5-BA results in the significant formation of +4 salicin. Like salicin, isosalicin also accumulated a +4 mass and a significant ($P = 0.002$) +2 label after administration of d^7-SB. Isosalicin is not naturally found in significant concentrations in poplar tissues. We also observed a significant accumulation of +7 SB-Glc in d^7-SB administered petioles. All administered substrates resulted in an increased concentration of +5 SB-Glc, ostensibly from the incorporation of free d^5-BA. Only d^7-SB contributed to a significant accumulation of +2 SB-Glc. For benzyl salicylate glucoside (BS-Glc) only d^7-BB contributed to the significant accumulation of a +5 and +6 label incorporation (Table 3.1). An expanded analysis of label incorporation patterns for dibenzenoid glucosides is presented in Section 3.3.4. To summarize, leaf disc and petiole administration of labelled d^7-BB and d^7-SB showed that BB can enter the salicortin biosynthetic pathway directly, but SB cannot. Although SB may theoretically be a bound biosynthetic intermediate in salicortin biosynthesis after

hydroxylation of BB, exogenous application of SB does not appear to enter salicortin biosynthesis.

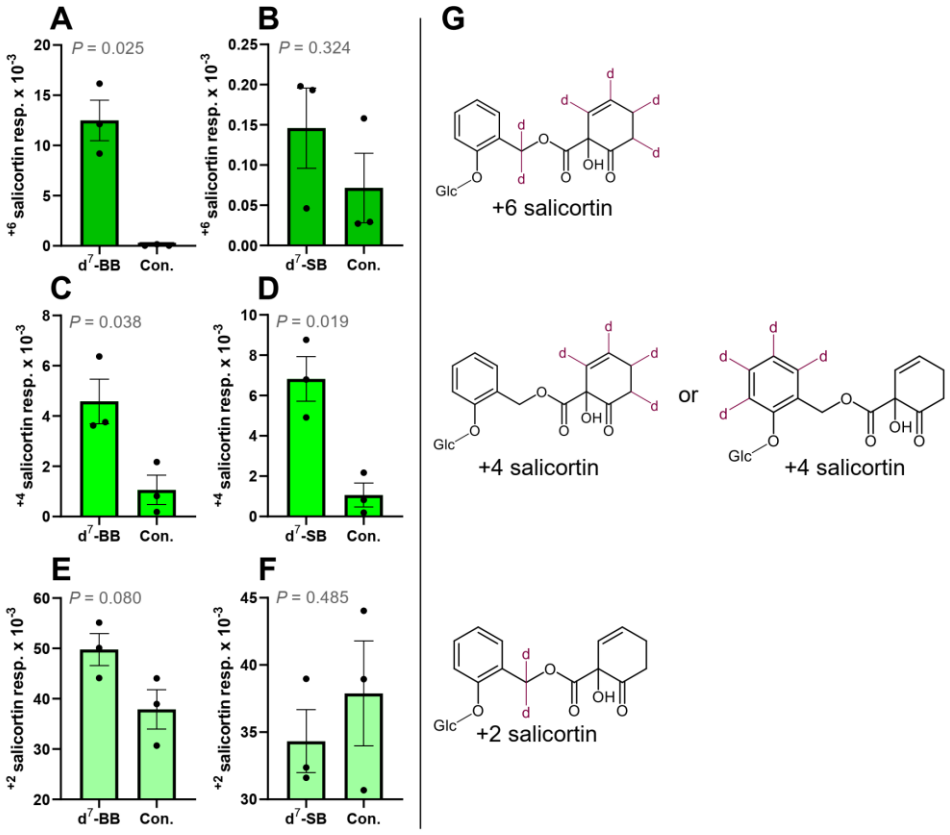


Figure 3.2. Labelled benzyl benzoate and salicyl benzoate incorporation into salicortin. Results are the deuterium incorporation patterns into salicortin in leaf discs from WT poplar. Incorporation is shown as mass detector peak area (response) of the extracted ion chromatogram. Graphs are the relative quantification of +6 salicortin following administration of d^7-BB (A) or d^7-SB (B). Relative quantification of +4 salicortin following d^7-BB (C) or d^7-SB (D) administration. Relative quantification of +2 salicortin following d^7-BB (E) or d^7-SB (F) administration. Error bars are ± SE, n = 3 experimental replicates. P values represent a Welch's *t*-test between responses of label administered and control leaf discs. G, salicortin structural diagrams with the inferred deuterium positions for +6, +4, and +2 label incorporations corresponding to graphs on the left. Deuterium atoms are drawn in purple, no hydrogens are included. Abbreviations: BB, benzyl benzoate; SB, salicyl benzoate; resp, response; Con, control; Glc, glucose.

3.3.3 Salicin can be synthesized from alcohol and acid moieties of benzyl benzoate and salicyl benzoate in poplar

As seen in above (Table 3.1) the administration of d^7-SB or d^7-BB resulted in significant label incorporation into +2 salicin in WT leaf petioles (Fig. 3.3A and Fig. 3.3B). The +2 label likely derived from the released alcohol moiety of the administered d^7-SB or d^7-BB. Labelled salicyl alcohol from d^7-SB incorporated into salicin and resulted in 10-fold more +2 salicin compared to d^7-BB administered petioles further suggesting SB is easily hydrolyzed but BB is not. The incorporation pattern indicates that released salicyl alcohol from SB provides one route to salicin. The benzyl alcohol component of BB can provide a precursor to salicortin or salicin. The detection of +4 salicin after d^7-SB and d^7-BB administration indicates that d^5-BA released from d^7-SB or d^7-BB likely incorporates into salicin ($P = 0.086$ for d^7-SB and $P = 0.112$ for d^7-BB) (Fig. 3.3C and Fig. 3.3D). We wanted to check the incorporation pattern in UGT71L1-KO plants, as they make a minimal amount of salicortin, but retain about 20% of salicin accumulation. The pattern of +2 and +4 salicin accumulation after d^7-SB and d^7-BB administration is similar in UGT71L1-KO and WT lines (Supplemental Figure S3.3A and Supplemental Figure S3.4A). In both UGT71L1-KO and WT petioles +2 isosalicin accumulated following d^7-SB administration from the release of labelled salicyl alcohol (Supplemental Figure S3.3B). Both d^7-SB and d^7-BB labelling resulted in the accumulation of +2 and +4 isosalicin; however, incorporation rates were lower and not significant in d^7-BB administered petioles (Supplemental Figure S3.3B and Supplemental Figure S3.4B).This may be due to the absence of released salicyl alcohol. The +4 isosalicin observed is most likely derived from released d^5-BA.

Our data suggests the BA moiety from administered SB and BB can contribute to salicortin and salicin biosynthesis (Table 3.1). Therefore, we tested directly if administered d^5-BA is incorporated into +4 salicin and +4 salicortin in WT petioles. Direct application of d^5-BA significantly contributed to the strong enhancement of +4 salicin and +4 salicortin (Fig. 3.3E and Fig. 3.3F). Therefore, it appears likely that the +4 labelled salicortin detected after administration of d^7-BB or d^7-SB is directly the result of d^5-BA released from d^7-BB or d^7-SB.

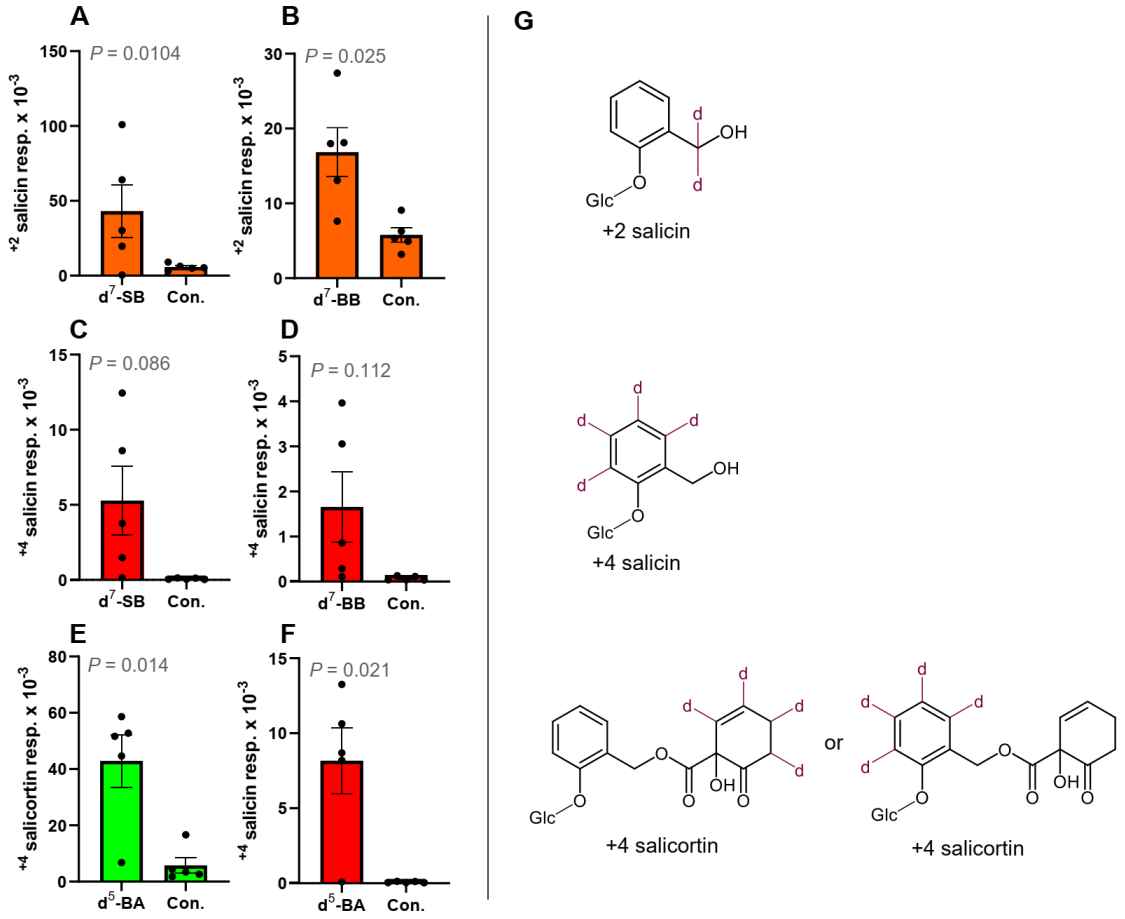


Figure 3.3. Incorporation of deuterium labelled substrates into salicin and salicortin. Results show label incorporation patterns in WT leaf petioles. Relative accumulation of +2 salicin following d^7-SB (A) or d^7-BB (B) administration. Relative accumulation of +4 salicin following d^7-SB (C) or d^7-BB (D) administration. Relative accumulation of +4 salicortin (E) and +4 salicin (F) following d^5-BA administration. Error bars are ± SE, n = 5 experimental replicates which are represented by black dots. P values were determined using a Welch's *t*-test between response areas of label administered and control samples. G, structural diagrams with the inferred deuterium positions for +2 and +4 salicin, and +4 salicortin incorporations corresponding to graphs on the left. Deuterium atoms are drawn in purple, no hydrogens are included. Abbreviations: BB, benzyl benzoate; SB, salicyl benzoate; resp, response; Con, control.

3.3.4 Poplar tissues accumulate salicyl benzoate and benzyl salicylate glucosides, but these may not be intermediates in salicortin biosynthesis

In addition to salicortin and salicin, we analyzed deuterium label incorporation into the potential salicortin biosynthetic intermediates SB-Glc and BS-Glc. For SB-Glc, a +7 label incorporation was observed in leaf petioles (Table 3.1). The absence of +6 salicortin despite the presence of +7 SB-Glc in d^7-SB administered petioles suggests that SB-Glc is not a normal intermediate in salicortin biosynthesis, as the incorporation patterns differ between SB-Glc and salicortin. Formation of an HCH ring from the benzoate component of SB would result in +6 salicortin. For SB-Glc a +2 label incorporation was detected for d^7-SB administered petioles, but there was no detectable incorporation of the benzyl alcohol from the administration of d^7-BB (Fig. 3.4A and Fig 3.4B). Unlike SB-Glc salicortin accumulated a +2 label after d^7-BB. These label incorporation patterns of SB-Glc do not match those of salicortin; therefore, we infer that SB-Glc is not likely an intermediate in salicortin biosynthesis. The accumulation of +5 SB-Glc from the administration of d^7-SB or d^7-BB indicates an incorporation of released d^5-BA into the benzoate moiety of SB-Glc (Fig. 3.4C and Fig. 3.4D). Labelled +5 SB-Glc is approximately 5-fold less abundant in d^7-BB administered petioles compared to d^7-SB administered petioles. The accumulation of +5 SB-Glc preferentially from d^7-SB administration suggests that d^7-SB may be more susceptible to ester bond hydrolysis in plant tissue.

We also measured label incorporation into BS-Glc. Administration of d^7-BB resulted in the accumulation of +6 BS-Glc and +5 BS-Glc in WT and UGT71L1-KO plants (Fig. 3.6A and Fig. 3.6B). The accumulation of +5 BS-Glc demonstrates the incorporation and conversion of the d^7-BB derived d^5-BA moiety into benzyl alcohol. The d^5-BA incorporation into the salicylate moiety would result in a +4 BS-Glc (Fig 3.5B). The labelled benzyl alcohol component of d^7-BB also appeared incorporate directly into BS-Glc resulting in +2 BS-Glc accumulation (Fig 3.5C). We did not detect significant accumulation +4 BS-Glc after administration of d^7-BB. The administration of d^7-SB resulted in no significant accumulation of +2 BS-Glc; however, +5 BS-Glc accumulation was observed (Fig 3.5D and Fig. 3.5E). The BA component of

administered d^7-SB appears to be converted to the benzyl alcohol as +5 BS-Glc was detected.

Both SB-Glc and BS-Glc are natural metabolites that were detected in unlabelled poplar leaf petioles (Fig 3.6A). SB-Glc had a higher relative response area than BS-Glc in WT plants; however, in UGT71L1-KO lines the response areas of SB-Glc and BS-Glc were equivalent. Therefore, it appears likely that UGT71L1 contributes to or influences the accumulation of SB-Glc in poplar. To assess the contribution of BA into SB-Glc and BS-Glc we administered d^5-BA into leaf petioles. After administration of d^5-BA elevated concentrations of +5 BS-Glc were detected (Fig. 3.6B). This indicates that free BA can incorporate into the benzyl alcohol moiety of BS-Glc. Unexpectedly, +4 BS-Glc was not detected in d^5-BA administered petioles. Therefore, the salicylate moiety of BS-Glc does not appear to incorporate free d^5-BA (Fig. 3.6C). As +4 salicortin accumulates after d^5-BA administration, the absence of +4 BS-Glc enrichment suggests that BS-Glc is not an intermediate in salicortin biosynthesis. Administered d^5-BA incorporated into the BA component of SB-Glc resulting in the significant accumulation of +5 SB-Glc (Fig 3.4D).

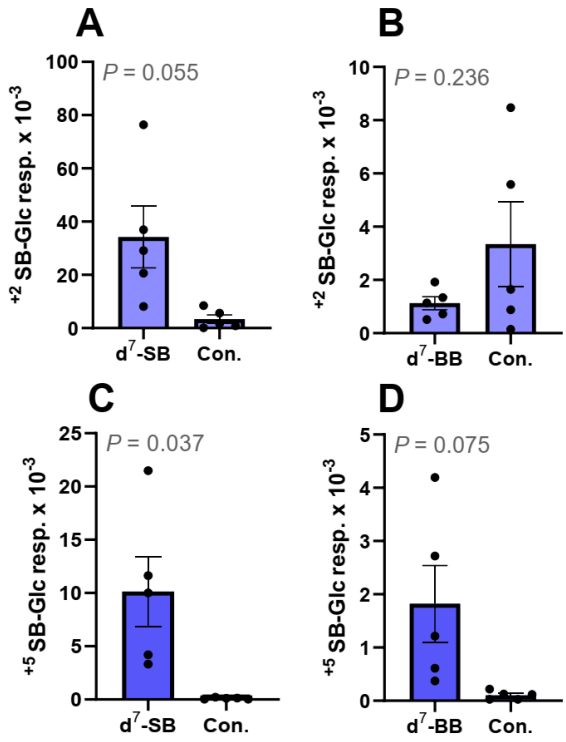


Figure 3.4. Incorporation patterns of SB-Glc from administered labelled salicyl benzoate or benzyl benzoate. The relative accumulation shown as extracted ion chromatogram peak area (response) of labelled SB-Glc in WT leaf petioles. Relative accumulation of +2 SB-Glc following d^7-SB (A) d^7-BB (B) administration. Relative accumulation of +5 SB-Glc following d^7-SB (C) or d^7-BB (D) administration. Black dots represent experimental replicates (n = 5). Error bars are ± SE. P values represent a Welch's *t*-test between response areas of label administered samples and control samples. Abbreviations: Glc, glucoside; BB, benzyl benzoate; SB, salicyl benzoate; resp, response; Con, control.

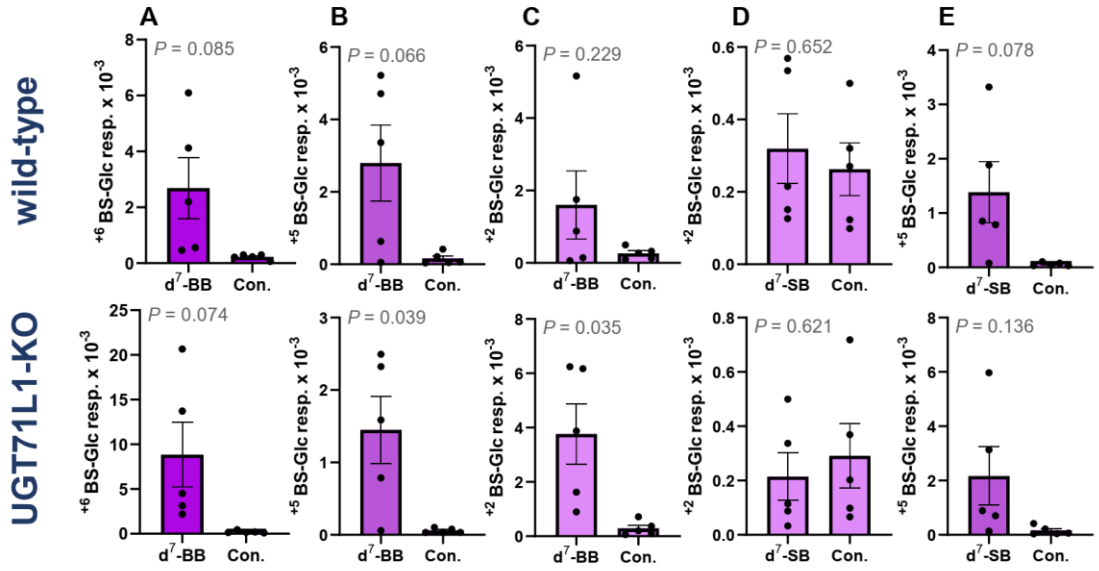


Figure 3.5. Contribution of benzyl benzoate and salicyl benzoate aglycones to benzyl salicylate glucoside accumulation. Deuterium labelled BB or SB was applied to petioles and the relative incorporation was measured. The top row shows analysis of leaf petioles in WT lines, the bottom row is analysis of leaf petioles in UGT71L1-KO lines. From left to right is the relative accumulation of +6 BS-Glc (A), +5 BS-Glc (B), and +2 BS-Glc (C) after d^7-BB administration in WT and UGT71L1-KO lines, and the relative accumulation of +2 BS-Glc (D) and +5 BS-Glc (E) after d^7-SB administration. Black dots represent experimental replicates (n = 5). Error bars are ± SE. P values in (A), (B), (C), (D) and (E) represent a Welch's *t*-test between response peak areas of label administered and control samples. Abbreviations: BA, benzoate; SB-Glc, salicyl benzoate glucoside; BS-Glc, benzyl salicylate glucoside; WT, wild-type; resp, response; Con, control.

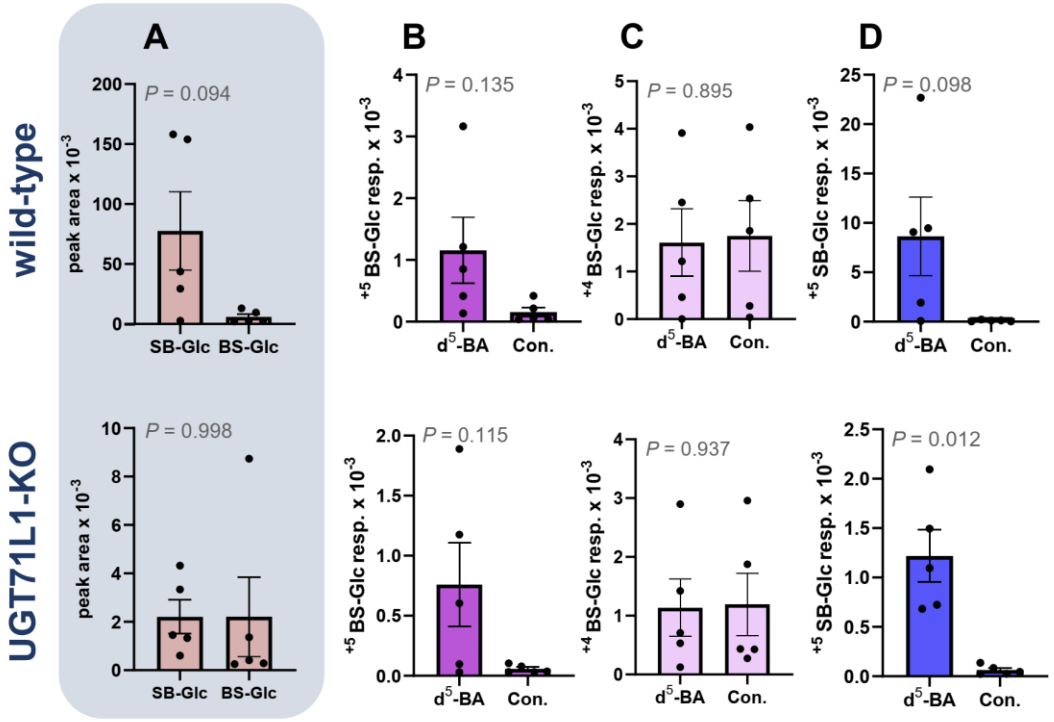


Figure 3.6. Salicyl benzoate glucoside and benzyl salicylate glucoside concentrations in poplar and the contribution of benzoate to their biosynthesis. Graphs represent the relative accumulation of analytes in leaf petioles from WT (top row) and UGT71L1-KO poplar lines (bottom row). Relative accumulation of unlabelled SB-Glc and BS-Glc in WT and UGT71L1 control lines is shown with the shaded background (A). The relative accumulation of +5 BS-Glc (B), +4 BS-Glc (C), and +5 SB-Glc (D) following d^5-BA administration. Black dots in each graph represent an experimental replicate (n = 5). Error bars are ± SE. P values represent a Welch's *t*-test between response areas of SB-Glc and BS-Glc in (A). P values in (B), (C), and (D) represent a Welch's t-test between response areas of label administered and control samples. Abbreviations: BA, benzoate; SB-Glc, salicyl benzoate glucoside; BS-Glc, benzyl salicylate-glucoside; WT, wild-type; resp, response; Con, control.

3.3.5 Deuterated benzoate, salicyl benzoate and benzyl benzoate can contribute to salicylic acid formation

We previously found that UGT71L1-KO lines accumulate 10-fold more salicylic acid (SA) and salicylic acid glucoside (SA-Glc) than WT poplar. Therefore, we administered d^5-BA, d^7-SB and d^7-BB to UGT71L1-KO and WT lines to test if exogenous benzyl benzoate or salicyl benzoate contribute to SA biosynthesis. As expected, the administration of d^5-BA resulted in the significant accumulation of +4 SA-Glc in UGT71L1-KO lines (Fig. 3.7A). Likewise, d^7-BB administration resulted in significant +4 SA-Glc accumulation (Fig. 3.7B). Therefore, the d^5-BA moiety from d^7-BB contributed to SA-Glc accumulation. No +2 SA-Glc was detected after administration of d^7-BB in WT or in UGT71L1-KO lines (Supplemental Figure S3.4B). The administration of d^7-SB resulted in the accumulation of significant amounts of +4 SA-Glc in UGT71L1-KO lines ($P = 0.014$) and a detected but non-significant and highly variable accumulation of +4 SA-Glc in WT lines (Fig. 3.7C). In both WT and UGT71L1-KO lines there was no detected accumulation of +2 SA-Glc following d^7-SB administration (Supplemental Figure 3.3B). Therefore, based on the label incorporation in UGT71L1-KO lines it appears that the d^5-BA component of d^7-SB and d^7-BB can contribute to SA-Glc accumulation, while the benzyl alcohol or salicyl alcohol components do not. Similarities in +4 SA-Glc accumulation between d^5-BA, d^7-SB, and d^7-BB may be indicating that free BA released from BB and SB in UGT71L1-KO lines contributes to SA-Glc formation.

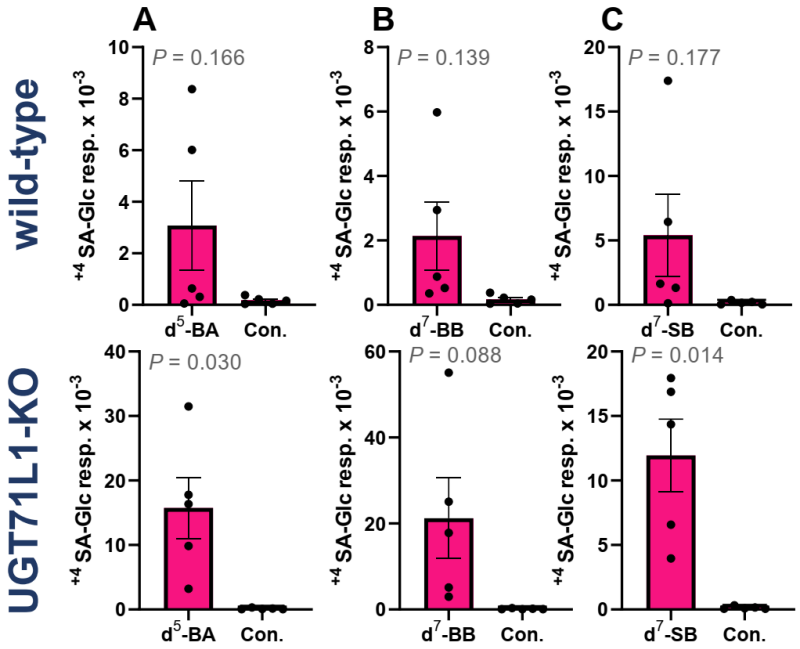


Figure 3.7. Contributions of deuterated benzoate, salicyl benzoate, and benzyl benzoate to salicylic acid glucoside accumulation. The top row shows analysis of WT leaf petioles, and the bottom row is analysis of UGT71L1-KO leaf petioles. The relative accumulation of +4 SA-Glc in WT and UGT71L1-KO lines after administration of d^5-BA (A), d^7-BB (B) or d^7-SB (C). Black dots in each graph represent an experimental replicate (n = 5). Error bars are ± SE. P values represent a Welch's *t*-test between the response areas of label administered and control samples. Abbreviations: WT, wild-type; resp, response; Con, control.

3.4 Discussion

3.4.1 Benzyl benzoate contributes to salicortin biosynthesis

Administration of d^7-BB, d^7-SB, and d^5-BA in WT leaf discs (Fig. 3.2) and leaf petioles (Table 3.1) resulted in label incorporation patterns into salicinoids with distinct patterns. Administration of d^7-BB resulted in the accumulation of +6 salicortin, whereas the administration of d^7-SB did not result in +6 salicortin accumulation. Therefore, our data indicates that benzyl benzoate incorporated directly, and is thus a precursor or intermediate in salicortin biosynthesis. In contrast, exogenously applied salicyl benzoate does not directly contribute to salicortin biosynthesis as a precursor or intermediate. Although intact d^7-BB and d^7-SB enter plant metabolic pathways, a portion of the molecules undergo ester bond cleavage and can provide the released alcohol or benzoate moieties as precursors to salicortin and salicin biosynthesis. All administered labelled benzenoids (d^5-BA, d^7-BB, and d^7-SB) resulted in the accumulation of +4 salicortin. The accumulation of +4 salicortin indicates that the labelled dibenzenoid substrates d^7-SB and d^7-BB can release d^5-BA, which can be utilized for salicortin biosynthesis. Babst et al. (2010) demonstrated the incorporation of ^{13}C-BA into either the salicyl alcohol component or the HCH moiety of salicortin. Our data support the role of BA as a precursor to salicortin biosynthesis. Additionally, +2 salicortin accumulates following d^7-BB administration. Administration of d^7-SB does not generate +2 salicortin. Thus, benzyl alcohol released from BB can contribute to salicortin biosynthesis but salicyl alcohol released from SB cannot. The release of BA from benzyl benzoate may represent a natural process in poplar, as a similar release has been hypothesized in petunia plants (Orlova et al., 2006; Klempien et al., 2012).

The salicinoid biosynthetic enzyme UGT71L1 coexpresses with the SABT and BEBT acyltransferase genes (Fellenberg et al., 2020). SABT and BEBT generate salicyl benzoate and benzyl benzoate, respectively (Chedgy et al., 2015). UGT71L1 is essential for salicinoid biosynthesis in poplar and the enzyme can glucosylate salicyl benzoate (Fellenberg et al., 2020). Therefore, we hypothesized that SB-Glc may be a salicortin biosynthetic intermediate (Gordon et al., 2022). However, our experiments show that SB

and SB-Glc do not readily convert to salicortin (Fig. 3.3 and Fig. 3.5). Incorporation patterns in leaf discs and leaf petioles show that introduced BB can contribute to salicortin biosynthesis. Therefore, we postulate that salicinoid biosynthesis requires BB as a central entry point into the biosynthesis of salicortin. BB is likely generated from benzyl alcohol and benzoyl-CoA precursors facilitated by the acyltransferase BEBT (Chedgy et al., 2015). Furthermore, the absence of +2 salicortin following d^7-SB administration (Fig 3.2), suggests that the generation of salicyl benzoate by SABT using salicyl alcohol and benzoyl-CoA may not be a significant contributor to salicortin biosynthesis.

Benzyl benzoate as a salicortin precursor or intermediate must undergo *ortho*-hydroxylation of the benzyl alcohol and *ortho*-oxidation of the benzoate moiety to achieve a salicortin structure (Fig. 3.1). Therefore, it seems likely that salicyl benzoate or benzyl salicylate are necessary intermediates based on structural requirements of salicortin biosynthesis (Gordon et al., 2022). If hydroxylation occurs on the benzoate moiety first SB would not be made, which could explain the inability of exogenously applied SB to contribute to salicortin biosynthesis. If the alcohol component is hydroxylated first, but SB cannot be introduced, the formed SB intermediate may be short lived and protected from non-specific enzymatic conversions. Label incorporation patterns in BS-Glc and SB-Glc suggest that neither molecule are stable intermediates in salicortin biosynthesis. BS-Glc does not accumulate a +4 label in the presence of d^5-BA but salicortin does. SB-Glc effectively incorporates a +2 and a +7 label from the application of d^7-SB but salicortin does not (Table 3.1).

UGT71L1 contributes to glucosylation of administered SB as the WT leaf petioles have approximately 10-fold more +7 SB-Glc than UGT71L1-KO lines following d^7-SB administration. However, UGT71L1 is also capable of glucosylating BS and salicyl salicylate (Gordon et al., 2022). There may be additional oxidized dibenzenoid intermediate that are natural substrates for UGT71L1. The inability to introduce exogenous SB as an intermediate suggests that the hydroxylated dibenzenoid intermediates in salicortin biosynthesis are not pathway entry points. Therefore, we postulate that the complete biosynthesis of salicortin may be undertaken by a biosynthetic protein complex that begins with benzyl benzoate and is dependent on the presence of

UGT71L1 (Fig. 3.1B). Salicyl benzoate, with the salicyl alcohol moiety may not be capable of entering the metabolon due to hydrophobicity shifts. An example of metabolic channel requiring a functional UGT enzyme is the biosynthesis of the monophenolic glycoside dhurrin in sorghum (*Sorghum bicolor*). Dhurrin biosynthesis requires oxidation and deamination of tyrosine to form the aglycone intermediate cyanohydrin. Cyanohydrin is then glucosylated by SbUGT85B1 in a biosynthetic complex (Laursen et al., 2016). The biosynthetic complex requires the interactions of CYP450 enzymes, complexed with SbUGT85B1 and an electron donor. Growing evidence supports the widespread presence of metabolic channels, or metabolons, in plant specialized metabolism (Winkel, 2004; Dastmalchi et al., 2016; Mucha et al., 2019). The nearly complete elimination of salicinoids in UGT71L1-KO lines, despite the presence of enzymes with identical substrate activities such as UGT78M1 suggests that dibenzenoid glucosylation activity alone is not sufficient to contribute to a salicinoid biosynthetic complex (Fellenberg et al., 2020; Gordon et al., 2022). Therefore, there may a unique signature within the UGT71L1 enzyme that confers the capacity for organization within a salicinoid biosynthetic complex.

Our labelling work demonstrates the capacity for benzyl benzoate to contribute to salicortin biosynthesis. Benzyl benzoate can be directly converted into salicortin with the HCH-moiety forming from the benzoate ring and the salicin moiety forming from benzyl alcohol. We provide evidence that salicyl benzoate may not be capable of entering into the salicortin biosynthetic process. Salicyl benzoate glucoside nor benzyl salicylate glucoside are likely biosynthetic intermediates as their label incorporation patterns do not match that of salicortin. Our experimental incorporation patterns in leaf discs generally resulted in more precise data than petiole uptake; however, the two separate tissues presented similar results regarding salicortin label incorporation. Overall, the label incorporation data presented herein led us to generate an updated salicinoid biosynthetic pathway utilizing benzyl benzoate as an entry point into a postulated biosynthetic channel (Fig. 3.8).

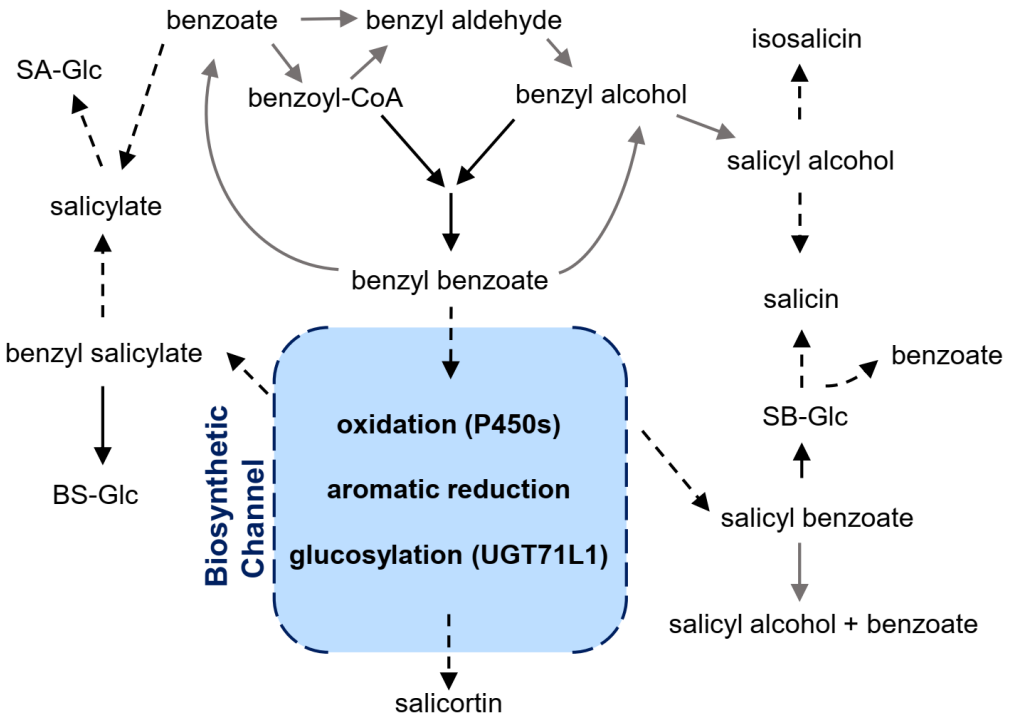


Figure 3.8. An updated model of salicyl benzoate and benzyl benzoate connections to salicortin and salicin biosynthesis. The schematic was generated based on literature and our experimental data. Black arrows represent established enzymatic activity (Fellenberg et al., 2020; Gordon et al., 2022); grey arrows represent established or proposed isotopic connections consistent with experimental results (Zenk, 1967; Babst et al., 2010), and dashed black arrows are proposed metabolic connections. The blue background represents an unestablished number of enzymatic steps producing salicortin from benzyl benzoate. Abbreviations: Glc, Glucoside.

3.4.2 Salicin biosynthesis can utilize salicyl alcohol, benzyl alcohol, or benzoate

Salicin is a glucosylated salicyl alcohol and a centrally shared salicinoid substructure. Salicin is significantly reduced with functional loss of UGT71L1 but not eliminated (Fellenberg et al., 2020). Our work demonstrates that exogenous application of d^7-BB, d^7-SB, or d^5-BA all contribute labelled precursors to salicin biosynthesis in WT and UGT71L1-KO plants. The presence of +2 or +4 salicin from administered d^7-BB and d^7-SB demonstrates that salicin biosynthesis can utilize the released benzyl alcohol moiety from d^7-BB, released salicyl alcohol from d^7-SB, and the released d^5-BA from d^7-BB or d^7-SB (Fig. 3.3). Administration of d^7-BB resulted in preferential biosynthesis of salicin over isosalicin but d^7-SB administration resulted in significant accumulation of +2 isosalicin (Supplemental Figure S3.3 and Supplemental Figure S3.4). Glucosylation of salicyl alcohol or the *ortho*-hydroxylation of benzyl alcohol followed by glucosylation into salicin has been shown previously (Babst et al., 2010). The accumulation of the unnatural isomer isosalicin from salicyl alcohol could signal involvement of a non-specific UGT in the observed pattern of salicin synthesis. The preferential accumulation of salicin over isosalicin following BB administration suggests the benzyl alcohol may be the natural substrate for salicin biosynthesis compared to salicyl alcohol.

The incorporation of d^5-BA into +4 salicin provides evidence for the biological conversion of BA into benzyl alcohol, followed by salicyl alcohol, and then salicin. By necessity BA would require multiple unidentified enzymatic processes to generate a direct salicin precursor (Widhalm and Dudareva, 2015). Fellenberg et al. (2020) demonstrated that UGT71L1 is a biosynthetic enzyme necessary for the major portion of salicin biosynthesis. Functional UGT71L1 also appears to be necessary for the preferential accumulation of salicin as opposed to isosalicin (Gordon et al., 2022). In UGT71L1-KO lines administered either d^7-BB or d^7-SB the +2 isosalicin concentrations are elevated compared to WT plants (Supplemental Figure S3.3B and Supplemental Figure S3.4B). Further exploration into the localization of salicinoid biosynthetic enzymes may provide insight into the metabolic connections of BA and salicyl alcohol (Widhalm and Dudareva, 2015; Lackus et al., 2021).

WT plants accumulate more +2 SB-Glc and +2 salicin following d^7-SB administration than the UGT71L1-KO plants (Fig. 3.3, Fig. 3.4). There may be a connection between +2 or +7 SB-Glc and salicin accumulation. Separation of the central ester bond on labelled +7 SB-Glc would release +2 salicin. Salicin has also previously been shown to arise from the turnover of salicortin via esterase activity (Julkunen-Tiitto, 1992; Julkunen-Tiitto and Sorsa, 2001). Therefore, SB-Glc turnover may be a minor contributor to salicin accumulation in poplar. The identification of a salicyl benzoate or salicortin specific esterase may provide insight into the biosynthesis of salicin. Our data demonstrate that benzenoid metabolism has significant interchange between monophenolic metabolites. The network of benzenoids within poplar likely contributes to the production of salicortin precursor molecules salicinoid biosynthesis (Lackus et al., 2021).

3.4.3 Salicyl benzoate and benzyl benzoate contribute benzoate to salicylic acid biosynthesis

Disrupted salicinoid biosynthesis in UGT71L1-KO lines resulted in the accumulation of 10-fold more SA than WT poplar (Gordon et al., 2022). Through the administration of d^7-SB and d^7-BB we tested if these hypothesized salicinoid biosynthetic intermediates were responsible for the overaccumulation of SA in UGT71L1-KO lines. We demonstrate that the incorporation of d^5-BA from either d^5-BA, d^7-SB, or d^7-BB resulted in the formation of +4 SA-Glc (Fig. 3.7). Therefore, free BA contributes to the formation of SA-Glc. To test if the hypothesized salicinoid intermediate BS-Glc can contribute to SA-Glc formation we compared the label incorporation patterns of BS-Glc and SA-Glc. Surprisingly, administration of d^7-BB or d^5-BA did not result in the accumulation of +4 BS-Glc. Thus, based on the absence of a +4 labelled SA-Glc substructure in BS-Glc it seems unlikely BS-Glc contributes to SA-Glc formation (Fig. 3.1A). As +4 BS-Glc was not accumulated during d^5-BA administration, the BS-Glc biosynthetic process may utilize unknown precursors such as salicoyl-CoA (Chong et al., 2001). Salicyl salicylate glucoside (SS-Glc) is a previously proposed source of SA in UGT71L1-KO lines (Gordon et al., 2022). We thus attempted to determine the label

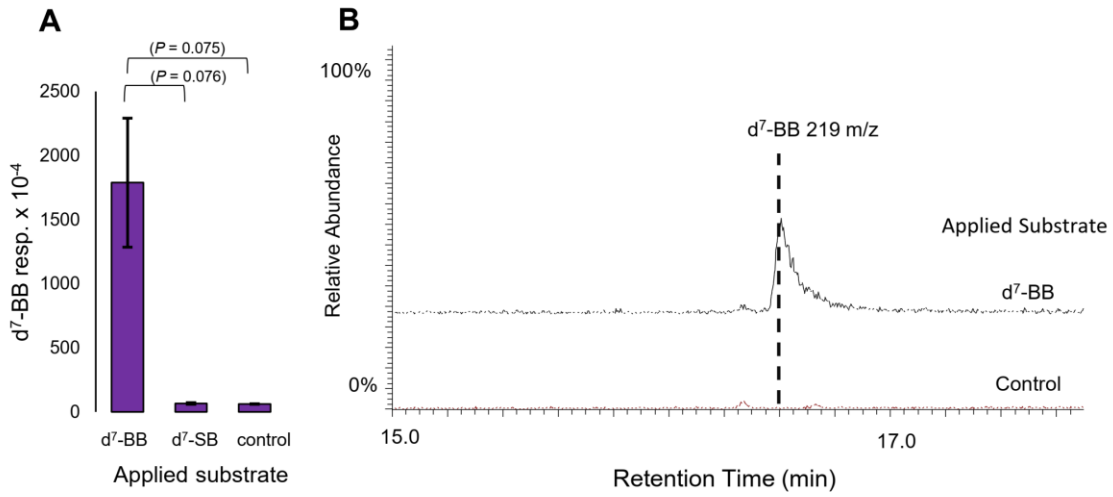
incorporation patterns of SS-Glc but were unable to detect the molecule in leaf disc or petiole extracts. However, the similar accumulation patterns of +4 SA-Glc from benzoate and the dibenzenoid administered substrates do not support the direct conversion of salicortin intermediates into SA. If salicortin biosynthetic intermediates were directly responsible for overaccumulated SA in UGT71L1-KO lines we may expect to see a significantly higher contribution of d^7-BB to +4 SA-Glc as BB is a salicortin precursor. While SS-Glc as a source of SA remains possible, it appears that free d^5-BA is the major source of +4 SA-Glc in UGT71L1-KO and WT poplar lines.

In Arabidopsis the majority of SA accumulates from the shikimate pathway via isochorismate and only an approximately 10% of SA derives from the phenylpropanoid pathway (Ding and Ding, 2020). Research in soybean and tobacco demonstrate that a substantial portion of SA is phenylpropanoid derived (Chong et al., 2001; Shine et al., 2016). Our data supports the idea that the elevated concentrations of SA in UGT71L1-KO poplars are phenylpropanoid derived, as +4 SA-Glc is significantly increased in concentrations in UGT71L1-KO and must derive from the applied d^5-BA. Benzoate is structurally distinct from an isochorismate precursor (Rekhter et al., 2019; Lackus et al., 2021). Although alternative phenylpropanoid derived SA biosynthetic routes are proposed, the significant generation of + 4 SA-Glc may suggest that SA accumulates directly from BA (Chong et al., 2001; Maururi-Lopez et al., 2019). BA conversion to SA would require an *ortho*-hydroxylation. A soluble benzoic acid 2-hydroxylase has previously been described in tobacco (Leon et al., 1995). The absence of an established benzoate 2-hydroxylase gene in poplar limits our understanding of the biosynthesis of phenylpropanoid derived SA. However, RNA-Seq data from UGT71L1-KO plants combined with recombinant enzyme activity of enriched hydroxylase genes may provide insight into the presence and identity of a central SA biosynthetic gene (Gordon et al., 2022). Our work demonstrates that in UGT71L1-KO lines the overaccumulation of SA is likely from BA. Benzyl benzoate or benzyl salicylate are unlikely to be major contributors to SA-Glc formation.

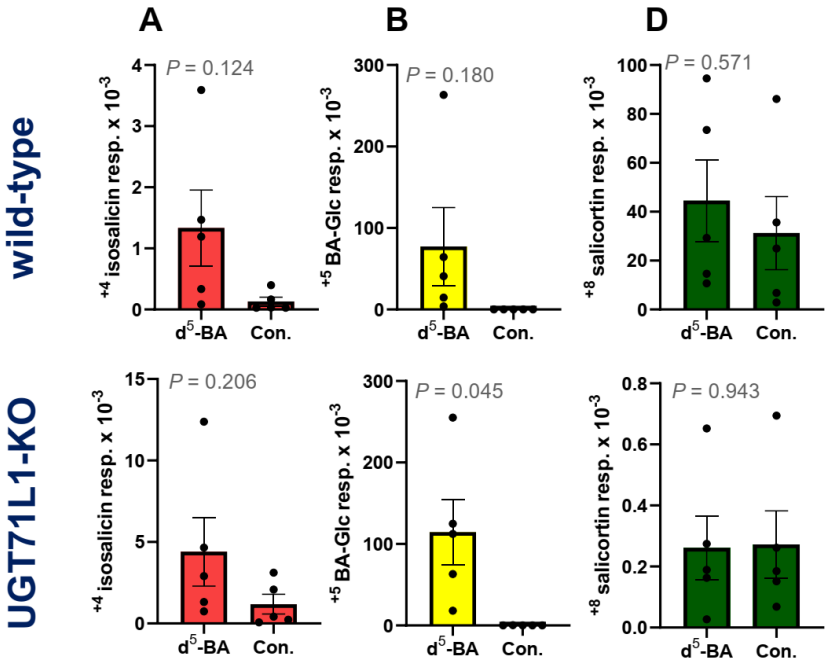
3.4.4 Summary

Labelling data within this report resulted in alterations to our previously proposed salicortin biosynthetic route. The previously hypothesized route of SB-Glc as a salicinoid intermediate now appears unlikely. SB-Glc and BS-Glc appear to be biosynthetically distinct from BB and salicortin biosynthesis. Furthermore, based on the inability to introduce salicyl benzoate into salicortin biosynthesis there may be a salicinoid biosynthetic complex reliant on UGT71L1. Our work highlights the importance of benzyl benzoate in salicortin biosynthesis and benzoate in SA biosynthesis.

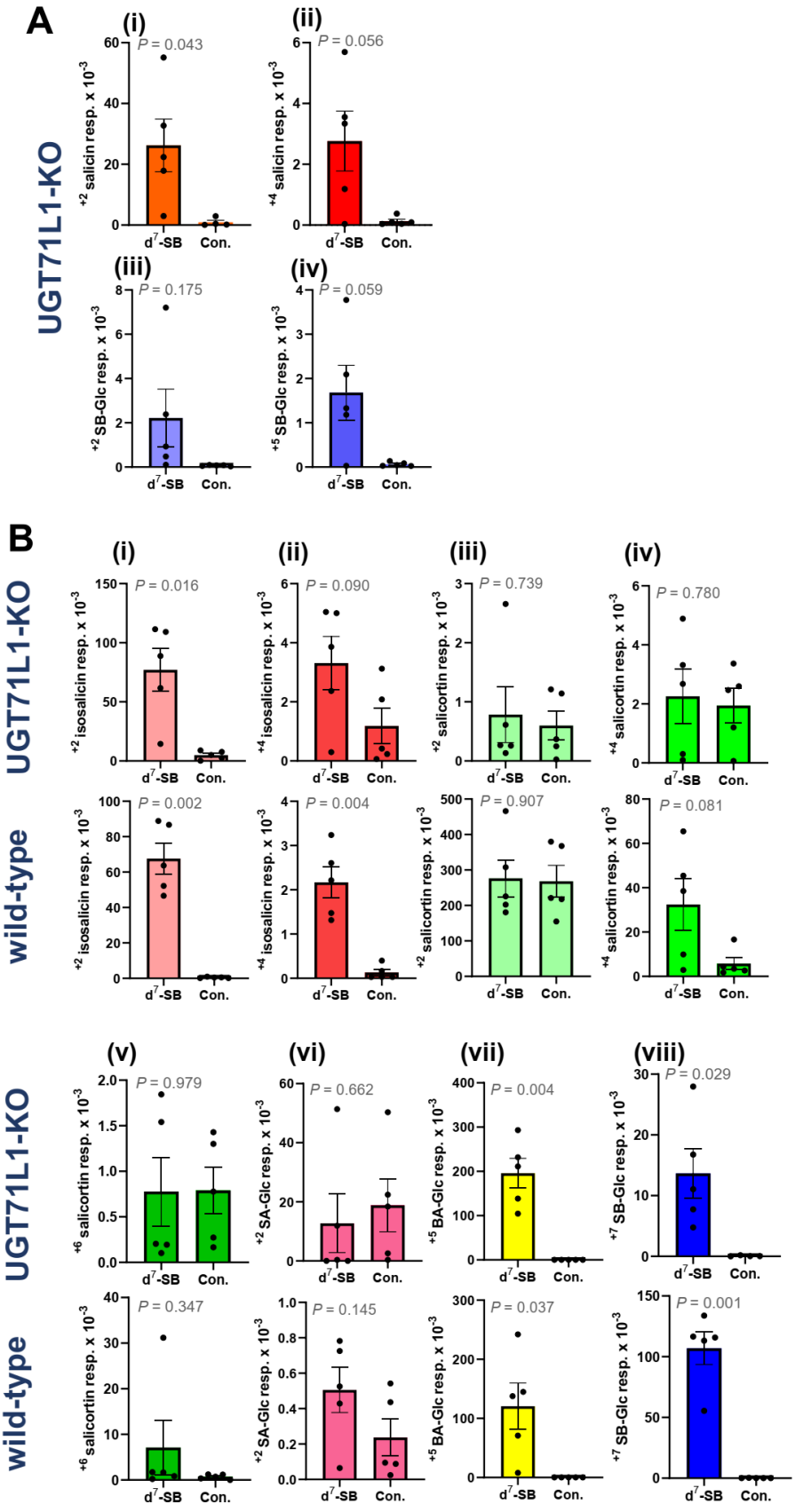
3.5 Supplementary Material



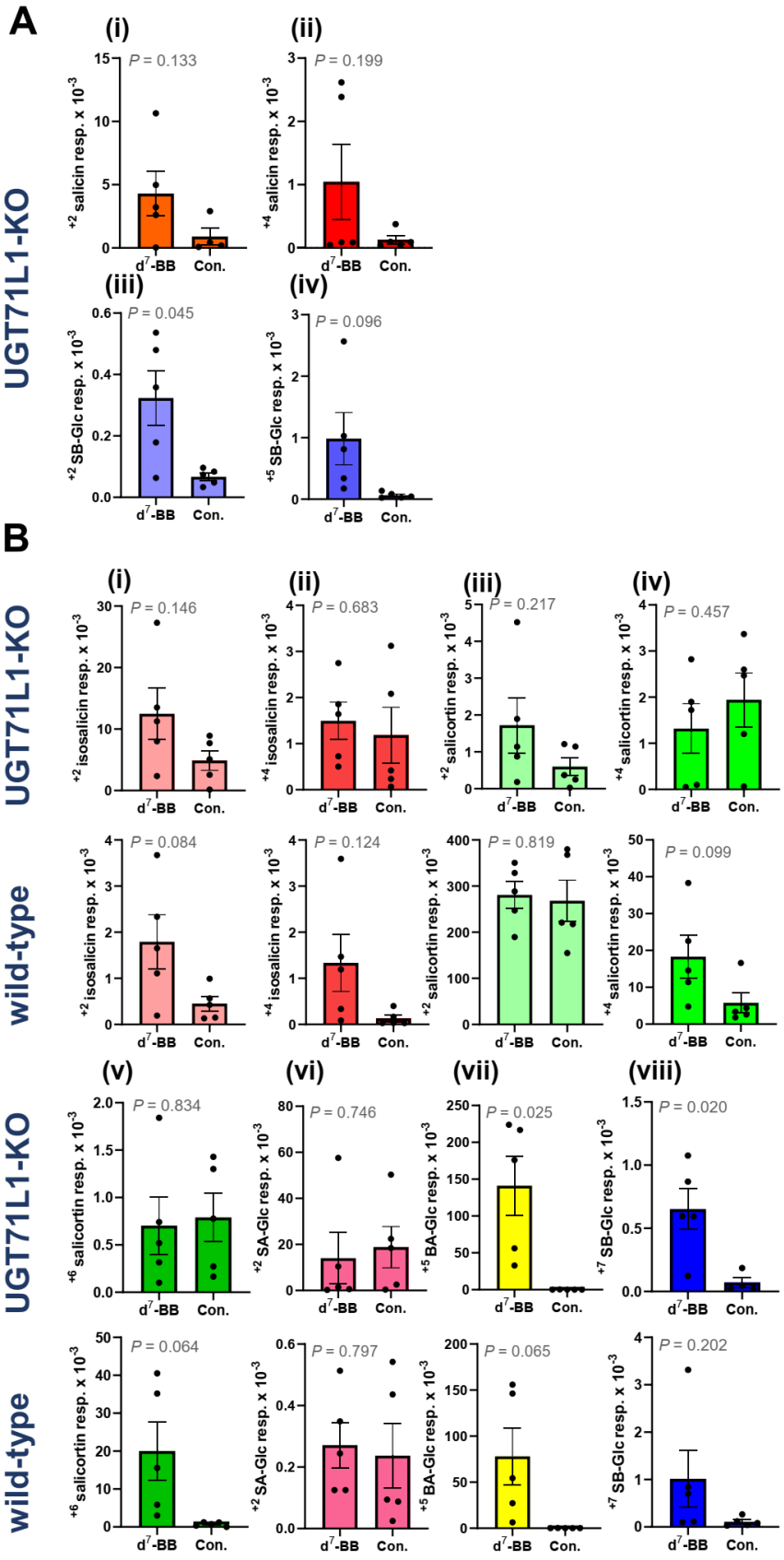
Supplemental Figure S3.1. Detection of d^7-benzyl benzoate in leaf discs. A, peak area for d^7-benzyl benzoate (BB) in leaf discs with administered d^7-BB and d^7-salicyl benzoate (SB). B, extracted ion chromatogram of m/z = 219 of d^7-BB and water (control) leaf discs. Abbreviations: ns = not-significant as per Welch's *t*-test. Error bars represent ± SE.



Supplemental Figure S3.2. Labelled benzoic acid contributions into phenolic analytes. The top row is analyte response areas for the WT line. The bottom row is the analyte response areas for UGT71L1-KO lines. A, detection of +4 isosalicin following d^5-BA administration; B, detection of +5 BA-Glc following d^5-BA administration; C, detection of +8 salicortin following d^5-BA administration. Mass incorporations are presented on individual y-axis labels, numbers represent response area x 10^{-3}. Black dots in each graph represent an experimental replicate (n = 5). Error bars are ± SE. P values represent a Welch's *t*-test between response areas of labelled and control samples. Abbreviations: SB, salicyl benzoate; BA, benzoate; BS, benzyl salicylate; Glc, glucoside; resp, response; Con, control.


A
UGT71L1-KO
(i)
P = 0.043
+2 salicin resp. x 10-3
(ii)
P = 0.056
+4 salicin resp. x 10-3
(iii)
P = 0.175
+2 SB-Glc resp. x 10-3
(iv)
P = 0.059
+5 SB-Glc resp. x 10-3
d7-SB
Con.
B
UGT71L1-KO
wild-type
(i)
P = 0.016
+2 isosalicin resp. x 10-3
(ii)
P = 0.090
+4 isosalicin resp. x 10-3
(iii)
P = 0.739
+2 salicortin resp. x 10-3
(iv)
P = 0.780
+4 salicortin resp. x 10-3
P = 0.002
+2 isosalicin resp. x 10-3
P = 0.004
+4 isosalicin resp. x 10-3
P = 0.907
+2 salicortin resp. x 10-3
P = 0.081
+4 salicortin resp. x 10-3
(v)
P = 0.979
+6 salicortin resp. x 10-3
(vi)
P = 0.662
+2 SA-Glc resp. x 10-3
(vii)
P = 0.004
+5 BA-Glc resp. x 10-3
(viii)
P = 0.029
+7 SB-Glc resp. x 10-3
P = 0.347
+6 salicortin resp. x 10-3
P = 0.145
+2 SA-Glc resp. x 10-3
P = 0.037
+5 BA-Glc resp. x 10-3
P = 0.001
+7 SB-Glc resp. x 10-3

Supplemental Figure S3.3. Labelled salicyl benzoate contributions into phenolic analytes in wild-type and UGT71L1-KO poplar tissue. All samples are the analysis of leaf petioles following 24 h administration of d^7-SB. A, analyte detection specific to UGT71L1-KO lines. A (i), response area of +2 salicin; A (ii), response area of +4 salicin; A (iii), response area of +2 SB-Glc; A (iv), response area of +5 SB-Glc. B, analyte detection in both UGT71L1-KO (upper graphs) and wild-type (lower graphs). B (i), response area of +2 isosalicin; B (ii), response area of +4 isosalicin; B (iii), response area of +2 salicortin; B (iv), response area of +4 salicortin; B (v), response area of +6 salicortin; B (vi), response area of +2 SA-Glc; B (vii), response area of +5 BA-Glc; B (viii), response area of +7 SB-Glc. Mass incorporations are presented on individual y-axis labels, numbers represent response area x 10^{-3}. Black dots in each graph represent an experimental replicate (n = 5). Error bars are ± SE. P values represent a Welch's *t*-test between response areas of labelled and control samples. Abbreviations: SB, salicyl benzoate; BA, benzoate; SA, salicylate; BS, benzyl salicylate; Glc, glucoside; resp, response; Con, control.


A
UGT71L1-KO
(i)
P = 0.133
+2 salicin resp. x 10-3
(ii)
P = 0.199
+4 salicin resp. x 10-3
(iii)
P = 0.045
+2 SB-Glc resp. x 10-3
(iv)
P = 0.096
+5 SB-Glc resp. x 10-3
d7-BB
Con.
B
UGT71L1-KO
(i)
P = 0.146
+2 isosalicin resp. x 10-3
(ii)
P = 0.683
+4 isosalicin resp. x 10-3
(iii)
P = 0.217
+2 salicortin resp. x 10-3
(iv)
P = 0.457
+4 salicortin resp. x 10-3
wild-type
P = 0.084
+2 isosalicin resp. x 10-3
P = 0.124
+4 isosalicin resp. x 10-3
P = 0.819
+2 salicortin resp. x 10-3
P = 0.099
+4 salicortin resp. x 10-3
UGT71L1-KO
(v)
P = 0.834
+6 salicortin resp. x 10-3
(vi)
P = 0.746
+2 SA-Glc resp. x 10-3
(vii)
P = 0.025
+5 BA-Glc resp. x 10-3
(viii)
P = 0.020
+7 SB-Glc resp. x 10-3
wild-type
P = 0.064
+6 salicortin resp. x 10-3
P = 0.797
+2 SA-Glc resp. x 10-3
P = 0.065
+5 BA-Glc resp. x 10-3
P = 0.202
+7 SB-Glc resp. x 10-3
d7-BB
Con.

Supplemental Figure S3.4. Labelled benzyl benzoate contributions into phenolic analytes in wild-type and UGT71L1-KO poplar tissue. All samples are the analysis of leaf petioles following 24 h administration of d^7-BB. A, analyte detections specific to UGT71L1-KO lines. A (i), response area of +2 salicin; A (ii), response area of +4 salicin; A (iii), response area of +2 SB-Glc; A (iv), response area of +5 SB-Glc. B, analyte detection in both UGT71L1-KO (upper graphs) and wild-type (lower graphs). B (i), response area of +2 isosalicin; B (ii), response area of +4 isosalicin; B (iii), response area of +2 salicortin; B (iv), response area of +4 salicortin; B (v), response area of +6 salicortin; B (vi), response area of +2 SA-Glc; B (vii), response area of +5 BA-Glc; B (viii), response area of +7 SB-Glc. Mass incorporations are presented on individual y-axis labels, numbers represent response area x 10^{-3}. Black dots in each graph represent an experimental replicate (n = 5). Error bars are ± SE. P values represent a Welch's *t*-test between response areas of labelled and control samples. Abbreviations: SB, salicyl benzoate; BA, benzoate; SA, salicylate; BS, benzyl salicylate; Glc, glucoside; resp, response; Con, control.

Supplemental Table T3.1. Parameters for the detection of labelled analytes using LC/MS separation and electrospray ionization in negative mode

Target compound (Adduct)[1]	Deuterium incorporation	m/z	Retention time (min)
Salicin (M + FA)	-	331	2.90
	+2	333	
	+4	335	
Salicortin (M+FA)	-	469	5.70
	+2	471	
	+4	473	
	+5	474	
	+6	475	
	+8	477	
Isosalicin (M-H)	-	285	3.28
	+2	287	
	+4	289	
BA-Glucoside (M+FA)	-	329	3.65
	+5	334	
BS-Glucoside (M+FA)	-	435	7.98
	+2	437	
	+4	439	
	+5	440	
	+6	441	
SB-Glucoside (M+FA)	-	435	8.07
	+2	437	
	+5	440	
	+7	442	
SA-Glucoside (M-H)	-	299	3.31
	+2	301	
	+4	303	

[1] Abbreviations: M, molecular ion; FA, formate; BA, Benzoate; BS, benzyl salicylate, SB, salicyl benzoate

Chapter 4 – *Populus* root salicinoid phenolic glycosides are not mobilized to support metabolism and regrowth under carbon limited conditions

This chapter has been submitted in a modified format to Tree Physiology for publication as Hillabrand, R.M.; Gordon, H.; Hynes, B.; Constabel, C.P., and Landhäusser, S.M. HG and RH are co-first authors. RH, HG, CPC, SML conceived and planned experiments. HG and RH co-wrote the manuscript, CPC and SML edited the manuscript. RH and BH conducted re-sprouting experiments and NSC analysis. HG grew initial plant material, conducted salicinoid analysis and statistical comparisons.

4.1 Introduction

For perennial plants such as trees, the ability to store and remobilize carbon reserves is an important adaptation to the seasonal fluctuations of perennial lifecycles, as well as biotic and abiotic stresses that impact carbon gain (Jamieson et al., 2015). Nonstructural carbohydrate (NSC) reserves are critical to perennial plants because they can be allocated and mobilized to support maintenance, growth, and defense when the carbon supply from photosynthesis is limited (Chapin et al., 1990; Hartman and Trumbore, 2016). Starch and simple water-soluble sugars are considered the primary form of NSC stores in plants, but they can also be comprised of more complex compounds such as oligosaccharides, lipids, sugar alcohols, polysaccharides, or phenol glycosides, which may have multiple functional roles (Martin and Smith, 1995; Schädel et al., 2009; Fischer et al., 2015; Schoonmaker et al., 2021). Understanding which NSC components are available for remobilization is a recurring question in plant biology (Dietze et al., 2014; Hartmann and Trumbore, 2016). However, our understanding of compound remobilization is limited. This impacts our ability to accurately assess and predict the role of carbon reserves in tree mortality or resistance to stress (Hartmann et al., 2018). For example, there is evidence for the remobilization of lipids towards respiratory processes when plants are exposed to carbon limiting shade stress (Fischer et

al., 2015). Likewise, the seasonal remobilization of the polysaccharide hemi-cellulose, which is considered a structural carbon compound, has also been reported and proposed to act as a carbon reserve similar to starch (Schädel et al., 2009).

Other types of compounds such as sugar alcohols and phenolic glycosides could contribute to the storage, flux, and remobilization of non-structural carbon (Fischer et al., 2015; Schoonmaker et al., 2021). This idea is reinforced by recent observations that in many plants the distinctions between specialized, primary, and regulatory metabolites have become blurred (Erb and Kliebenstein, 2020). For example, in cassava (*Manihot esculanata*), the amino acid-derived cyanogenic glycosides are potent anti-herbivore defenses but were found to also have a crucial role in nitrogen metabolism. Cassava plants with RNAi-suppression of cyanogenic glycoside biosynthesis affected growth capacity under nitrogen limited conditions, which led to the discovery of primary roles in nitrogen transport from shoot to tubers (Siritunga and Sayre, 2004; Jørgensen et al., 2005). The capacity for releasing and recycling the nitrogen and glucose stored in cyanogenic glycosides has also been demonstrated in sorghum (*Sorghum bicolor*) (Bjarnholt et al., 2018). These studies suggest that relationships between nutrients, carbon, and specialized metabolites may be more widespread across plants than currently reported.

Light deprivation studies, which force plants to utilize stored carbon reserves, are useful in determining the extent of NSC reserve remobilization and redistribution. Most observations from these studies suggest that storage reserve compounds, such as starch are largely remobilized and utilized before plant death, while soluble sugars appear to be maintained at low levels, likely due to their additional functions in osmoregulation, metabolism, and signaling (Piper and Fajardo, 2016; Kebrom, 2017; Weber et al., 2018, Wiley et al., 2019; Goetz et al., 2021). While starch is considered the primary component of stored carbon for most plants, it is unclear to what extent alternative carbohydrate compounds can be remobilized and contribute to survival when photosynthetically derived carbon becomes limited. Although quantification methods that target specific soluble sugar components are available, other common techniques such as the anthrone and the phenol-sulfuric acid assays are rather unselective and will quantify a large range of water-soluble compounds including oligosaccharides and other glucans as well as

phenolic glycosides (Landhäusser et al., 2018). Research by Wiley et al. (2019) indicates that, apart from starch and mono- and disaccharides, other non-conventional carbon compounds captured by this quantitation method may contribute to the carbon storage pool and provide another potential source of carbon for aspen recovery.

In aspen trees a prime candidate for unidentified contributions to carbon storage pool are the specialized metabolites unique to Salicaceae called salicinoids, which are also known as phenolic glycosides. Salicinoids function in insect herbivory defense (Boeckler et al., 2011, Rubert-Nason et al., 2015) and are classified by a core glucosylated salicyl alcohol moiety. Salicinoids are constitutively produced, not considered inducible, and are abundant (Ruuhola and Julkunen-Tiitto, 2003; Stevens and Lindroth, 2005; Fabisch et al., 2019). The abundant and central glucose moiety can theoretically provide a carbon (C) source if salicinoid catabolic recycling can be induced through carbon limitation stress, such as the environmental conditions experienced by aspen roots prior to suckering after an above ground disturbance. The availability of salicinoid-deficient transgenic aspens called UGT71L1-KO (*P. alba x P. tremula*) provides the opportunity to compare the potential remobilization of C from salicinoids during C limitation. In *P. tremula x P. alba* hybrids salicinoids accumulate at 0.13% fresh weight (FW) in leaves, whereas in UGT71L1-KO plants salicinoid concentrations are minimal (Gordon et al., 2022). Additionally, the UGT71L1-KO plants, which display a dramatic growth phenotype, have not been examined for effects below ground, thus utilization of this line can expand our understanding of salicinoid interruption on NSC, starch and sugar concentrations. While the carbon and energy cost associated with salicinoid production implies that an investment in defense occurs in competition with growth, the constitutive nature of salicinoid production, and the direct connection with plant growth and development highlights the central role of salicinoids in plant metabolism (Gordon et al., 2022). Salicinoid content has also been shown to be negatively correlated with growth among aspen genotypes grown under resource limited conditions (Osier and Lindroth, 2006; Barker et al., 2019). We hypothesize that constitutive salicinoid production may contribute to growth and suckering capacity, where the salicinoid glucose moiety is utilized during carbon limited growth (Neilson et al., 2013).

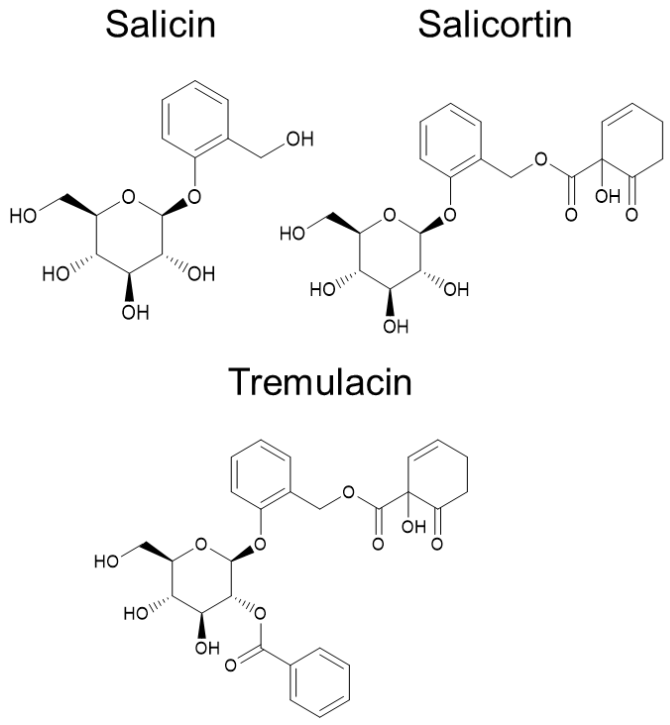


Figure 4.1. Structures of salicin, salicortin, and tremulacin. Abundant salicinoids found in the leaves of the hybrid aspen *P. tremula* x *P. alba.*

In a carbon limitation study, we exposed the root systems of UGT71L1-KO and control lines of a *P. tremula x P. alba* hybrid to a dark resprouting treatment, to allow their root systems to produce suckers and to directly test if salicinoids are catabolized under these conditions. We compared the relative growth (sprouting), concentrations of NSCs, and salicinoid concentrations at different phenological stages in both roots and shoots. Our objectives were i) to determine if salicinoid remobilization for growth and metabolism is supported by observations of salicinoid concentration changes together with other NSCs during carbon limitation, and ii) to test if the shoot regrowth rate (resiliency) is affected by the absence of salicinoid biosynthesis.

4.2 Materials and Methods

4.2.1 Plant cultivation and resprouting conditions

Three individual UGT71L1-KO lines of a *P. tremula x P. alba* hybrid, previously described in detail by Gordon et al. (2022), were pooled and used as a single UGT71L1-KO group for the study (n = 15). Control plants consisted of pooled data from clonal lines of *P. tremula x P. alba* INRA 717 (n = 8) and three independent empty vector control lines (n = 12) for total of 20 control plants. Control and UGT71L1-KO trees were transitioned out of in vitro cultures by transplanting to a peat mix (Sunshine Mix 4, Sungro, Seba Beach, AB, Canada) in pots (5 cm diameter) and grown for three weeks in a mist chamber. Following mist chamber growth plants were transferred to larger pots (4 L) containing the same soil mixture and a slow-release fertilized blend at the concentrations of 8.9 $g \cdot L^{-1}$ controlled release 8-6-12 NPK plus micronutrients (Acer, Delta, BC, Canada), 0.46 $g \cdot L^{-1}$ superphosphate 0-20-0 (Green Valley, Surrey, BC, Canada), 1.21$\cdot$g L^{-1} Micromax Micronutrients (Scotts-Sierra, Marysville, OH, USA), and 4.75 $g \cdot L^{-1}$ Dolomite lime (IMASCO, Surrey, BC, Canada) (Major and Constabel, 2006). Plants were grown in a greenhouse between 20 ºC and 23 ºC with supplemental lighting providing an average of 300 $\mu mol \cdot m^{-2} \cdot s^{-1}$ for a 16-hour photoperiod as described by Gordon et al. (2022). After three months of growth, control plants were trimmed once when their height exceeded one meter to prevent them from falling over. Trees were moved to an open fenced enclosure (mid-August to mid-October) for two months to induce fall dormancy. Since UGT71L1-KO plants remained shorter, they were not trimmed.

In mid-October 2019, the above ground portion of all trees was removed at the base of the stem and the root systems were transferred to a dark and temperature-controlled growth room, wherein the temperature underwent a gradual reduction over six weeks from 4 °C to 0.5 °C. Starting in January, temperatures were raised to 12 °C over three weeks, after which the root systems were moved into larger pots containing vermiculite. This was considered the starting point of the resprouting experiment. A final temperature of 20 °C was reached after an additional 2 weeks (see Fig. 4.2 for

experimental timeline). Throughout the experiment, the potted root systems were left to sprout in complete darkness. Survival duration was recorded as the time from repotting the root systems to their death stage, and adjusted survival duration was calculated by dividing by the final root biomass (see below). Three weeks after the cold storage period, six additional UGT71L1-KO and five additional control plants were transferred to a separate growth chamber with 800 $\mu mol \cdot m^{-2} \cdot s^{-1}$ PAR light, 16 h day length and following the same temperature regime as the dark growth chamber. These plants are referred to as the 'light-control' and 'light-UGT71L1-KO'. The light grown controls were used to establish a salicinoid concentration reference point in above-ground shoot tissue not subject to carbon limitation. Plants in both dark and light chambers were watered as required. When all plants had visible sprouts, plants were fertilized once with a 15:30:15 NPK soluble fertilizer with chelated micronutrients (Plant-Prod Inc., Brampton, ON, Canada). For the dark plants the total number of individual root sprouts was counted and the combined height of the five tallest sprouts was measured for every plant in the dark treatment every two days. The total number of sprouts was last recorded after 50 days of resprouting

4.2.2 Harvesting of plant material

Root samples for NSC and salicinoid analyses were harvested from all plants during cold storage for an initial pre-sprouting measurement. A sub-sample of one gram of fine root tissue (roots 1-2 mm in diameter) was harvested from each root system and freeze-dried for 72 h. Two other phenological stages were chosen for harvesting: (1) at the time when shoot growth ceased (end-of-growth stage) and (2) at the onset of shoot death (death stage) when all shoots showed signs of tissue necrosis. However, not all tissues of the plant might have been functionally dead at the death stage, particularly the larger woody structures of the root system. We believe that this stage can be considered the functional death of the whole root system, as no new sprouts were produced and the existing sprouts were starting to die back (Wiley et al., 2019). The timing of these sample collection stages was on an individual plant basis, as each plant reached each phenological stage at slightly different times. The 'end-of-growth' samples were

collected when no new sprouts were observed and the summed height growth of the five tallest sprouts did not differ more than one cm over five consecutive days. Samples corresponding to the plant death stage were harvested when every sprout had a minimum of one cm of necrotic tissue at the tip. While necrotic black tissue was readily identifiable in the control plants, dead sprout tissue of UGT71L1-KO plants tended to maintain the pale yellow/white colour while desiccating and withering (Supplemental Figure S4.1). We collected stem tissue from all plants grown in dark conditions because the etiolated UGT71L1-KO sprouts produced very little leaf tissue (Fig. 4.2). After the root and shoot samples at the death stage were collected, the whole plant was harvested. All remaining above and below ground tissue was collected and dried for one hour at 100 °C followed by 72 h at 70 °C. Dried tissue was weighed and total root and shoot mass recorded. Relative sprout mass was recorded as shoot weight divided by root weight (Fig. 4.5). Plants grown in the light were harvested on a single day for final salicinoid and biomass measurements following five months of growth. At the five-month time point, the majority of resprouting carbon-limited plants had reached the death stage and the experiment was terminated three weeks later when the last two etiolated controls were harvested.

4.2.3 Analysis of non-structural carbohydrates

A subset (UGT71L1-KO n = 15, control n = 12) of freeze-dried root tissues was ground to a fine powder using a ball mill (TissueLyser II, Qiagen Inc, Mississauga, ON, Canada) and analyzed for NSC, starch, and soluble sugar concentrations using the phenol–sulfuric acid assay. Briefly, total soluble sugars were extracted in 80% ethanol, hydrolyzed by sulfuric acid, and measured colourmetrically at 490 nm after a phenol addition (Landhäusser et al., 2018). The starch remaining in the pellet was broken down enzymatically into glucose. Following the addition of a peroxidase-glucose oxidase-o-dianisidine solution, the concentration was measured colourmetrically at 525 nm, and a conversion factor of 0.9 was used to convert the glucose concentration to starch concentration (Landhäusser et al., 2018). Concentrations were expressed as a percentage of sample dry weight. Total NSC concentrations were calculated as the sum of the total

soluble sugar and starch concentrations. Total NSC and the subsets total sugar, and total starch concentrations were averaged for the control and UGT71L1-KO trees. Additionally, the concentrations of glucose, fructose, and sucrose (GFS) were determined separately using a previously described NAD-linked enzymatic assay (Landhäusser et al., 2018). Remobilized NSC mass measured was defined as the NSC mass at the initial stage minus the NSC mass remaining at the death stage.

4.2.4 Salicinoid extraction and analysis

A subset of tissue samples from roots, stems, and leaves was collected, freeze-dried, and milled as per the NSC analysis. Extraction of dried samples (10 mg) was conducted in 96-well plates with 950 µL of methanol. The plates were sealed with silicone lids and vigorously shaken for 1 min at 15 cycles·s^{-1}, followed by gentle agitation on an orbital shaker for 30 min. Plates were centrifuged for two min at 2,465 *g*, and 800 µL of supernatant was transferred to a fresh 96-well plate for LC/MS analysis. Salicinoids were analyzed using separation and quantification methodology based on Fellenberg et al. (2020) on a Waters Acquity UPLC-MS. In brief, a 2 µL injection was separated using a Waters Acquity Ethylene Bridged Hybrid column (C18, pore size: 1.7 µm, dimensions: 2.1 x 50 mm). A binary solvent gradient separation was used with ultra-pure water containing 0.1% formate (v/v) (A) and LC-MS grade acetonitrile with 0.1% formate (B). The flow rate was 0.5 mL·min^{-1}. Separation gradient was 99.9% A, 0-0.5 min; 99.9-80.0 % A, 0.5-5.0 min; 80.0-50.0% A, 5.0-8.0 min; 50.0-10.0% A, 8.0-9.0 min; 10% A, 9.0-10.0 min; 10-99.9% A, 10.0 – 10.1 min; 99.9% A, 10.1-11.0 min. A single quadrupole MS in electrospray ionization negative mode was used to detect and quantify salicortin, tremulacin, tremuloidin and salicin. The salicin formate adduct at m/z = 331.0 with cone voltage (V) of 10 was detected at 2.8 min. Salicortin (-H) at m/z = 423.0 with V = 20 was detected at 5.0 min. Tremulacin (-H) at m/z = 537.0 with V = 18 was detected between 7.3 min. Salicortin and tremulacin were quantified using linear regression of authentic standard peak areas provided by Dr. Richard Lindroth. Salicin was quantified using commercially available standards from Sigma-Aldrich (Oakville,

ON, Canada). Peak integration was conducted using MassLynx 4.2 (Waters, Mississauga, ON, Canada).

4.2.5 Calculations and data analysis

Welch's *t*-test was used to test for significant differences between control and UGT71L1-KO trees grown in the dark for growth characteristics, NSC components, and salicinoid concentrations. Regression analysis was performed with R Statistical Software (R Core Team, 2018). To normalize the amount of shoot tissue produced by the different root masses we used relative sprout mass, obtained by dividing the total dried sprouted shoot tissue mass (g) by the total dried root tissue mass (g). Remobilized NSC was calculated as the difference between the NSC mass from the initial pre-sprouting and the final sampling at death. A linear regression was used to test the significance of the relationship between relative sprout mass and remobilized NSC mass, as well as starch, total sugars, and sucrose. The formula for the calculation of percentage starch utilized for salicinoid production is presented below.

$$\%\boldsymbol{Ss} = \left(\left(\frac{\boldsymbol{St}}{\boldsymbol{St}-\boldsymbol{SS}}\right) - \mathbf{1}\right) * \mathbf{100}$$

Equation 4.1. Starch utilization for salicinoid biosynthesis. %Ss is the percentage starch mass utilized for shoot salicinoid biosynthesis; St is Δ starch mass in roots between the initial and death stage in g; SS is shoot salicinoid total mass at the death stage in g.

4.3 Results

4.3.1 Low salicinoid UGT71L1-KO poplars respond differently to carbon limitation during resprouting compared to wild-type poplar

Sprouts of UGT71L1-KO and control lines grown in the dark were etiolated, with white stems and small pale yellow-white leaves (Fig. 4.2a). The UGT71L1-KO produced a significantly larger number of sprouts per root system (P =0.02; Welch's *t* test), but the leaves were extremely reduced in size. There was no significant difference between control and UGT71L1-KO lines in the total height of the five tallest sprouts combined; however, sprouts in the UGT71L1-KO line were thinner and more flexible compared to the control plants resulting in greater sprout mass in the control lines (Table 4.1 and Fig. 4.2). Significant differences in the time to reach the death stage (survival duration) were recorded between UGT71L1-KO and the control lines. Control plants survived an average of 56 days longer than the UGT71L1-KO plants, which reached the death phenological stage after an average of 71 days of carbon limited conditions. Control plants had five-times more root and shoot mass than the UGT71L1-KO plants (Table 4.1); however, when survival duration was normalized with final root biomass, UGT71L1-KO plants survived three times longer per gram of root biomass supporting sprouts than the control plants. In both control and UGT71L1-KO plants no additional new fine root growth was observed for the duration of the study. When sprouting under light, all plants showed significant new root growth and control plants developed larger and round leaves, while UGT71L1-KO plants produced leaves that were narrow, but overall had a higher leaf number. The phenotypes of light-UGT71L1-KO lines that resprouted were consistent with observations described by Gordon et al. (2022).

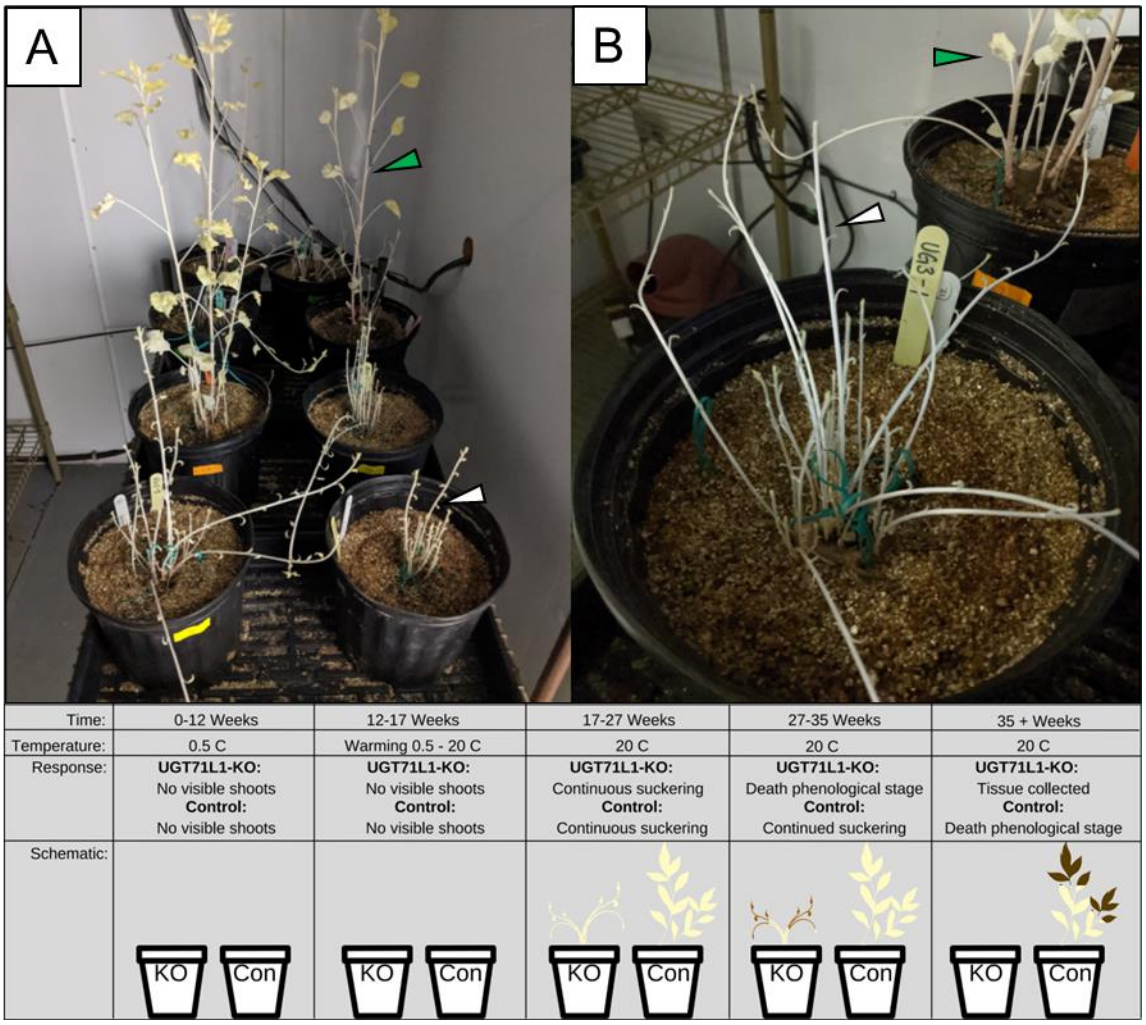


Figure 4.2. Resprouting plant set-up and experimental timeline. (A) shows experimental control (rear) and UGT71L1-KO (forefront) hybrid aspens (*P. tremula x P. alba*) resprouting in a dark growth chamber. An enlarged image of UGT71L1-KO is shown in (B). White arrows indicate the slender, etiolated stems with highly reduced leaves of the UGT71L1-KO line. Green arrows point to etiolated stem and reduced leaves of control plants. Bottom segment cartoon highlights experimental stages and approximate timeline. Abbreviations: KO; UGT71L1-KO, Con; Control.

Table 4.1. Phenotypic traits of poplar during carbon-limited resprouting

	UGT71L1-KO	Control
Root collar diameter (mm)[a]	10.52 (±0.27)*	20.03 (±0.42)
Root biomass (g)[b]	7.51 (±0.46)*	41.61 (±1.87)
Shoot biomass (g)[b]	0.74 (± 0.06)*	3.50 (± 0.38)
Height (cm)[c]	201.04 (± 6.07)	224.15 (± 25.73)
No. sprouts[d]	25.53 (± 2.40)*	18.05 (± 2.15)
Survival duration (days)[e]	71 (±1.4)*	127 (±6.1)
Survival - adjusted (days g^{-1})	9.5*	3.1

[a] Root collar diameter recorded prior to resprouting.
[b] Root and shoot mass recorded after the final harvest which occurred when plants reached the death phonological stage.
[c] The height represents the combined height of the five tallest sprouts per root system.
[d] The total number of individual sprouts from a single root stock following 50 days of resprouting.
[e] The number of days from the start of the experiment until death (survival days)
* denotes $P \leq 0.05$ using a Welch's *t*-Test. Numbers represent the mean ± SE. UGT71L1-KO n = 15; Control n = 20

4.3.2 Carbon limitation progressively depletes but does not eliminate non-structural carbohydrate root stores

Prior to resprouting, both lines had similar NSC (30-37%), starch (8-10%), and total soluble sugars (23-27%) concentrations in their roots as measured by phenol acid. In both lines, the concentration of C reserves decreased during resprouting. However, the decline in the UGT71L1-KO line was greater and resulted in significantly lower concentrations of NSC, starch, and soluble sugars compared to the control plants at both the end-of-growth and death stages (Fig. 4.3a-c). Initially, the glucose, fructose, and sucrose (GFS) components represented 62% of the total sugar concentrations for both control and UGT71L1-KO lines. Their proportion declined to 4.2% in controls and 7.5% in UGT71L1-KO at the death stage (Fig. 4.3c). Glucose concentrations in UGT71L1-KO were significantly lower at the initial stage than the control (Fig. 4.3d). Initial sucrose concentrations in roots were over six-fold higher than fructose and glucose and made up the bulk of GFS concentration (Fig. 4.3d-f). Overall sucrose comprised 49% of all soluble sugars in the roots of the UGT71L1-KO and control lines, and concentrations were not significantly different between the lines (Fig. 4.3f). Notably, sucrose was the only NSC

component that was depleted to below the detection limit in the roots of both plant lines at the end-of-growth and death stage (Fig 4.3f). At the death stage, starch concentrations were reduced to 0.1% dry wt for UGT71L1-KO lines and 2.2% dry wt in control lines. At this stage, total sugars were reduced to 3.2% and 6.2% for UGT71L1-KO and control lines, respectively.

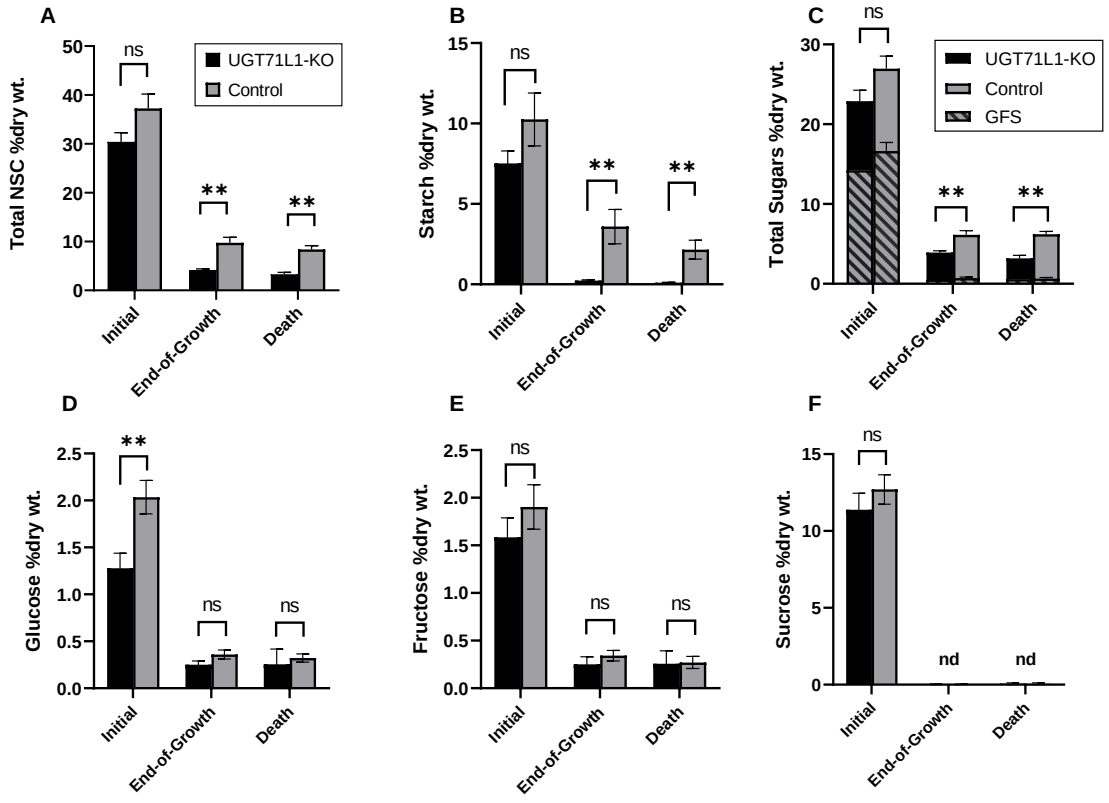


Figure 4.3. Total NSC concentrations and individual NSC component in roots. Collected from resprouted aspen at three phenological stages as defined in Materials and Methods. (A), total NSC; (B), total starch; (C), total sugars including the combined glucose, fructose, and sucrose (GFS) component (hatched bar, determined enzymatically); (D), total glucose; (E), total fructose, and (F), total sucrose per each line ± SE (sample sizes: UGT71L1-KO, n = 15; control, n = 12). Data represents mean ± SE. Asterisks denotes significant differences as determined by a Welch's *t*-test (* = $P < 0.05$, ** = $P < 0.01$, ns = non-significant, nd = not-detected)

4.3.3 Carbon limitation does not significantly reduce salicinoid content in dark-sprouted aspen shoots and roots

Salicinoid analysis of etiolated sprouts and root material demonstrated UGT71L1-KO plants had low salicinoid concentrations (< 1 µg·mg DW^{-1}) (Fig. 4.4.). In control plants, shoot salicinoid concentration was comparable to light grown control plants, and there were no declines in salicinoid concentration between end-of-growth and the death phenological stage during carbon limited growth (Fig. 4.4a-c). Interestingly, tremulacin was the most abundant salicinoid in shoots (Fig. 4.4c), whereas salicortin was more abundant in roots (Fig. 4.4e). Root salicinoid concentrations were not reduced during the experiment, indicating that salicinoid concentrations were maintained in roots despite carbon limited growth. Salicin concentrations in roots increased significantly between initial and the end-of-growth phenological stages (Fig. 4.4d). Salicortin and tremulacin maintained stable concentrations throughout the experiment (Fig. 4.4e-f). In the light-grown UGT71L1-KO plants, leaves had higher salicinoid content than the other above-ground tissues (Fig. 4.4b). Salicin concentrations were also dramatically increased in leaves on light-grown sprouts of UGT71L1-KO compared to light-control plants (Fig. 4.3a). This was the only salicinoid where this was observed; resprouting from initial roots evidentially stimulated salicin accumulation, an effect that was amplified in UGT71L1-KO lines.

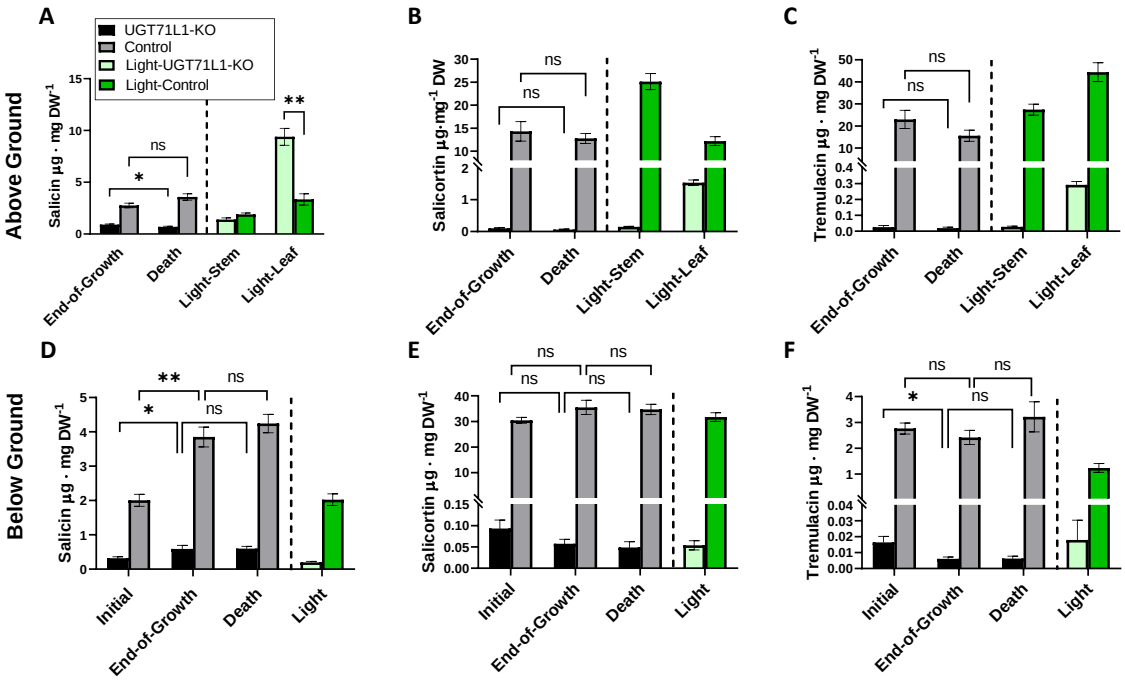


Figure 4.4. Salicinoid concentrations in *P. tremula x P. alba* shoots (top row) and roots (bottom row) from root systems allowed to resprout in the dark at three phenological stages. The initial stage are samples from roots prior to sprouting, whereas the end-of-growth and death include shoots and represent phenological stages described in detail in the methods. Additionally, samples were collected from light resprouted tissue (green color) which was collected at the termination (five months of growth) of the study. (A), above ground (shoot) salicin; (B), shoot salicortin; (C), shoot tremulacin; (D), root salicin; (E), root salicortin; (F), root tremulacin. Dashed vertical line represents separation of dark resprouted and light resprouted plants. Data represents mean ± SE. Asterisks denotes significant differences as determined by a Welch's *t*-test (* = $P < 0.05$, ** = $P < 0.01$, ns = non-significant). All salicinoid analysis samples had $n \geq 5$ biological replicates (complete list in Supplemental Table S4.1).

4.3.4 Relative sprout mass is not equivalent between control and salicinoid deficient plants

In control plants, relative sprout mass (shoot biomass divided by root biomass) was significantly correlated with remobilized NSC mass ($R^2 = 0.41$, $P = 0.02$). Relative sprout mass increased as the amount of potentially remobilized NSC increased (Fig. 4.5). By contrast, in the UGT71L1-KO line relative sprout mass did not correlate with the remobilized NSC mass; however, UGT71L1-KO root systems produced more shoot biomass per gram of remobilized NSC (Fig. 4.5). Similar relationships were found when the relative sprout mass of the control line was compared with remobilized starch mass ($R^2 = 0.43$, $P = 0.01$) and sucrose mass ($R^2 = 0.37$, $P = 0.02$) however, this relationship was not significant for total soluble sugar mass (Supplemental Figure S4.2). During resprouting, the highest total salicinoid content was found in control plants that produced the greatest relative sprout mass (Supplemental Figure S4.3). To estimate the potential carbon costs of salicinoid production in the two lines, we calculated the proportion of total starch that was estimated as a requirement for salicinoid production (Eqn. 4.1). In UGT71L1-KO lines, this proportion was 0.1%, whereas control lines utilized at least 3.1% of consumed starch for salicinoid production in etiolated shoots.

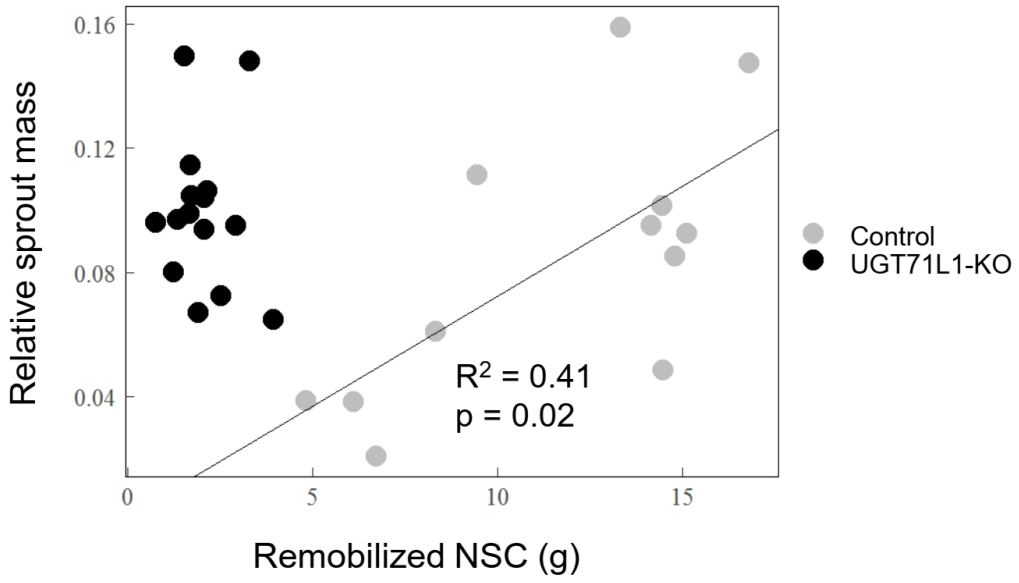


Figure 4.5. Relative sprout mass at death as a function of remobilized non-structural carbohydrates (NSC) mass. UGT71L1-KO over-performed in relative sprout mass production per gram of remobilized NSC, relative to control plants. Relative spout mass was calculated by dividing each individual plants total shoot mass by total root mass. Remobilized NSC mass was calculated as the difference between the NSC mass measured at the initial stage and the death stage NSC mass. No significant correlations were observed between relative sprout mass and remobilized NSC in UGT71L1-KO plants.

4.4 Discussion

A major objective of this work was to determine if the salicinoid phenolic glycosides can be catabolized and subsequently used as an energy source during carbon limited stress in aspens. In control plants, salicinoid concentrations (disproportionately made up of salicortin) in roots remained stable from the initial timepoint until the plant death stage, while root NSC levels declined strongly during that period. In the etiolated shoot tissues (control and UGT71L1-KO lines), the salicortin and tremulacin concentrations remained stable. This indicates that aspen plants maintained a similar production of constitutive salicinoid defense molecules in sprouts despite carbon limitation and the energetic costs associated with biosynthesis. This finding is consistent with previous research demonstrating the metabolic stability and constitutive production of salicinoids (Ruuhola and Julkunen-Tiitto, 2002, Fabisch et al., 2019). Salicinoid concentrations in the carbon limited plants were comparable to our light grown trees and were inferentially not catabolized as an energy source. Although the overall concentrations of salicinoids in roots generally remained constant (Fig. 4.4), the salicin concentration in the roots of UGT71L1-KO and control lines increased during the initial phases of the experiment, indicating a greater production of salicin during the early onset of the resprouting process. Surprisingly, the salicin concentration in light-UGT71L1-KO leaves was also significantly elevated after resprouting, an effect not observed for salicortin or tremulacin. Although unexpected, this is consistent with previous work showing that salicin is synthesized and regulated, at least in part, via a distinct pathway from salicortin and tremulacin (Fellenberg et al., 2020). Our work also documented differences in salicinoid profiles between shoot and root tissues in *P. tremula x P. alba*. In shoots, tremulacin was the most abundant salicinoid, with salicortin present at lower concentrations (Fig. 4.4), whereas in roots the salicortin was the most abundant salicinoid. Tremulacin is a modified salicortin with an additional benzoate moiety (Fig. 4.1). The distinct salicinoid profiles could reflect different adaptive salicinoid functions in roots and shoots.

Although the starch component of the NSC pool significantly declined in both lines, it was reduced to the limit of detection in UGT71L1-KO plants. However, starch

was still 2.2% DW in the control plants at the death stage. Variation in the ability of plants to convert all stored starch during the final stage of severe carbon limitation stress has been observed in other studies and has been related to disruptions of the metabolic breakdown of starch or to differences in pool size and the accessibility of some starch reserves in older, more distant, or woody tissues (Landhäusser and Lieffers, 2012; Wiley et al., 2017). Total soluble sugars were significantly reduced at the end-of-growth and death stage but remained higher in concentration than starch for UGT71L1-KO and WT lines. This has also been observed in other studies and is likely related to the need of soluble sugars in other metabolic roles in the function of plant cell or to disruption of phloem transport (Sevanto et al., 2014; Martínez-Vilalta et al., 2016). While we maintained detection of soluble sugars using the phenol method, the nearly complete depletion of the GFS component of the soluble sugar pool at the death stage highlights the importance of GFS as a measure of available carbon during severe carbon limitation. Our data show that in control plants, the death stage was reached before starch was completely depleted in the root system. It may have been the absence of sucrose that resulted in the cessation of growth and the onset of widespread shoot necrosis (our indicator of the functional death of the root systems), as low concentrations of glucose, fructose, and starch were still detectable at the death stage, whereas sucrose was already below detectable limits at the end-of-growth stage. We speculate that since sucrose plays a central role in starch metabolism and associated carbon transport (Hartmann and Trumbore, 2016), the absence of sucrose in the root system may indicate that basal metabolic demands of root tissue prevented sucrose accumulation. The monosaccharides (glucose, fructose) may be utilized for immediate metabolic requirements instead of being converted to sucrose for transport (Fig. 4.6) (Sevanto et al., 2014; Martinez-Vilalta et al., 2016; Kebrom, 2017).

Although apical dominance and bud activation has been historically attributed to auxins, there is growing evidence for an interaction of sucrose and photosynthates in regulating plant development and growth. For example, in peas, sufficient sucrose is required for bud release and activation (Mason et al., 2014), while in peach roots growth and branching increases when provided with supplemental sucrose (Zhang et al., 2020). Sucrose synthesis requires an initial investment in energy through first the hydrolysis of

starch granules and the production of UDP-glucose (Ruan, 2014; Macneill et al., 2017). Sucrose transport through the phloem relies on symporters which also require a net input of energy (Stein and Granot, 2019). The end-of-growth phase may represent the timepoint where starch and sugar availability are only sufficient to match the metabolic demands of the root tissue hence the presence of glucose and fructose but the absence of sucrose.

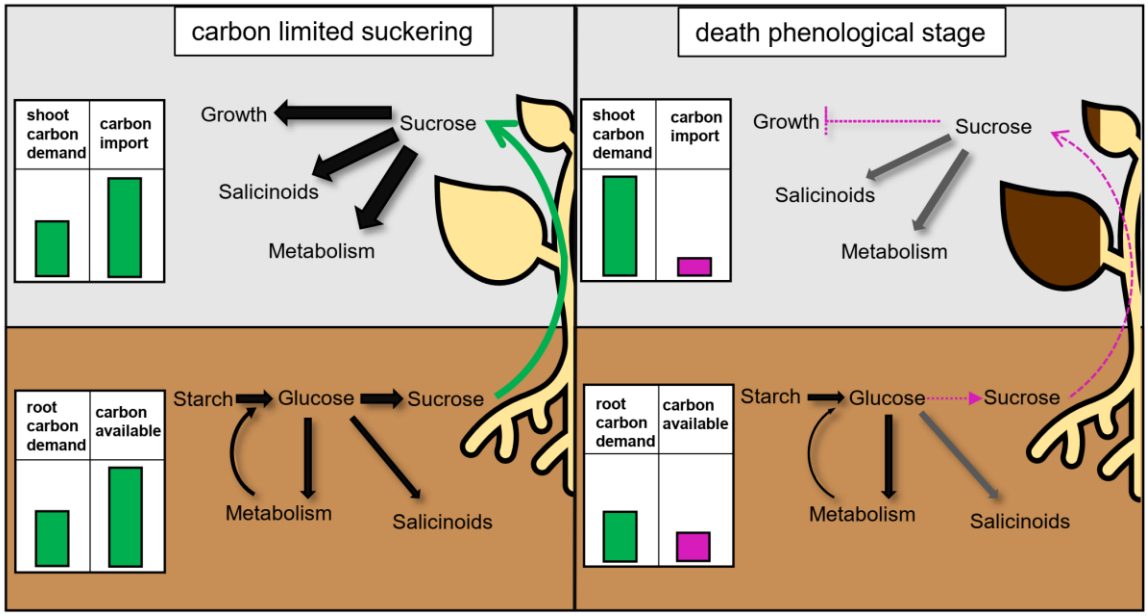


Figure 4.6. Conceptual model of carbon flow from roots to shoots during aspen resprouting (left panel) and the carbon dynamic during the death phenological stage (right panel). Black arrows represent active metabolic connections, with thicker arrows representing increased flux, green arrows over the plant cartoon represent active phloem transport of sucrose, whereas greyed arrows represent reduced metabolic connections. Magenta arrows represent halted physiological and metabolic process during the death stage.

To examine the impact of constitutive salicinoid production on growth and carbon allocation, we examined the differences in NSC utilization between the control and salicinoid deficient plants. Although the salicinoid-deficient UGT71L1-KO plants were smaller, when normalized for root biomass they had greater relative growth rates and survival duration compared to control line plants (Table 4.1). This could reflect the metabolic cost of salicinoid biosynthesis and suggests a growth-defense trade-off: our estimations show that control plants required approximately 3% of the total starch mass for salicinoid production (Eqn. 4.1.). In other studies, a negative correlation of growth and defense compound accumulation has also been observed for aspen genotypes with differing salicinoid content (Osier and Lindroth, 2006; Kruger et al., 2020; Cope et al., 2021) but the cost of the trade-off was not quantified. In control trees, the greater the relative sprout growth, the greater total salicinoid content at death, a relationship absent in the UGT71L1-KO line, suggesting that biosynthesis was interdependent with growth (Supplemental Figure S4.3). Studies in *P. tremuloides* have recognized that salicinoid production is constitutive and developmentally regulated, but strongly dependent on genotype (Osier et al., 2000; Osier and Lindroth, 2006). Here, we extend these observations to show that salicinoid accumulation continues even in the absence of photosynthesis and when trees are near death from carbon limited growth. A potential biological role for increased salicinoid concentration during aspen suckering is for defense against mammalian herbivory. Analysis of aspen regrowth in a beaver pond found that an unidentified phenolic compound, which was present in higher concentrations in smaller aspens, was a deterrent to beaver feeding (Basey et al., 1988). This unidentified phenolic may have been a salicinoid. High salicinoid content is thought to be most important to survival in early life stages when herbivore impact is greatest (Stevens and Lindroth, 2005; Lastra et al., 2017) a relationship that has been shown between browsing and defense for other boreal trees such as paper birch (*Betula papyrifera)* (Bryant et al., 2009).

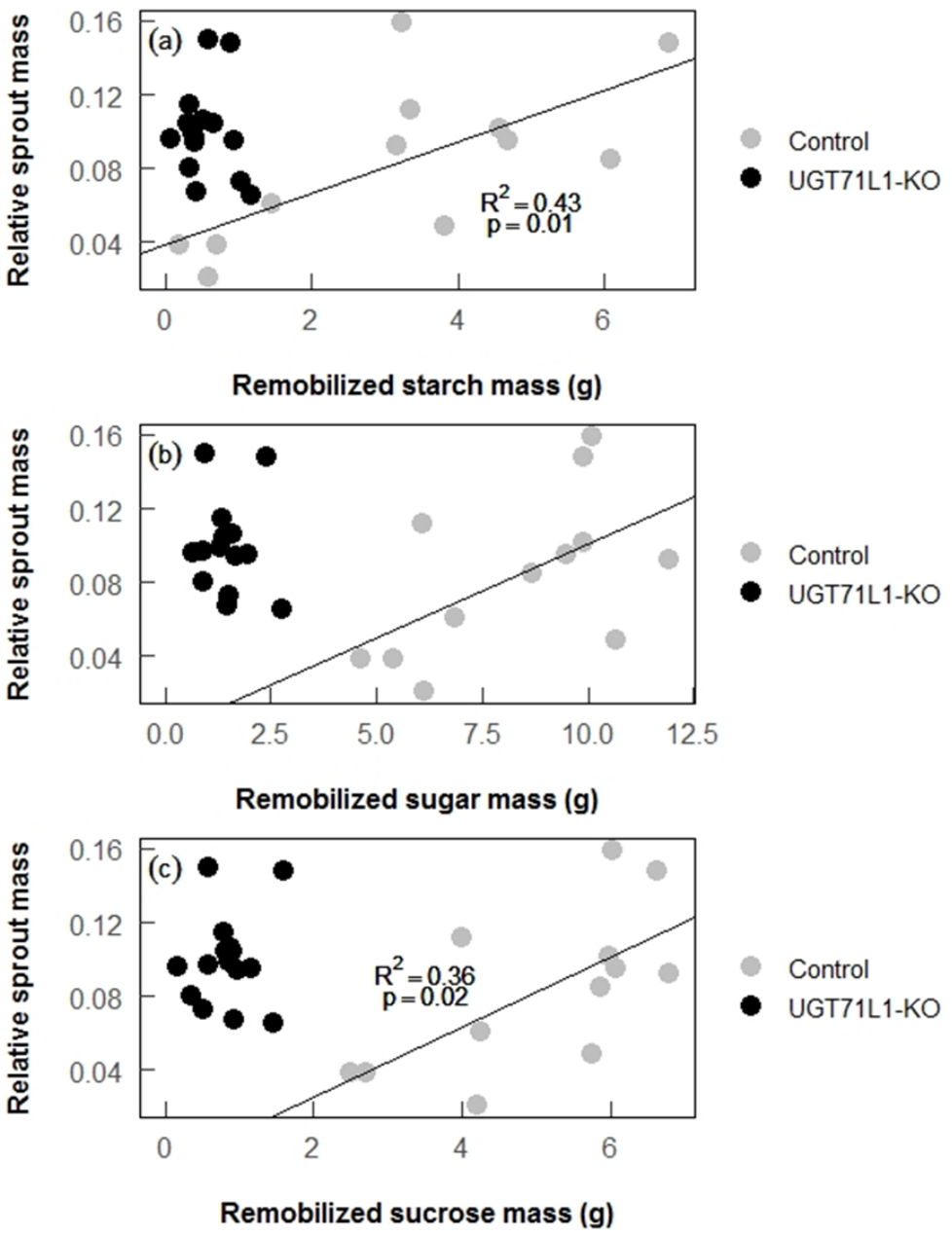
Our results demonstrated a clear lack of tremulacin and salicortin concentration plasticity when carbon is limited. Salicin appears to be under a separate regulatory mechanism. The pattern of salicin accumulation in light-grown plants was distinct from accumulation of salicortin and tremulacin. Although UGT71L1-KO lines have overall

reduced salicin levels, this increased significantly in shoots grown in light (Fig. 4a), despite disruption of UGT71L1. This increase was not detected in light-grown control plants. The high concentration of salicin observed in the light-grown UGT71L1-KO sprouts suggests that that additional UGT71L1-like enzymes must be active in this tissue to support salicin biosynthesis. For example, UGT78M1 has similar substrate preferences to UGT71L1 and may contribute to salicin biosynthesis in aspen sprouts (Fellenberg et al., 2020). In addition, the substantial accumulation of salicin in light-grown UGT71L1-KO plants suggests that the regulatory mechanisms for salicin synthesis and accumulation are perturbed, or that under these conditions salicin may act as an overflow for salicinoid flux. Transcriptomic analysis of resprouting UGT71L1-KO lines may prove useful in identifying salicin regulators or additional biosynthetic genes.

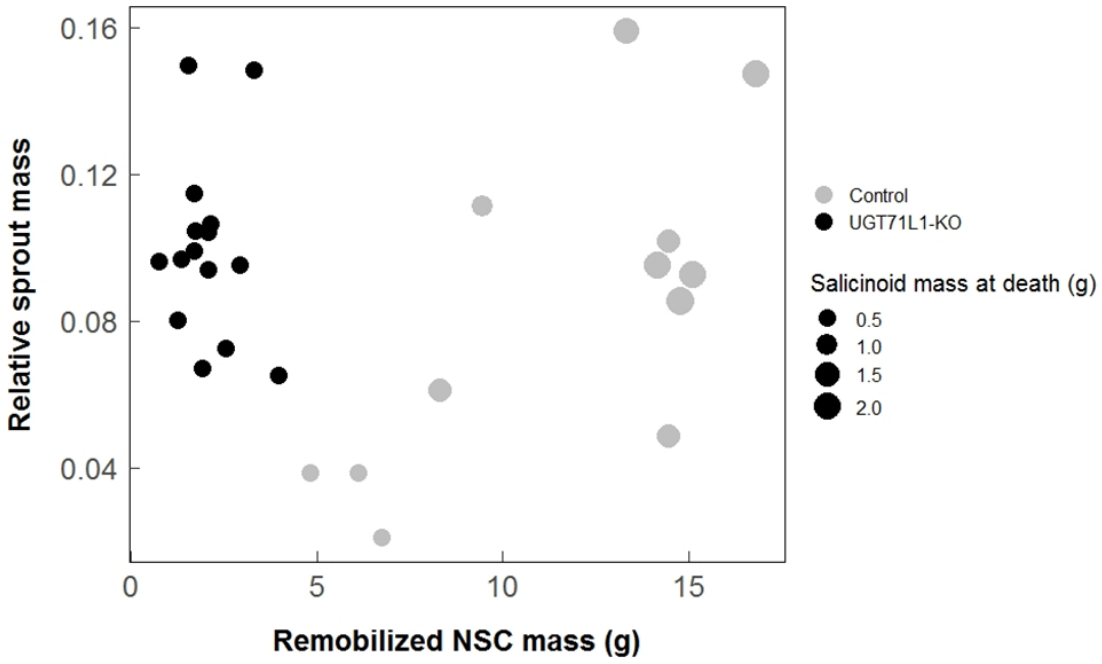
Overall, salicinoid concentrations remained relatively stable in shoot and root tissues during severely carbon-limited conditions. Therefore, it appears that the glucose moiety of salicinoids is a non-accessible carbohydrate sink. To date there is no evidence for transport of salicinoids between root and shoot tissues, although further investigation may provide additional insight. In summary, our experiment demonstrates that i) the glucose moiety of salicinoids is not catabolized or remobilized from roots to fuel growth as an energy source during carbon limitation, ii) aspen tissues maintain constitutive production of salicinoids even during carbon deprivation, thus reducing the efficiency of converting root biomass into shoot tissue, and iii) *P. tremula x P. alba* roots show distinct salicinoid profiles from shoots, and accumulate predominantly salicortin rather than tremulacin.

4.5 Supplementary Material

Supplemental Figure S4.1. UGT71L1-KO individuals harvested at the death stage. Dead shoot tissue of UGT71L1-KO plants tended to desiccate and wither while remaining flexible rather than darken. Photograph taken by Rachel Hillabrand.



Supplemental Figure S4.2. Relationships between relative spout mass and remobilized starch (a), sugar (b), and sucrose (c). Linear regression represents relationship of control lines with relative sprout mass. Non-significant ($\alpha = 0.05$) linear relationships are not presented.



Supplemental Figure S4.3. Relative sprout mass as a function of remobilized NSC mass with salicinoid mass at death indicated by the size of each point. NSC, non-structural carbohydrates.

Supplemental Table S4.1. Records of samples collected for tissue subsets during salicinoid analysis.

No. carbon limited biological replicates			
	Initial[1]	**End-of-growth**	**Death**
UGT71L1-KO shoots	-	10	10
UGT71L1-KO roots	14	10	10
Control shoots	-	9	9
Control roots	14	9	9
No. light-grown biological replicates			
	Initial[1]	Final	
UGT71L1-KO shoots	-	6	
UGT71L1-KO roots	14	6	
Control shoots	-	5	
Control roots	14	5	

[1] Initial analysis of samples was pooled for both light and dark sprouted samples as treatment had not begun.

Chapter 5 – Disruption of UDP-glycosyltransferase UGT78M1 using CRISPR/Cas9 promotes the accumulation of salicyl benzoate glucoside in poplar

This chapter is an ongoing research project intended for eventual publication and herein the roles and contributions of the UDP-dependent glycosyltransferase UGT78M1 to salicinoid biosynthesis and benzenoid metabolism in poplar is investigated. Experiments were conducted by Harley Gordon and the chapter was written by Harley Gordon. Identification of a putative gRNA site was conducted by Christin Fellenberg. Additional UGT78M1 in vitro substrate screening was not included in this chapter; however, the preliminary data is presented in Appendix 6.

5.1 Introduction

Salicaceae is a family of woody plants that includes poplars, cottonwoods, willows, and aspens. Salicaceous species produce salicinoid phenolic glycosides (salicinoids). Salicinoids are a group of specialized metabolites consisting of over thirty-five different compounds that can function as anti-herbivory molecules (Boeckler et al., 2011; Keefover et al., 2014; Rubert-Nason et al., 2015). Salicinoids are defined by a central glucosylated salicyl alcohol, which is often modified through the addition of benzenoid, coumaroyl, 6-hydoxy-2-clyclohexen-on-oyl (HCH), and acetyl groups (Keefover, et al., 2014). Aspens such as *Populus tremuloides* possess relatively high concentrations of salicinoids, which can accumulate to approximately 20% dry weight in young leaves (Donaldson et al., 2006). The aspens *P. tremula* and *P. alba* are both abundant, widespread, and ecologically important Salicaceae species in Eurasia and North America. Salicinoid production by poplars contributes to the global carbon cycle and ecosystem biodiversity (Rogers et al., 2020).

Isotopic and radioisotopic in vivo labelling experiments have demonstrated that salicinoids are produced from the phenylpropanoid pathway. Cinnamic acid provides an early precursor for the aromatic and aromatic derived components of the salicinoids

salicin and salicortin (Zenk, 1967; Babst et al., 2010). Cinnamic acid requires *ortho*-hydroxylation and side-chain shortening, prior to incorporation into the salicyl alcohol or HCH component of salicortin (Boeckler et al., 2011; Lackus et al., 2021). Benzoic acid also contributes to salicortin biosynthesis (Chapter 3). The dibenzenoid benzyl benzoate is synthesized from benzyl alcohol and benzoyl-CoA by a BAHD acyltransferase (Chedgy et al., 2015). Benzyl benzoate is likely a central intermediate molecule in salicortin biosynthesis (Babst et al., 2010; Chapter 3). The acyltransferase enzymes capable of salicyl benzoate and benzyl benzoate biosynthesis were designated as SABT (benzoyl-CoA: salicyl alcohol O-benzoyltransferase) and BEBT (benzoyl-CoA: benzyl alcohol O-benzoyltransferase). Salicyl benzoate appears connected to salicinoid biosynthesis but may not be a direct precursor (Chapter 3). Coexpression analysis of available poplar transcriptomic data sets with PtSABT and PtBEBT acyltransferase genes generated a partially overlapping list of coexpressed genes which is part of an ongoing research program to elucidate the complete salicinoid biosynthetic pathway (Chedgy et al., 2015; Fellenberg et al., 2020).

Two glycosyl transferase genes (designated as UGT71L1 and UGT78M1) strongly coexpressed with SABT (Fellenberg et al., 2020). Both UGT71L1 and UGT78M1 were identified as putative salicinoid biosynthetic enzymes. UGT71L1 had been previously suggested as a potential salicinoid biosynthetic gene by Tsai et al. (2011); whereas UGT78M1 had not been predicted as a salicinoid biosynthetic gene. UGT78M1 had been previously identified in poplar through coexpression with condensed tannin biosynthesis, however no in vitro glucosylation activity of UGT78M1 was found on flavonoid substrates (Table 5.1) (Veljanovski and Constabel, 2013). Both UGT71L1 and UGT78M1 demonstrated peak activity at pH 7, with glucosylation substrate preference for salicyl benzoate, and minor activity on salicyl alcohol, *ortho*-hydroxy cinnamic acid and salicyl aldehyde. UGT71L1 and UGT78M1 enzymes only share 28% amino acid similarity. However, the recombinant proteins have the same in vitro substrate preference (Fellenberg et al., 2020). Recent exploration of additional substrates demonstrated glycosylation activity of UGT78M1 and UGT71L1 on the hydroxylated dibenzenoids salicyl salicylate and benzyl salicylate (Gordon et al., 2022).

CRISPR/Cas9 genome editing was used to create functional knockouts of UGT71L1 in poplar hairy roots (Fellenberg et al., 2020) and whole plants (Gordon et al., 2022). Poplar hairy roots with disrupted UGT71L1 had no detectable salicortin and tremulacin, in contrast to control hairy roots (Fellenberg et al., 2020). Poplars with a functional knockout of UGT71L1 (UGT71L1-KO) showed a 95% reduction in salicortin and tremulacin concentrations in leaves. Salicin, the simplest salicinoid, was only reduced by 70%. Therefore, redundant salicinoid glycosylation activity is present in poplar. Significant chemotypic and phenotypic effects are also present in UGT71L1 knockout (KO) plants. UGT71L1-KO plants exhibit stunted growth, altered leaf morphology, and significant reductions in other phenolic compound concentrations (Gordon et al., 2022). A functional UGT71L1 thus appears essential for the normal growth, development, and defense of poplar trees. Given the residual production of salicinoids in UGT71L1-KO plants, and the similar substrate preference of UGT71L1 and UGT78M1, UGT78M1 remains a candidate for contributions to salicinoid biosynthesis. Despite strong evidence based on enzyme activity and coexpression analysis the role of UGT78M1 in salicinoid biosynthesis is doubtful. Furthermore, *Salix purpurea* synthesizes tremulacin and salicortin; but the *S. purpurea* genome does not contain a UGT78M1 homolog (Kulasekaran et al., 2020). This suggests that UGT78M1may not be a primary contributor to salicinoid biosynthesis during nominal growth. No functional disruption of UGT78M1 has yet been achieved in poplar. The role of UGT78M1 in salicinoid biosynthesis or benzenoid homeostasis is unestablished. As UGT71L1-KO plants maintain reduced salicin concentrations the disruption of UGT78M1 may eliminate the remaining salicin production capacity of UGT71L1-KO poplars.

Table 5.1. Glucosylation substrates for recombinant UGT78M1

Compound[1]	UGT78M1 Activity[2]	Reference
2,4,5 trichlorophenol	+	Veljanovski and Constabel, 2013
kaempferol	-	Veljanovski and Constabel, 2013
quercetin	-	Veljanovski and Constabel, 2013
catechin	-	Veljanovski and Constabel, 2013
cyanidin	-	Veljanovski and Constabel, 2013
salicyl alcohol	+	Fellenberg et al., 2020
salicyl aldehyde	+	Fellenberg et al., 2020
salicyl benzoate	++	Fellenberg et al., 2020
2-hydroxycinnamic acid	-	Fellenberg et al., 2020
salicylic acid	-	Fellenberg et al., 2020
benzyl salicylate	++	Gordon et al., 2022
salicyl salicylate	++	Gordon et al., 2022

[1] Recombinantly expressed UGT78M1 assessed in pH 7 solution and analyzed with HPLC-DAD.
[2] Activity is presented in as a +, active, ++, highly active, and -, no activity detected.

CRISPR/Cas9 has been used to disrupt UGT71L1 in *P. tremula x P. alba* with no detected off-target genome mutations (Gordon et al., 2022). The published *P. tremula x P. alba* genome can be used to select gRNA target sequences designed to eliminate functionality in UGT78M1 (Xue et al., 2015; Fellenberg et al., 2020). Our objective was to understand the biological role and salicinoid biosynthetic contribution of UGT78M1. We successfully utilized CRISPR/Cas9 genome editing to disrupt UGT78M1 in the hybrid research poplar *P. tremula x P. alba*, as well as the UGT71L1-KO lines to generate double knockout plants. We hypothesized that UGT78M1 contributes partially to the biosynthesis of salicin, salicortin, and tremulacin. Through phytochemical analysis we found that UGT78M1 disruption did not affect salicortin or tremulacin biosynthesis, but may have had a minor effect on salicin biosynthesis. Notably, UGT78M1 disruption results in significant accumulation of salicyl benzoate glucoside, an established in vitro substrate for UGT78M1.

5.2 Materials and Methods

5.2.1 Sequence analysis of the UGT78M1 gRNA site and introduced mutations

The guide RNA (gRNA) sequence site for the disruption of UGT78M1 was selected by searching potential gRNA sequences within exons of UGT78M1 using the *P. tremula x P. alba* sPta717aV1 database (https://www.aspendb.org). After site selection we confirmed the genomic sequence of the *P. tremula* and *P. alba* UGT78M1 alleles was conducted. Genomic DNA was isolated from in vitro cultures of wild-type *P. tremula x P. alba* INRA 717-1B4. Approximately 100 mg of leaf tissue was collected for genomic DNA extraction using the DNeasy Plant Mini kit (Qiagen, Toronto, ON, Canada). For polymerase chain reaction (PCR) amplification of the gRNA target sequence, 2 µL of genomic DNA was used as a template. Amplicons were generated using a forward (TGCATTTCTTTGGTTTGCAG) and reverse (CATATCTTCCAAGCCTAGTTCG) oligonucleotide primer (Integrated DNA Technologies, Coralville, IA, USA). PCR thermocycler settings were 1: 98 ºC, 1 min; 2: 98 ºC, 20 s; 3: 59.8 ºC, 30 s; 4: 72 ºC, 1 min; 5: repeat steps two to four for 32 cycles; 6: 72 ºC, 5 min. PCR products were purified using an EZ-10 Spin Column PCR Product Purification Kit (Biobasic, Markham, ON, Canada). Prior to cloning and sequence analysis purified PCR products (approx. 300 ng) were A-tailed using 1mM dATP, 20 mM $MgSO_4$, and Taq Polymerase with Taq Buffer (Biobasic). A-tailing reactions were incubated at 72ºC for 20 min. 5 µL of A-tailed PCR product was ligated into a pUCM-T vector overnight at 4 ºC. Ligated pUCM-T plasmids were transferred into *Escherichia coli* strain: XL-1 Blue (Agilent Technologies, Mississauga, ON, Canada) and transformants were selected using ampicillin resistance and blue/white colony screening. Positive colonies were cultured in 5 mL of Luria-Bertani broth (LB) containing 100 $mg \cdot L^{-1}$ carbenicillin. Plasmids were purified from cultures using an EZ-10 Plasmid Purification Kit (Biobasic). Purified plasmids were Sanger sequenced using a gene specific sequencing primer (CAAATCACAGGCACACATCC) designed to amplify the selected gRNA site. Separate tremula and alba alleles were identified based on single nucleotide

polymorphism alignment using the in silico genome of *P. tremula x P. alba* (Xue et al., 2015). Sequence alignment was conduced using Clustal Omega (Madeira et al., 2022).

The same genomic DNA amplification methods were used for analysis of CRISPR/Cas9 induced mutations in transformed lines. The target region of UGT78M1 was amplified by PCR using a Q5 polymerase and the oligonucleotide primers as described above. Successful reactions were cleaned, and Sanger sequencing reactions were assessed using the Tracking of Indels by Decomposition (TIDE) webtool (Brinkman et al., 2014; Gordon et al., 2022; http://shinyapps.datacurators.nl/tide/). Indels greater 50 bp were assigned following manual assessment of sequencing traces that had non-consensus sequencing reactions and the absence of successful mutation identification through the TIDE webtool. A sample of genomic DNA from UGT78M1 transformants was subcloned and sequenced to confirm TIDE indel identification accuracy. In brief, clean PCR products were subcloned into a pUCM-T cloning vector and transformed into *E. coli* as described above. Plasmid preparations were Sanger sequenced using the previously mentioned UGT78M1 sequencing primer which allowed for confirmation of individual *P. tremula* and *P. alba* alleles in a subset of independent transformants (Fig. 5.1B).

5.2.2 Vector construction and plant culture

CRISPR/Cas9 edited UGT78M1 knockouts and UGT71L1/UGT78M1-KO double knockout plants were generated using a p201H vector (Jacobs et al., 2015) containing a gRNA sequence for UGT78M1 selected and confirmed as above (gRNA sequence: AAGCAGAAACACGAATCTG). The p201H vector was constructed as previously described (Fellenberg et al., 2020) (Supplemental Figure S5.1.). In brief, p201H was linearized using restriction enzymes SpeI and SwaI (New England Biolabs, Whitby, ON, Canada) and the DNA sequence for encoding the gRNA was ligated into the linear vector using Gibson assembly cloning (Gibson et al., 2010). Circular plasmids were electroporated into *E. coli* XL1-Blue (Agilent) and a positive colony was selected. The p201H:gRNA plasmids were purified from a bacterial culture using commercial kits

(Biobasic) Sequence integrity was confirmed using Sanger sequencing. The p201H:gRNA plasmid was inserted into *Agrobacterium tumefaciens* strain GV3101 via electroporation.

For transformation, poplar leaves from in vitro cultures of *P. tremula* x *P. alba* (INRA 717-1B4) were excised and used for Agrobacterium transformation as previously described (James et al., 2017). Explant leaves from UGT71L1-KO poplar lines were used for the generation of double knockouts (Gordon et al., 2022). Excised leaves were wounded using a scalpel and infiltrated with an *A. tumefaciens* GV3101 solution containing the p201H construct with gRNA by incubating the leaves in an *A. tumefaciens* culture for 3 h. Leaves were blotted dry and placed on half strength Murashige and Skoog (MS) medium in petri plates containing 4.4 $g \cdot L^{-1}$ MS Basal Salts, 30 $g \cdot L^{-1}$ sucrose, 0.2 $g \cdot L^{-1}$ glutamine, 5 $g \cdot L^{-1}$ agar, 100 $\mu g \cdot L^{-1}$ nicotinic acid, 100 $\mu g \cdot L^{-1}$ pyroxidine HCl, 100 $\mu g \cdot L^{-1}$ thiamine HCl, 100 $\mu g \cdot L^{-1}$ calcium pantothenate, 5 µM isopentenyladenine and 10 µM 1-napthaleneacetic acid for 14 days. Leaves were then placed on a selective callus induction medium containing all above ingredients with the addition of 100 $\mu g \cdot L^{-1}$ cefotaxime, 100 $\mu g \cdot L^{-1}$ carbenicillin and 10 $\mu g \cdot L^{-1}$ hygromycin. Following callus growth, calli were transferred to shoot induction media containing 4.4 $g \cdot L^{-1}$ MS Basal Salts, 30 $g \cdot L^{-1}$ sucrose, 0.2 $g \cdot L^{-1}$ glutamine, 5 $g \cdot L^{-1}$ agar, 100 $\mu g \cdot L^{-1}$ nicotinic acid, 100 $\mu g \cdot L^{-1}$ pyroxidine HCl, 100 $\mu g \cdot L^{-1}$ thiamine HCl, 100 $\mu g \cdot L^{-1}$ calcium pantothenate, 0.5 µM indole-3-butyric acid (IBA), and 0.4 $\mu g \cdot L^{-1}$ thidiazuron, 100 $\mu g \cdot L^{-1}$ cefotaxime, 100 $\mu g \cdot L^{-1}$ carbenicillin and 10 $\mu g \cdot L^{-1}$ hygromycin and the cultures were placed in an environmentally controlled chamber with a photoperiod of 16 h and a temperature of 23 ºC. Generated shoots were selected on woody plant medium containing 10 $\mu g \cdot L^{-1}$ hygromycin (James et al., 2017). One independent empty-vector control line expressing the Cas9 endonuclease and hygromycin resistance with no gRNA was generated, and an independent double empty vector line transformed with both p201N and p201H was also generated.

5.2.3 Plant Growth

Hygromycin resistant transformed lines were grown as in vitro cultures until roots were well established. Plantlets were then transferred to soil in two-inch pots and grown in a mist chamber for three-weeks. Plants were then transferred to one-gallon pots containing Sunshine Mix 4 soil (Sungro, Seba Beach, AB, Canada). Soil was supplemented with slow-release fertilizer blend containing of 8.9 $g \cdot L^{-1}$ controlled release 8-6-12 NPK plus micronutrients (Acer, Delta, BC, Canada), 0.46 $g \cdot L^{-1}$ superphosphate 0-20-0 (Green Valley, Surrey, BC, Canada), 1.21·g L^{-1} Micromax Micronutrients (Scotts-Sierra, Marysville, OH, USA), and 4.75 $g \cdot L^{-1}$ Dolomite lime (IMASCO, Surrey, BC, Canada) (Major and Constabel, 2006). Plants were grown in a greenhouse with supplemental lighting provided by 600 W high pressure sodium bulbs delivering an average intensity of 300 $\mu mol \cdot m^{-2} \cdot s^{-1}$ for a 16 h photoperiod. Plants were irrigated as required and rotated within the greenhouse weekly. Leaves were numbered starting at one from the terminal peak where leaf lamina breadth was greater than 1 cm across. Leaves #6-10 were collected, the midvein excised, and flash frozen in liquid nitrogen. Leaves were stored at -80 ºC prior to metabolite extraction and analysis.

5.2.4 Metabolite extraction and analysis

Frozen material from the #6-10 harvested leaves was combined and homogenized under liquid nitrogen with a mortar and pestle. Frozen tissue (~100 mg) was extracted with 1 mL of chilled methanol. The leaf-methanol suspension was sonicated for 2 min and then shaken on a benchtop orbital shaker for 30 min. Extracts were vortexed briefly, and centrifuged at 16,000 *g* for 2 min. The supernatant (800 µL) was placed in a deep 1 mL 96-well plate and stored at -20 ºC prior to analysis. Methanol extracts were separated and analyzed using established methodology on a Waters Acquity UPLC-MS (Waters, Mississauga, ON, Canada) (Chapter 4). A 2 µL injection was separated using a Waters Acquity Ethylene Bridged Hybrid (C18, pore size: 1.7 µm, dimensions: 2.1 x 50 mm) column at 40 ºC. A binary solvent gradient separation was used with water containing 0.1% formic acid (A) and LC/MS grade acetonitrile containing 0.1% formic acid (v/v)

(B). The flow rate was 0.5 $mL \cdot min^{-1}$ and the separation gradient was 99.9% solvent A, 0-0.5 min; 99.9-80.0 % A, 0.5-6.0 min; 80.0-50.0% A, 6.0-9.0 min; 50.0-10.0% A, 9.0-10.0 min; 10% A, 10.0-11.5 min; 10-99.9% A, 11.5 – 12.0 min; 99.9% A, 12.0-13.0 min. A single quadrupole MS in electrospray ionization negative mode was used to detect and quantify all analytes. The MS conducted a full scan from 70.00 to 850.00 m/z with a continuous cone voltage of 10 V. Compound identities were assigned using an in-house library based on standards (salicin, salicortin, SA, tremulacin), enzyme assay reactions (SB-Glc, BS-Glc, SS-Glc, BA-Glc, SA-Glc), m/z values, and retention times. A table of m/z values and retention times used in this study is presented in Supplemental Table T5.1. To confirm the identification of a salicyl benzoate glucoside UPLC-MS analysis was conducted on in vitro enzyme assays prepared using recombinant UGT71L1 and salicyl benzoate (Supplemental Figure S5.3) (Gordon et al., 2022). Peak integration was conducted using MassLynx v4.2 (Waters). Tremulacin was quantified using linear regression of an authentic standard peak areas provided by Dr. Richard Lindroth. Salicortin was quantified using a linear regression of an authentic commercial standard (Biosynth, Staad, SG, Switzerland). Salicin was quantified using commercially available standards from Sigma-Aldrich (Oakville, ON, Canada). Isosalicin was quantified using the salicin standard curve. SA-Glc was quantified using a salicylic acid standard curve. Additional compounds are presented in MS response $area \cdot mg\ FW^{-1}$ unless otherwise noted.

For data interpretation and statistical analysis, the growth and metabolite data from three independent transformed UGT78M1-KO and three independent UGT71L1/UGT78M1-KO lines was pooled. The pooled data is presented in figure 5.2 and figure 5.3 and represents a subsample of independently transformed lines.

5.3 Results

5.3.1 Disruption of UGT78M1 in *P. tremula* and *P. alba* alleles was achieved using CRISPR/Cas9 genome editing

To explore the biological role of UGT78M1 in wild-type and UGT71L1-KO poplars multiple independent knockout lines were generated in both genetic backgrounds. Transformation and hygromycin selection of CRISPR/Cas9 knockout lines effectively selected for positive transformants containing interruptions in the UGT78M1 coding sequence. Double transformation using the p201H vector in UGT71L1-KO poplar lines previously transformed with p201N was successful. Sixteen independent UGT78M1-KO lines and twelve independent UGT71L1/UGT78M1-KO double knockout lines were generated. Ten of the twelve independent UGT71L1/UGT78M1-KO lines and nine of the sixteen independent UGT78M1-KO lines were genotyped using Sanger sequencing to evaluate the CRISPR/Cas9 induced mutations. The most common insertions and deletions (indels) consisted of +1 bp insertions, which comprised 44% of identified mutations. Six larger (± 50 bp or larger) unidentified indels were assigned based on Sanger sequencing results (Fig. 5.1). Sequencing data demonstrates hygromycin selection of CRISPR/Cas9 modified plant lines is effective in selecting for a range of induced mutations in the UGT78M1 genomic site as all nineteen genotyped lines presented at least one mutated allele (Fig. 5.1). The transformation of UGT71L1-KO lines with p201H does not appear to be impacted by the CRISPR/Cas9 construct previously introduced into the plant genome. The in silico sequence of UGT78M1 from aspenDB (https://www.aspendb.org/) shows two copies of UGT78, labelled as UGT78M1 and UGT78M2. However, the detection of two UGT78 genes was never achieved using our UGT78 gRNA sequencing primers. Despite multiple independent transgenic lines, we failed to amplify more than a single tremula and single alba gRNA site from individual plants. Therefore, we assumed the in silico duplication was an artifact of genome assembly and not present in *P. tremula x P. alba*.

Chemical phenotyping by UPLC-MS was used to screen transformed lines. Each transformed line was grown in duplicate. To improve our capacity for statistical and

chemical phenotype interpretation, lines were grouped based on phytochemical quantification of salicinoids and salicyl benzoate glucoside (Supplemental Figure S5.3). Three independent transgenic lines were combined for statistical analysis of growth and phenotypic data. Growth and analytical data from the UGT78M1 single knockout lines; 78KO-5, 78KO-10, and 78KO-14 were combined. For the UGT71L1/UGT78M1-KO double knockout lines data from 71KO 78KO-8, 71KO 78KO-5, and 71KO 78KO-4 were pooled. The selected lines have been propagated and will be grown with additional biological replicates grown for future work.

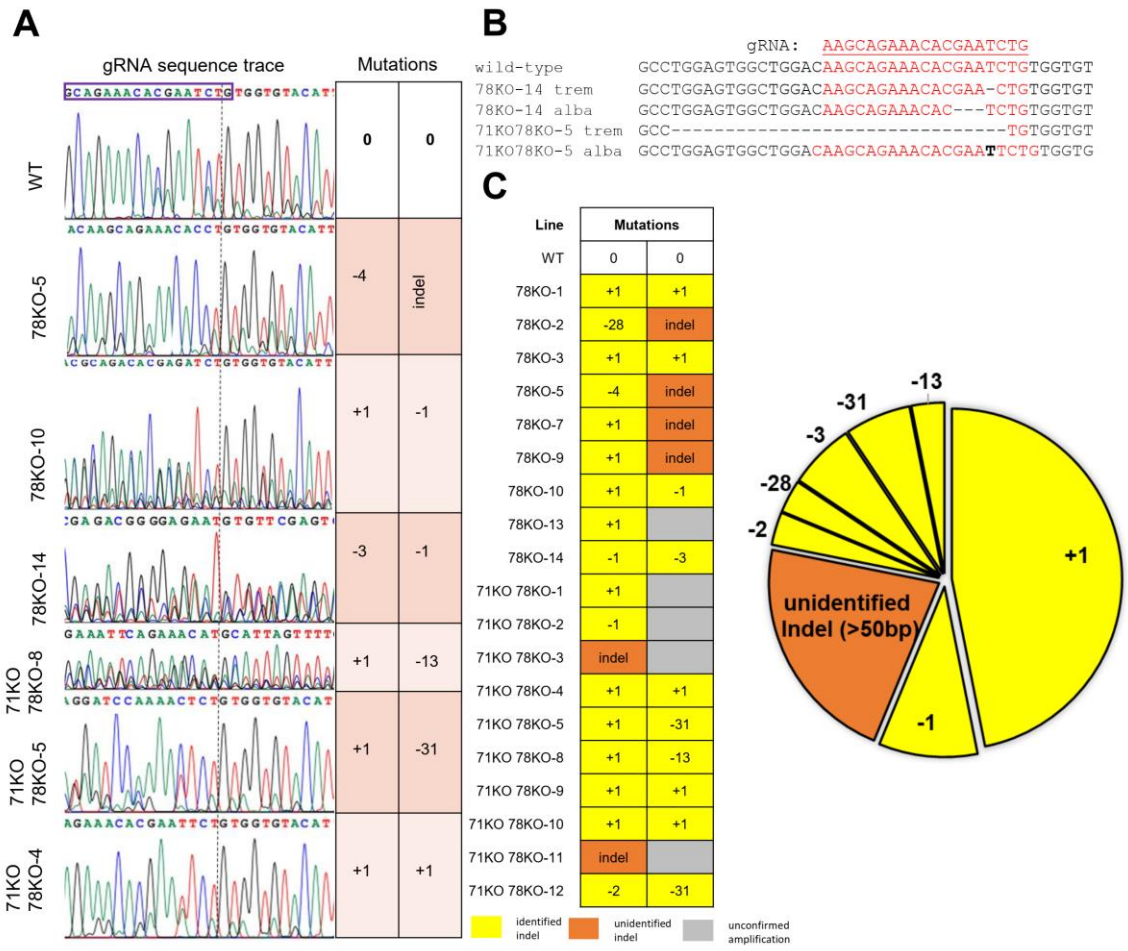


Figure 5.1. Mutation analysis and summary of UGT78 and UGT71L1/UGT78M1 knockout lines. A, Sanger sequencing traces of PCR amplicons from three combined UGT78M1-KO and three combined UGT71L1/UGT78M1-KO lines. The gRNA sequence is framed at the upper left portion of the figure. A vertical dashed line crosses through protospacer adjacent motif (PAM) of the gRNA sequence. The table on the right lists the indels corresponding to the adjacent sequencing traces identified through the TIDE webtool (Brinkman et al., 2014). B, Sequencing results for subcloned and sequenced PCR products from select transgenic lines. Confirmation of mutations in both the *P. tremula* (trem) and *P. alba* (alba) allele of UGT78M1 are presented. C, Table of all detected mutations in sequencing reactions of PCR amplified UGT78M1 from independently transformed lines. The pie chart presents the proportion of detected mutation among sequenced independent transformants.

5.3.2 Disruption of UGT78M1 does not appear to alter growth in wild-type poplar

Greenhouse grown UGT78M1-KO lines were monitored for macro-phenotypic growth effects. UGT78M1-KO lines remained visibly consistent with wild-type poplars throughout one-month of greenhouse growth. UGT78M1 disruption in UGT71L1-KO or WT lines did not appear to alter plant height from the genetic background (Fig. 5.2). By contrast, the double knockout UGT71L1/UGT78M1-KO lines had significantly more leaves per plant with an average leaf count of 26 compared to the average of 16 in UGT71L1-KO lines. Disruption of UGT78M1 in wild-type poplars did not appear to significantly alter plant height or the total number of leaves. The double knockout UGT71L1/UGT78M1-KO lines were visually consistent with UGT71L1-KO lines, but collected data revealed the difference in leaf number. UGT71L1/UGT78M1-KO lines exhibit reduced height compared to wild-type plants and have a slender, crinkled leaf morphology consistent with the UGT71L1-KO lines (Gordon et al., 2022). Representative leaf images are presented in Fig. 5.2B.

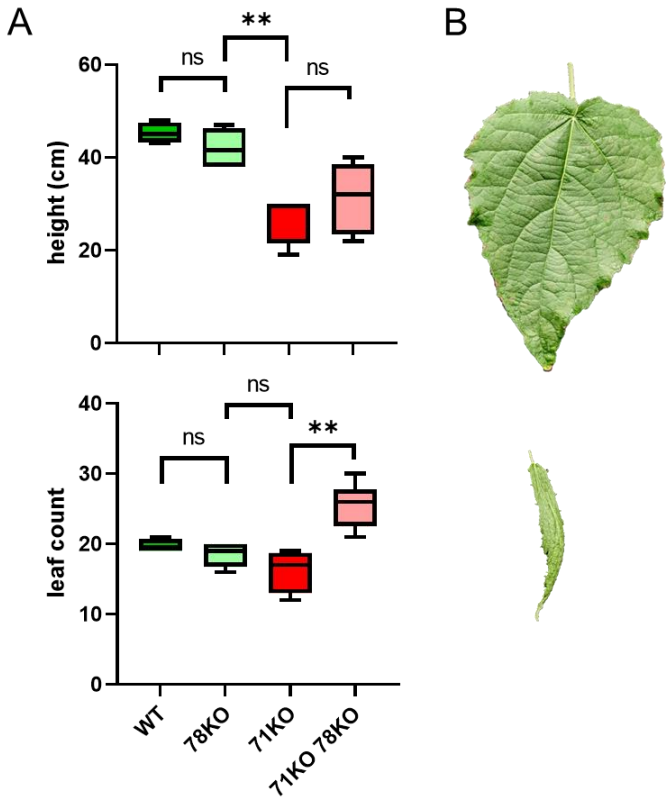


Figure 5.2. Height, leaf count, and leaf morphology, from one-month old greenhouse grown UGT78M1 knockout poplars. A; Height (top) and leaf count (bottom) of greenhouse grown poplars. Leaf count began from terminal leaves with a lamina breadth of > 1 cm. Solid black horizontal line is the median. Biological replicates: WT = 4; UGT78M1-KO = 3 independent lines with n = 2 each (6 total plants); UGT71L1-KO = 4; UGT71L1/UGT78M1-KO = 3 independent lines with n = 2 (6 total plants). Statistical comparisons were conducted using a Welch's *t*-test, with ns = non-significant; * = $P < 0.05$; ** = $P < 0.01$. Abbreviations: WT, wild-type; KO, knockout; 78, UGT78M1; 71, UGT71L1. B, upper leaf is representative of a UGT78M1-KO line, bottom leaf is representative of UGT71L1/UGT78M1-KO lines.

5.3.3 Disruption of UGT78M1 increases salicyl benzoate glucoside concentration but has a minimal impact on salicinoid concentrations

UPLC-MS analysis of targeted phenolic molecules in leaf extracts across sixteen independent UGT78M1-KO and twelve independent UGT71L1/UGT78M1-KO lines revealed several phytochemical effects of UGT78M1 interruption. Preliminary observation showed an apparent increase in salicyl benzoate glucoside concentrations and a small reduction in salicortin and tremulacin concentrations. (Supplemental Figure S5.2). To conduct statistical analysis, the analytical results from the biological replicates of three independent UGT78M1-KO lines and three independent UGT71L1/UGT78M1-KO lines were pooled as described in Section 5.3.1 and recalculated (Fig. 5.3). Following pooled analysis salicin concentrations were found to be significantly reduced in UGT78M1-KO lines compared to wild-type *P. tremula x P. alba*. No significant reduction in salicin concentrations was observed between UGT71L1-KO lines and UGT71L1/UGT78M1-KO double knockout lines (Fig. 5.3A). Salicortin and tremulacin were not significantly reduced in UGT78M1-KO lines compared to wild-type plants. Additionally, no significant reduction of the concentrations of tremulacin and salicortin in UGT71L1/UGT78M1-KO lines was detected (Fig. 5.3B and Fig. 5.3C).

Despite no observed effects of UGT78M1 disruption on gross morphology a significant enrichment of salicyl benzoate glucoside was observed in both UGT78M1-KO and the UGT71L1/UGT78M1-KO double knockout lines (Fig. 5.3D). The accumulation of salicyl benzoate glucoside was consistent across sixteen independent UGT78M1-KO lines and twelve independent UGT71L1/UGT78M1-KO double knockout lines (Supplemental Figure S5.2). Significant differences in the accumulation of benzyl salicylate glucoside were not observed between UGT78M1-KO and wild-type lines or UGT71L1/UGT78M1-KO and UGT71L1-KO lines (Fig. 5.3E). The accumulation of salicylic acid glucoside was also not significantly altered in UGT78M1-KO lines. The UGT71L1/UGT78M1-KO double knockout lines maintain elevated concentrations of salicylic acid glucoside and presented no significant concentration differences from the UGT71L1-KO lines.

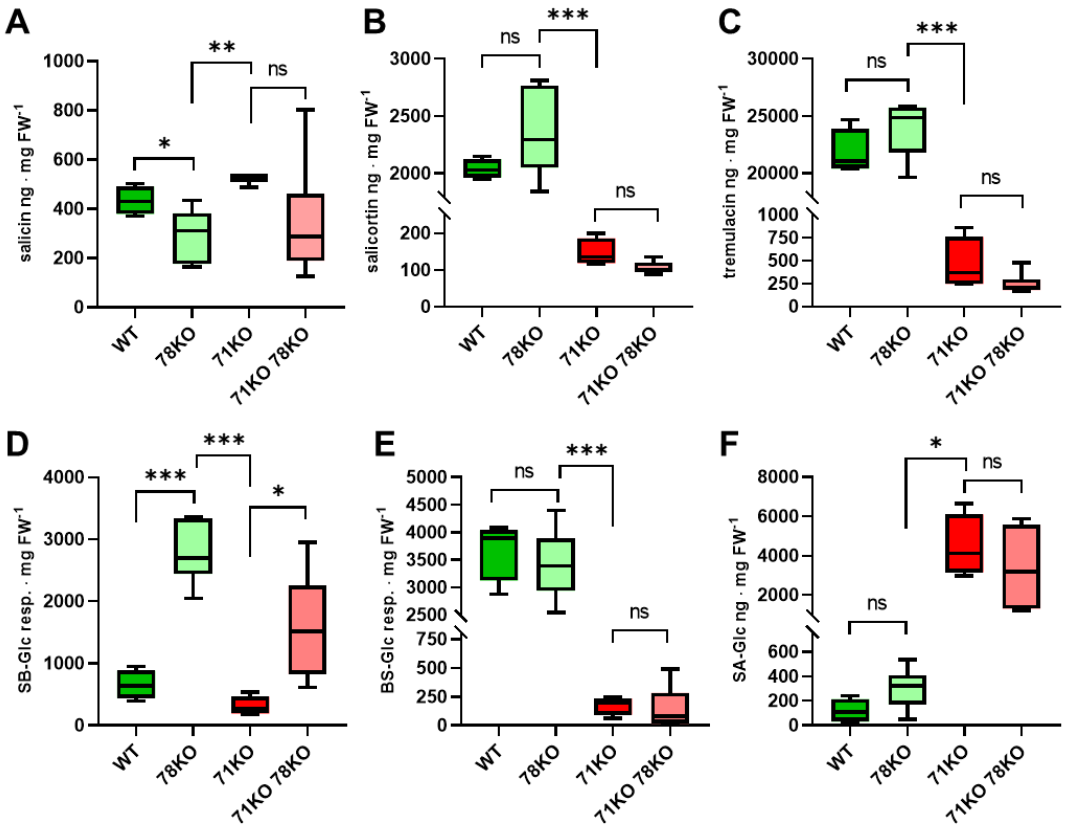


Figure 5.3. Analysis of targeted small molecules in one-month old UGT78M1 knockout poplar. Quantification of select phenolic analytes in UGT78M1-KO and UGT71L1/UGT78M1-KO lines. Solid black lines are the concentration median. A; salicin concentration; B, salicortin concentration; C, tremulacin concentration; D, salicyl benzoate glucoside (SB-Glc) concentration presented in MS response area·mg fresh weight $(FW)^{-1}$; E, benzyl salicylate glucoside (BS-Glc) concentration presented in peak area (response) ·mg $(FW)^{-1}$; F, salicylic acid glucoside (SA-Glc) concentration as calculated using a salicylic acid standard curve. Sample sizes represent individual plants: WT = 4; UGT78M1-KO = 3 independent lines with n = 2 for each biological replicate (6 total plants); UGT71L1-KO = 4; UGT71L1/UGT78M1-KO = 3 independent lines with n = 2 for each biological replicate (6 total plants). Statistical comparisons were conducted using a Welch's *t*-test, with ns = non-significant; * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$. Abbreviations: WT, wild-type; KO, knockout; 78, UGT78M1; 71, UGT71L1.

5.4 Discussion

5.4.1 Induced chemotypes in plants with disrupted UGT78M1 indicates successful genome modification despite uncertainty in UGT78 copy number

To study the biological role of UGT78M1, a gRNA construct and the Cas9 endonuclease was used to disrupt the in-frame genomic DNA sequence of UGT78M1 in *P. tremula x P. alba* plants (Fig. 5.1). Gene disruption was achieved in wild-type plants as well as UGT71L1-KO which was generated with a p201N CRISPR/Cas9 construct (Gordon et al., 2022). Although CRISPR/Cas9 multiplex genome editing using multiple gRNA sequences has successfully several genes in plants (Do et al., 2019; Bai et al., 2020), our work demonstrates the effectiveness of a secondary transformation in poplar which utilized a re-transformation with a full CRISPR/Cas9 transformation vector.

Nealy half of induced mutations identified indels contained a single base pair (bp) insertion (44%). Several variations indels were also identified at a lower rate and larger indels greater than 50 bp are likely present (Fig. 5.1). The abundance of +1 bp indels, and the two instances of -31 mutations, is likely a result of the specific gRNA sequence which influences double stranded breaks and the non-homologous end joining (NHEJ) repair mechanism. Elorriaga et al. (2018) demonstrated that in *P. tremula x P. alba* different gRNA sequences for the same gene influence the mutation profile. Interruption of a floral meristem identity gene using gRNA in poplar resulted in the detection of 40% -2 bp indels across numerous independent transformants, while a separate construct on the same gene only generated 11% -2 bp indels. Large CRISPR/Cas9 induced indels (> 50 bp) are a known consequences of the double stranded breaks induced by the Cas9 endonuclease (Elorriaga et al., 2018; Gordon et al., 2022). Fellenberg et al. (2020) had attempted to eliminate UGT78M1 and UGT71L1 in single and double knockout lines using a binary transformation vector. However, discrepancies within the available in silico *P. tremula x P. alba* and the in vivo genome resulted in initial gRNA sequence mismatching the UGT78M1 genomic sequence. Therefore, prior to gRNA sequence selection we subcloned the UGT78M1 alleles to confirm sequence fidelity. We identified inconsistencies in the in silico genomic sequence and our subcloned tremula and alba

allele amplicons. The UGT78 gene was published as a duplicated gene model in aspenDB (https://www.aspendb.org/) annotated as UGT78M1 (Potri.006G171100) and UGT78M2 (Potri.006G171200). We were unable to amplify and sequence more than a single UGT78, thus only a single UGT78 gene was found in each of the hybrid genomes. The *P. tremula x P. alba* genome may only contain a single copy of UGT78M1. The V1.1 *P. tremula x P. alba* genome published by AspenDB was assembled using v3 Of *P. trichocarpa* as a reference (Xue et al., 2015). However, a recent iteration of the *P. trichocarpa* genome V4.1 has changes in annotations and the coding region of UGT78M1 (Tuskan et al., 2006). The v3 P. trichocarpa genome used as a reference for *P. tremula x P. alba* genome construction may have contributed to in silico discrepancies identified in the UGT78 sequence. As we did not sequence the entire genomic region, UGT78M2 may still be present, but it would likely be incompatible with our amplification primers. Potential off target sites for the gRNA used were not sequenced. However, a recent search for potential gRNA off target sites revealed no sites with a single base-pair mismatch and 15 sites with a double base-pair mismatch (*P. tremula* x *P. alba* v5.1, DOE-JGI, http://phytozome.jgi.doe.gov/). The detection of a significant chemotype in multiple independent lines, despite only a single detected copy of disrupted UGT78 provides evidence for the presence of a single non-redundant copy of UGT78M1 in *P. tremula x P. alba*.

5.4.2 UGT78M1 is not a major contributor to salicin biosynthesis

As the salicinoid deficient UGT71L1-KO lines maintain a low concentration of salicinoids, UGT78M1 was a strong candidate for minor contributions to salicinoid biosynthesis. Both UGT71L1 and UGT78M1 are family 1 UDP-dependent glycosyltransferases that appear to have evolved for glucosylation of hydroxylated dibenzenoid substrates (Fellenberg et al., 2020; Gordon et al., 2022). UGT78M1 also has similar coexpression patterns compared to UGT71L1 (Fellenberg et al., 2020). Our work shows that disruption of UGT78M1 in UGT71L1-KO lines does not eliminate salicin biosynthesis or reduce the concentrations of salicortin or tremulacin. In single knockout

UGT78M1-KO lines salicin concentrations are significantly reduced but no reduction in salicortin or tremulacin is observed.

Our work thus provides the first direct evidence for the lack of UGT78M1 contributions in salicortin and tremulacin biosynthesis in poplar. UGT71L1 and UGT78M1 share an amino acid sequence similarity of only 28% but have similar substrate preference (Fellenberg et al., 2020; Gordon et al., 2022). The divergent amino acid sequences likely contribute to distinct in plant activity regardless of shared expression patterns and similar substrate preference. Protein localization may also contribute to differences in UGT78M1 and UGT71L1 in vivo activity. In silico queries for chloroplast, mitochondrial, or nucleus signal sequence peptides of both UGT78M1 and UGT71L1 returned no predicted subcellular localization (Sperschneider et al., 2017). The contribution of UGT78M1 to salicin accumulation may represent only a minor route as no significant reduction in salicin concentrations was observed during the disruption of UGT78M1 in UGT71L1-KO lines. Salicin biosynthesis is separate from salicortin and tremulacin biosynthesis (Gordon et al., 2022; Chapter 4). Therefore, UGT78M1 may have a minor contribution to salicin accumulation. However, an increase in sample size, and selection of numerous tissue ages may provide stronger confirmation to the role of UGT78M1 in salicin biosynthesis. The lack of UGT78M1 involvement in the biosynthesis of complex salicinoids is supported by the observations of Kulasekaran et al. (2020), who noted that UGT78M1 was not identified in the *S. purpurea* despite the synthesis of salicortin and tremulacin.

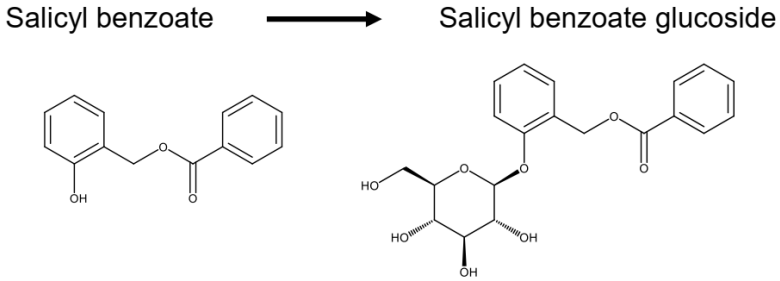
The significant increase in the number of leaves between UGT71L1-KO and UGT71L1/UGT78M1-KO lines was only apparent during statistical analysis and was not noted during routine plant maintenance (Fig. 5.2). Disruption of a family 1 UDP-dependent glycosyltransferase in Arabidopsis resulted in an increased rosette leaf count which was hypothesized to be connected to the effect of flavonoid glycosides on auxin distribution (Wang et al., 2012). Therefore, phytohormone analysis of UGT78M1-KO lines may reveal additional insights into any regulatory roles of UGT78M1. The glucosylated phytohormone salicylic acid glucoside was not impacted by UGT78M1 disruption. Therefore, unlike UGT71L1-KO lines, salicylic acid perturbations may not be causing the increased leaf number phenotype. The increased leaf count in

UGT71L1/UGT78M1-KO lines may be an amplification of the established UGT71L1-KO phenotype, which presented reduced internode length and increased branching.

5.4.3 UGT78M1 disruption results in the accumulation of the dibenzenoid salicyl benzoate glucoside

Recombinant UGT78M1 glucosylates salicyl benzoate, benzyl salicylate, and salicyl salicylate in vitro (Table 5.1). Unexpectedly, UGT78M1 disruption in *P. tremula* x *P. alba* resulted in the significant accumulation of salicyl benzoate glucoside. The UGT71L1/UGT78M1-KO double knockout lines also accumulate significantly more salicyl benzoate glucoside (Fig. 5.3). This result was unexpected as UGT78M1 displays optimal in vitro glucosylation activity on salicyl benzoate (Fellenberg et al., 2020). This overaccumulation of salicyl benzoate glucoside in UGT78M1 knockouts presents several open questions. Why does salicyl benzoate glucoside, a product of UGT78M1, increase in concentration when the UGT78M1 is disrupted? By contrast, in UGT71L1-KO lines salicyl benzoate glucoside is reduced in concentration as expected (Fig. 5.3)(Chapter 3). The presence of increased salicyl benzoate glucoside in UGT71L1/UGT78M1-KO lines indicates that *P. tremula x P. alba* must contain additional UGT enzymes capable of glucosylating the dibenzenoid salicyl benzoate (Fig. 5.4). Furthermore, the biological processes cause the 5-fold increase salicyl benzoate glucoside concentration appear relatively unaffected by the disruption of UGT71L1. The increased expression of UGT enzymes capable of glucosylating salicyl benzoate is a possible mechanism for compound overaccumulation. However, the accumulation of precursor molecules involved in salicyl benzoate biosynthesis may also result in the elevated accumulation of the aglycone, which could be glucosylated through yet unidentified glycosyltransferase enzymes. If salicyl alcohol was overaccumulating, we would expect increase in both salicin and isosalicin concentrations (Zenk, 1967; Babst et al., 2010; Chapter 3). There was no detectable increase in the accumulation of isosalicin in UGT78M1-KO lines (Supplemental Figure S5.2).

The connection of UGT78M1 with salicyl benzoate glucoside accumulation provides an opportunity to examine the in planta roles of the coexpressed acyltransferase genes SABT and BEBT. SABT forms salicyl benzoate, which is a precursor for salicyl benzoate glucoside. Further metabolomic analysis targeting benzenoid and dibenzenoid concentration changes as a result of UGT78M1 disruption may provide valuable insight into unidentified chemotypic changes present in the leaves of UGT78M1-KO and UGT71L1/UGT78M1-KO lines.



Genotype	Salicyl benzoate glucoside FC
UGT71L1-KO	0.5
UGT78M1-KO	5.2
UGT71L1/UGT78M1-KO	4.2

Figure 5.4. Salicyl benzoate glucoside concentrations comparisons between gene edited poplars. Concentration differences are presented by average response area fold-change (FC). UGT71L1-KO and UGT78M1-KO concentration fold-change calculated compared to wild-type values. UGT71L1/UGT78M1-KO double knockout concentration fold-change calculated compared to UGT71L1-KO values.

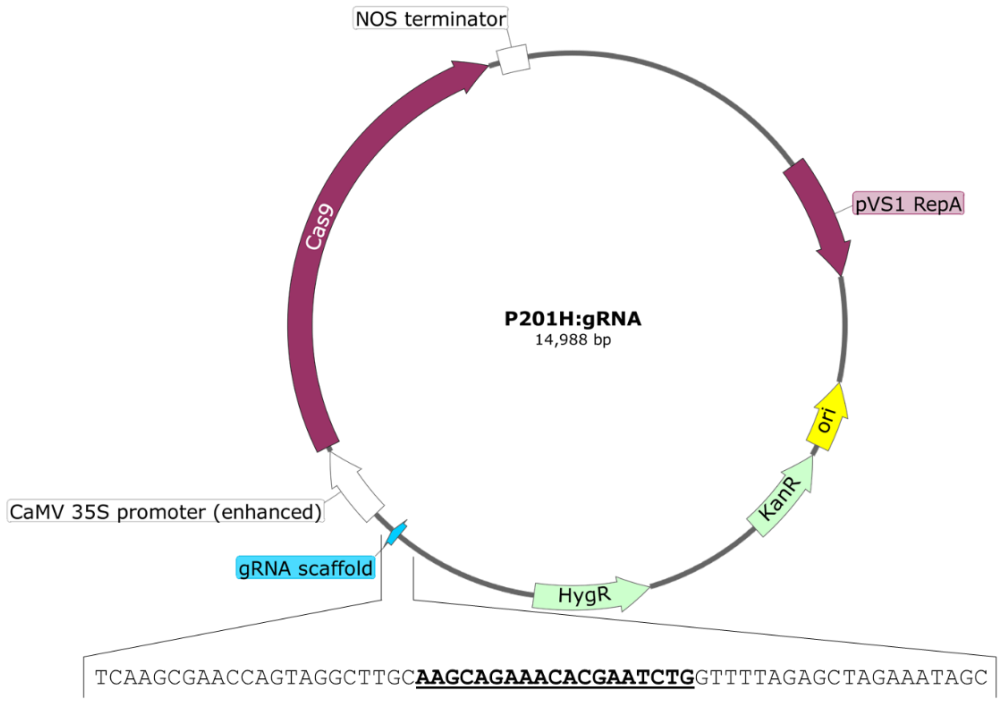
Protein BLAST analysis of the *P. trichocarpa* UGT78M1 (Potri.006G171100) predicts UGT78M1 as a flavonoid and anthocyanin 3-O-glycosyltransferase enzyme, which is functionally inaccurate (Johnson et al., 2008; Veljanovski and Constabel, 2013). The most similar identified protein sequence in *Arabidopsis thaliana* is AtUGT78D2, with a 46% sequence similarity. While AtUGT78D2 is capable of flavonoid-3-

glucosyltransferase activity, there is evidence for the involvement of flavonoid glycosyltransferase activity in hydroxybenzenoid metabolism. The absence of functional AtUGT78D2 in Arabidopsis can increase the concentration of the ubiquinone precursor 4-hydoxybenzoate through the peroxidative cleavage of kaempferol (Soubeyrand et al., 2018). The AtUGT78D2 catalyzed glucosylation is important for the ubiquinone benzenoid supply in Arabidopsis (Soubeyrand et al., 2021). Further exploration of hydroxybenzenoid metabolism in UGT78M1 knockout plants may provide insight into the biological roles of UGT78M1. Our results establish a connection between UGT78M1 and dibenzenoid metabolism, but the mechanism for this connection remains undetermined.

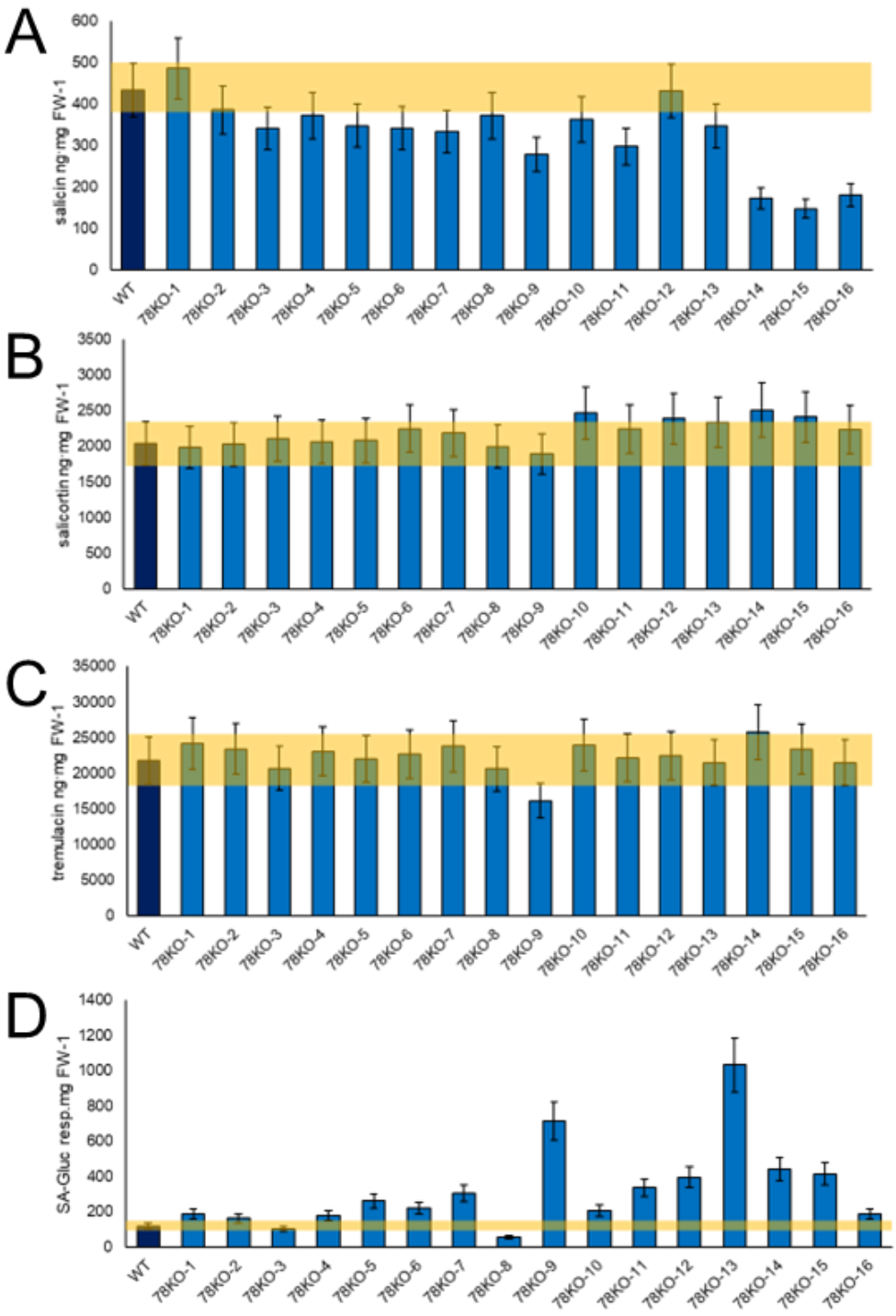
5.3.4 Summary

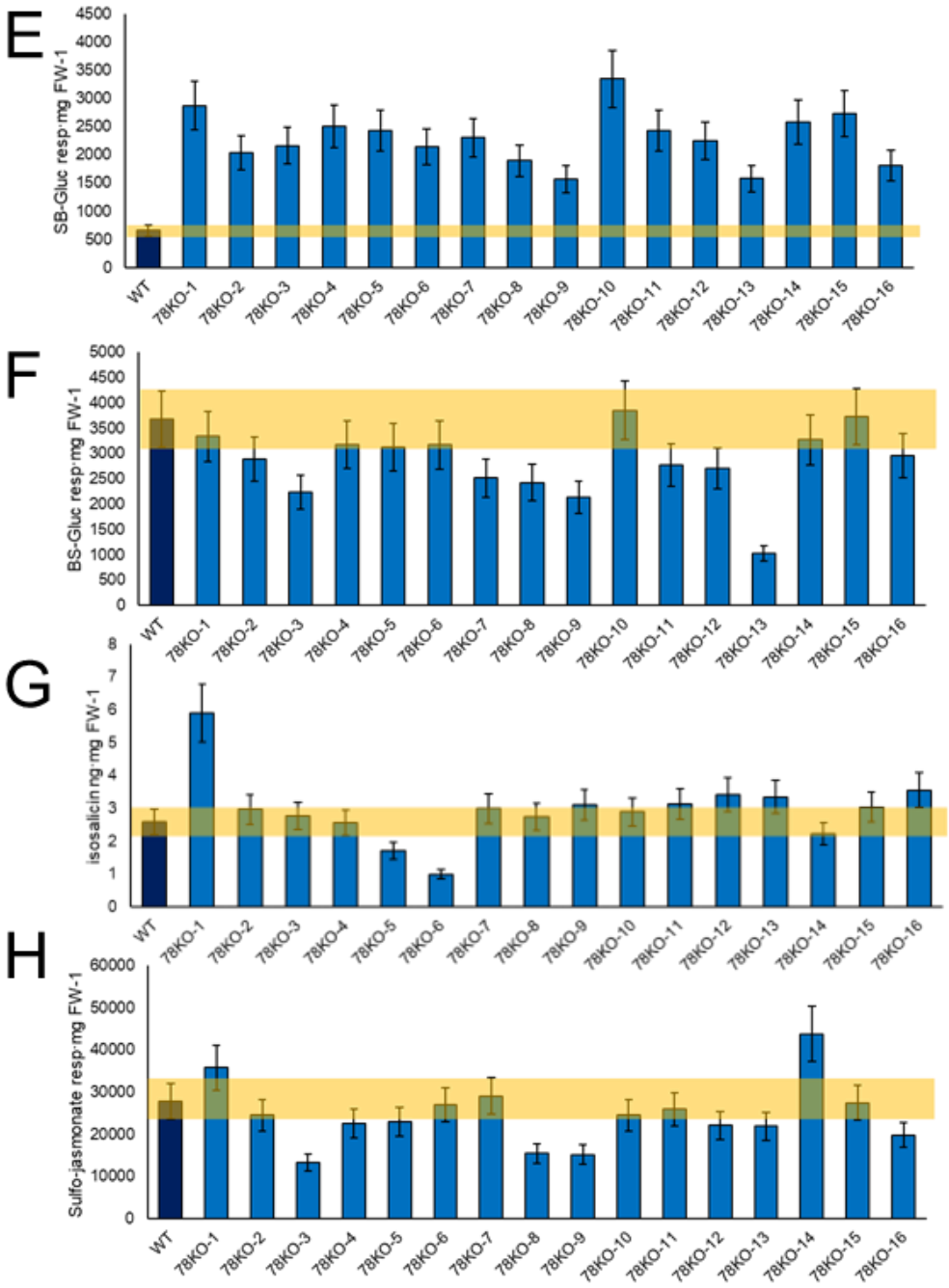
Our work has demonstrated functional use of CRISPR/Cas9 constructs to efficiently transform the UDP-dependent glycosyltransferase UGT78M1 in hybrid poplar with a broad range of insertions and deletions. We demonstrate that hygromycin selection of CRISPR/Cas9 edited plants in a genetic background previously transformed with a p201N vector is a viable way to generate multi-gene knockouts in poplar. We also show that UGT78M1 does not appear to contribute to the biosynthesis of salicortin or tremulacin. Functional UGT78M1 is essential for the normal accumulation of salicyl benzoate glucoside and salicin in one-month old greenhouse grown poplars.

5.5 Supplementary Material

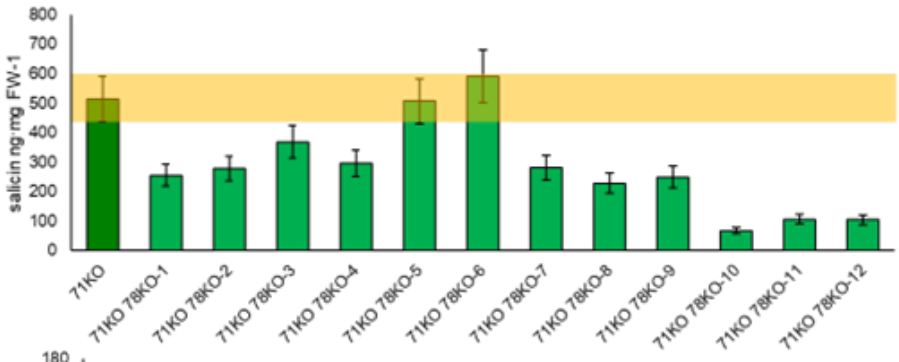


Supplemental Figure S5.1. Vector map of the p201H construct containing the CRISPR/Cas9 gRNA sequence used to disrupt UGT78M1. The gRNA sequence is bolded and underlined, flanked by the construct scaffold. Abbreviations: HygR, hygromycin resistance marker; kanR, kanamycin resistance marker; ori, origin of replication site; CAMV, Cauliflower Mosaic Virus; RepA, replication protein; bp, base pairs; NOS, nopaline synthase. Vector map was prepared using a modified image from SnapGene (Dotmatics, Boston, MA, USA).

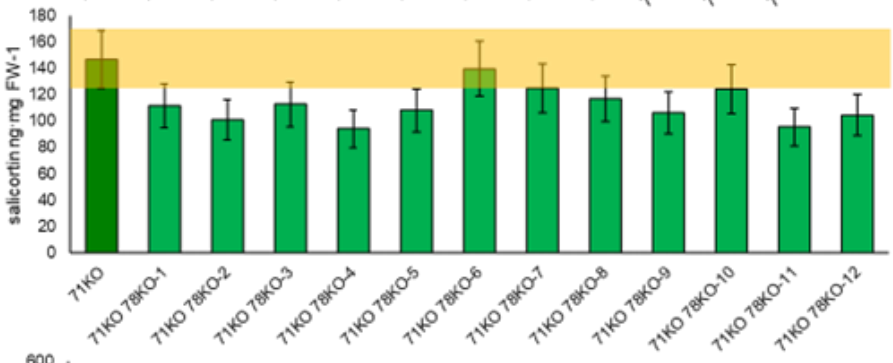

A
salicin ng·mg FW-1
600
500
400
300
200
100
0
WT
78KO-1
78KO-2
78KO-3
78KO-4
78KO-5
78KO-6
78KO-7
78KO-8
78KO-9
78KO-10
78KO-11
78KO-12
78KO-13
78KO-14
78KO-15
78KO-16
B
salicortin ng·mg FW-1
3500
3000
2500
2000
1500
1000
500
0
C
tremulacin ng·mg FW-1
35000
30000
25000
20000
15000
10000
5000
0
D
SA-Gluc resp.mg FW-1
1400
1200
1000
800
600
400
200
0


E
SB-Gluc resp·mg FW-1
4500
4000
3500
3000
2500
2000
1500
1000
500
0
WT
78KO-1
78KO-2
78KO-3
78KO-4
78KO-5
78KO-6
78KO-7
78KO-8
78KO-9
78KO-10
78KO-11
78KO-12
78KO-13
78KO-14
78KO-15
78KO-16
F
BS-Gluc resp·mg FW-1
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
G
isosalicin ng·mg FW-1
8
7
6
5
4
3
2
1
0
H
Sulfo-jasmonate resp·mg FW-1
60000
50000
40000
30000
20000
10000
0

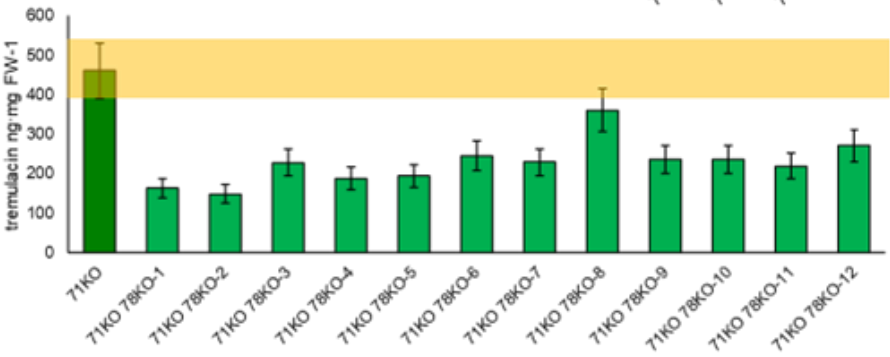
I


salicin ng·mg FW-1
800
700
600
500
400
300
200
100
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12

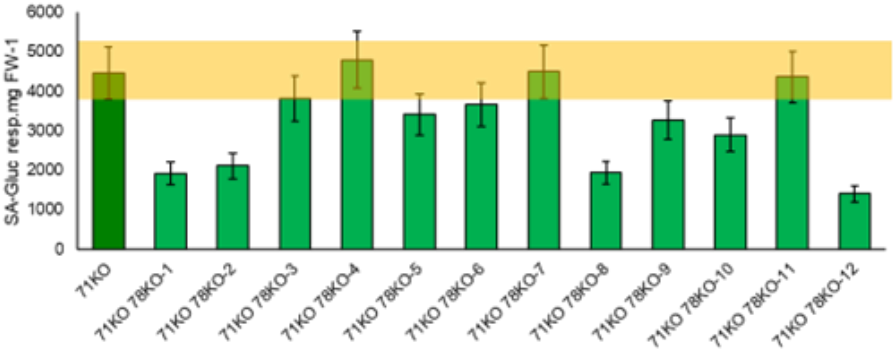
J

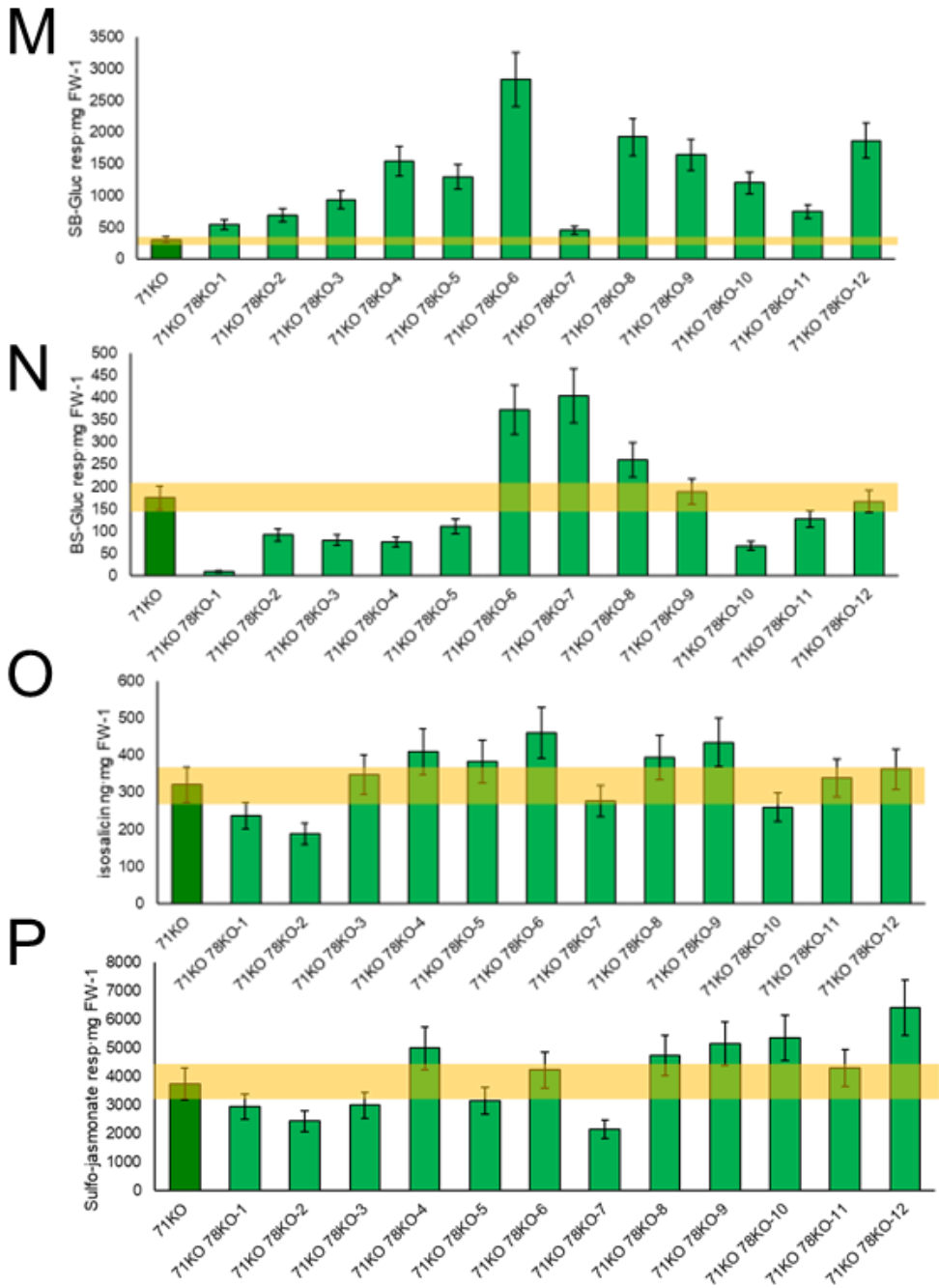

salicortin ng·mg FW-1
180
160
140
120
100
80
60
40
20
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12

K

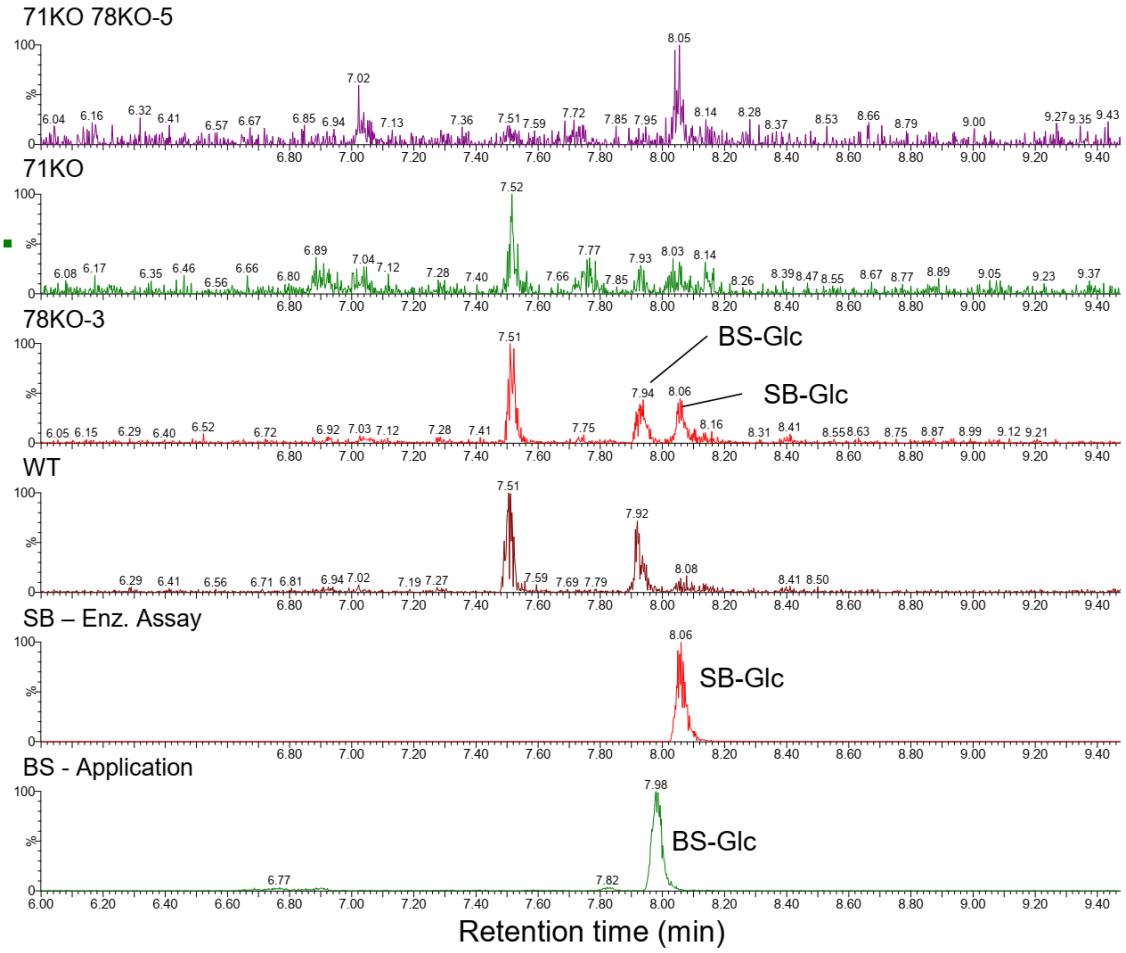

tremulacin ng·mg FW-1
600
500
400
300
200
100
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12

L


SA-Gluc resp.mg FW-1
6000
5000
4000
3000
2000
1000
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12


M
SB-Gluc resp·mg FW-1
3500
3000
2500
2000
1500
1000
500
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12
N
BS-Gluc resp·mg FW-1
500
450
400
350
300
250
200
150
100
50
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12
O
isosalicin ng·mg FW-1
600
500
400
300
200
100
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12
P
Sulfo-jasmonate resp·mg FW-1
8000
7000
6000
5000
4000
3000
2000
1000
0
71KO
71KO 78KO-1
71KO 78KO-2
71KO 78KO-3
71KO 78KO-4
71KO 78KO-5
71KO 78KO-6
71KO 78KO-7
71KO 78KO-8
71KO 78KO-9
71KO 78KO-10
71KO 78KO-11
71KO 78KO-12

Supplemental Figure S5.2. Chemotypic analysis of UGT78M1 knockout and UGT71L1/UGT78M1 knockout lines. Blue graphs are analytical results from UGT78M1 knockout lines, green graphs are analytical results from UGT71L1/UGT78M1 double knockout lines. A; salicin concentration, B; salicortin concentration, C; tremulacin concentration, D; Salicylic acid glucoside response area·FW^{-1}, E; salicyl benzoate glucoside response area·FW^{-1}, F; benzyl salicylate glucoside response area·FW^{-1}, G, isosalicin concentration, H; sulfo-jasmonate response area·FW^{-1}. I; salicin concentration, J; salicortin concentration, K; tremulacin concentration, L; Salicylic acid glucoside response area·FW^{-1}, M; salicyl benzoate glucoside response area·FW^{-1}, N; benzyl salicylate glucoside response area·FW^{-1}, O, isosalicin concentration, P; sulfo-jasmonate response area·FW^{-1}. WT bars represent the mean of 4 biological replicates, 71KO lines represent 3 biological replicates. Each independently transformed line was analyzed in duplicate. The yellow shaded bars represents a 15% relative standard deviation from the mean of the genetic background line (WT or 71KO) and was used to visually assess chemotypic variation. Error bars represent ± 15% mean values.



Supplemental Figure S5.3. Extracted ion chromatograms (EIC) for the identification of salicyl benzoate glucoside and benzyl salicylate glucoside. LC/MS separation and detection of targeted analytes at 435 m/z. Chromatograms in descending order: 71KO 78KO-5; representative of UGT71L1/UGT78M1 knockout lines, 71KO; representative EIC of UGT71L1-knockout line, 78KO-3, UGT78M1 knockout line, WT; wild-type line, SB – Enz. Assay; LC/MS analysis of salicyl benzoate enzyme assay with recombinant UGT71L1 as described by Gordon et al. (2022), BS – Application, analysis of exogenous benzyl salicylate administration to wild-type poplar leaf tissue as previously described (Chapter 3). Abbreviations: 71KO, UGT71L1 knockout; 78KO, UGT78M1 knockout; WT, wild-type; SB, salicyl benzoate; BS, benzyl salicylate; Glc, glucoside; Enz., enzyme.

Supplemental Table T5.1. Targeted LC/MS analyte mass and retention times

Compound[1](adduct)	m/z	ret. time (min)
Salicin (M + FA)	331	2.90
Isosalicin (M-H)	285	3.28
Salicortin (M+FA)	469	5.70
SA-Glc (M-H)	299	3.31
SB-Glc (M+FA)	435	8.07
BS-Glc (M+FA)	435	7.98
Tremulacin (M-H)	527	8.40
Sulfo-jasmonate[2] (M-H)	305	4.45

[1]Abbreviations: M, molecular ion; FA, formate; ret., retention; SA-Glc, salicylate glucoside; SB-Glc, salicyl benzoate glucoside; BS-Glc, benzyl salicylate glucoside.
[2]Putative identification based on m/z values and response area

Chapter 6 – Overall Conclusions

6.1 Summarized Findings

As the understanding of plant metabolism and biochemistry expands and evolves, if there is a central theme is coalescing it is “connection”. With the advent of reasonably accessible multi-omic data, and the rapid rise in studies using non-model species, unexpected connections in biology and plant biochemistry are regularly discovered. Salicinoid biosynthesis, and the studies contained within this dissertation illustrate one such example. My work builds on this trend and shows that salicinoids in poplar are more than defense metabolites with a single function. Salicinoids and their biosynthesis have an impact on the growth, development, and defense of poplar trees. Genome edited transgenic trees, metabolomic analysis, RNA-Sequencing, and conventional biochemical assays offer whole plant resolution of the breadth of salicinoid metabolic connections.

My dissertation recognizes UGT71L1 as a fundamental gene for the biosynthesis of salicinoids. I have demonstrated the importance of functional UGT71L1 in poplar, as disruption impacts plant growth and development. My work demonstrates a metabolic connection between disrupted salicinoid biosynthesis and the phytohormone salicylic acid. Perturbations in salicylic acid resulting from salicinoid disruption led to discovery of the connection of salicinoid biosynthesis and growth in whole plants. My work also highlights the constitutive production of salicinoids in poplar despite carbon limitations. During carbon starvation poplar trees continued to produce salicinoids. This work demonstrated the inaccessibility of the glucose moiety on the salicinoid molecules for energy metabolism, thus salicinoids appear to be a metabolic end product with minimal capacity for remobilization (Chapter 4). Furthermore, this dissertation builds upon the work of Fellenberg et al. (2020) and demonstrates that the UDP-dependent glycosyltransferase enzyme UGT78M1 is not required for the core biosynthesis of salicinoids. However, UGT78M1 does impact salicin and dibenzenoid accumulation (Chapter 5).

I have used CRISPR/Cas9 genome editing with *Agrobacterium* mediated transformation to focus on the in planta identification of gene function. It was through the

generation of stable transformant plant in the target species, *P. tremula x P. alba*, that I was able to explore the connections in salicinoid biosynthesis that were the result of interrupted salicinoid biosynthesis (Gordon et al., 2022). To assess connections and the biosynthetic roles of our target genes we utilized transcriptomic, targeted metabolomic, and non-targeted metabolomic approaches for the analysis of whole plants. Administration of stable isotopically labelled substrates and the UPLC-MS analytical capacity analysis was leveraged to grow our understanding of the connections present within the salicinoid metabolic network (Chapter 3). I provide the first direct evidence of benzyl benzoate as an entry point and intermediate in salicinoid biosynthesis.

6.2 Impact of Work

My work has generated significant advances on our understanding of the relationship of salicinoids and poplar tree growth. This dissertation highlights the advantages of gene analysis in target species. The use of only heterologous expression systems such as *N. benthamiana* can limit the discovery of biologically relevant connections (Jacobowitz and Weng., 2020). In a broader perspective, my dissertation provides insight into possible biosynthetic routes of phenylpropanoid derived salicylic acid (Gordon et al., 2022). Notably, I have demonstrated how perturbations to glucosylation of phenolic derived dibenzenoid molecules connects to salicylic acid overaccumulation which, in turn, can drive larger developmental changes through agonistic connections to jasmonic acid. My work on phytohormone connections of phenylpropanoid glucosylation and salicylic acid accumulation may have broad impacts in plant biology. Interruption of a related chalcone UDP-glycosyltransferase in apple (*Malus domestica*) perturbs salicylic acid homeostasis and contributes to fungal pest resistance (Dare et al., 2017; Zhou et al., 2019). As a phytohormone, phenylpropanoid derived salicylic acid contributes to systemic acquired resistance in plants and Arabidopsis has a limited capacity for understanding fundamental molecular interactions in perennial woody plants (Ullah et al., 2019; Torrens-Spence et al., 2019). The work presented in Chapter 2 demonstrates discovery of a small molecule trichotomy within

salicinoid biosynthetic intermediates involving specialized metabolites, hormones, and primary metabolites (Erb and Kliebenstein., 2020; Gordon et al., 2022).

Chapters 2 and 3 demonstrate the networked connections of benzenoids and dibenzenoids in poplar. Furthermore, my work proposes the presence of a metabolon or enzyme complex for the conversion of benzyl benzoate to salicortin. In a sorghum metabolon, a UGT is essential for the assembly of an enzyme complex (Laursen et al., 2016). UGT71L1 may be an integral component of a salicinoid biosynthetic metabolons (Vogt, 2010). Additionally, my work provides evidence that salicyl benzoate or benzyl salicylate glucosides are not intermediates in salicortin biosynthesis. The salicyl benzoate aglycone is not an entry point to salicortin biosynthesis; however, salicyl benzoate may be an intermediate that is formed during salicinoid biosynthesis by the putative metabolon (Chapter 3). Differential incorporation of benzenoid precursors into salicinoids and glucosylated phenolic metabolites are consistent with the networked conversions of benzenoid molecules in poplar (Orlova et al., 2006; Widhalm and Dudareva, 2015). The benzenoid connections I have identified through labelled molecule tracking demonstrate the flexibility of benzenoid metabolism as benzoate can contribute to both rings of salicortin, salicin, isosalicin, and dibenzenoid glucosides (Chapter 3). One open possibility in dibenzenoid and salicinoid biosynthesis remains untested, and that is the possible contribution of a molecule salicoyl-CoA as an intermediate. While salicoyl-CoA has not yet been observed or synthesized, if created, it could be screened with recombinant SABT and BEBT as a possible substrate.

The analysis of resprouted poplar roots under carbon limiting conditions demonstrated differences in salicinoid profiles between above and below ground tissues. Continued salicinoid production during carbon starvation demonstrated constitutive salicinoid biosynthesis in whole trees. Significantly increased salicin accumulation during resprouting in UGT71L1-KO poplars further distinguishes salicin biosynthesis from salicortin and tremulacin biosynthesis. This presents an excellent targeted for differential gene expression analysis identify additional salicin biosynthetic genes. Poplars can undergo carbon starvation despite having a small concentration of root starch that remains. The glucose moiety in the salicinoids salicortin and tremulacin appear inaccessible, and salicinoid biosynthesis competes directly with primary metabolism

during carbon starvation. During poplar resprouting in carbon limited conditions salicinoid deficient UGT71L1-KO lines survive longer per gram of available biomass.

Interruption of the UDP-glycosyltransferase gene UGT78M1 shows that UGT78M1 has minimal contributions to core salicinoid biosynthesis in poplar. This data is consistent with the absence of UGT78M1 in willow, despite the production of salicin, salicortin, and tremulacin (Kulasekaran et al., 2020). Furthermore, disruption of UGT78M1 results in an overaccumulation of salicyl benzoate glucoside. I have expanded the known substrates of two UDP-dependent glycosyltransferase enzymes UGT78M1 and UGT71L1. Veljanovski and Constabel (2013) found no flavonoid specific activity with recombinant UGT78M1, but Fellenberg et al. (2020) demonstrated a strong preference for the glucosylation of salicyl benzoate. My work further expanded the known substrates of UGT78M1 and UGT71L1 to include salicyl salicylate and benzyl salicylate. Preliminary data presented in Appendix 6 shows an expanded substrate repertoire for UGT78M1, which includes gentisyl benzoate.

In summary, the small molecule trichotomy in plant specialized metabolism is an excellent theme for this dissertation (Erb and Kliebenstein, 2020). Work in Chapter 2 shows the connection of salicinoid specialized metabolites to the phytohormones salicylic and jasmonic acid. In Chapter 3, I have expanded our understanding of the metabolic connections that appear for benzenoid molecules. Chapter 4 shows the direct competition and interaction of salicinoid biosynthesis and primary metabolism. Finally, Chapter 5 highlights how additional enzymatic connections are present but remain unelucidated.

6.3 Future Research

An under-investigated component of salicinoid biosynthesis is the subcellular and tissue specific localization of the biosynthetic enzymes. My work demonstrates that poplar root and shoot tissue are likely capable of in situ salicinoid biosynthesis. Matrix assisted laser desorption/ionization has been used to demonstrate that the salicortin and tremulacin are concentrated in vascular tissue of leaves (Costa et al., 2013). Further refinement to better understand biosynthesis in root tips, phloem, vascular, and mesophilic localization is needed. In Chapter 4 I demonstrate the difference in salicinoid

profiles between above and belowground tissues. Tremulacin is the most abundant salicinoid in above ground tissue, whereas in roots salicortin is most abundant. Understanding differences in expression patterns of UGT71L1 may provide insight into the precise localization of biosynthesis. To aid in this, I have generated a β-glucuronidase (GUS) expression construct under the control of the native UGT71L1 promoter in a modified PMDC32 expression vector (Appendix 3). Generating poplar transformants using the provided expression construct and conducting the subsequent GUS assays may provide information into the tissue specific localization of salicinoid biosynthesis. Gene coexpression analysis with SABT generated a list of genes that may contribute to salicinoid biosynthesis or metabolism. The list contains two uncharacterized proteins annotated as transporters. Investigation into the substrates of transporter proteins Potri.011G002500 and Potri.014G136600 may provide insight into the subcellular localization of salicinoid pools.

The salicinoid biosynthetic pathway is hypothesized to have biosynthetic enzymes spatially organized into a complex or metabolon. Benzyl benzoate contributes to salicortin biosynthesis; however, salicyl benzoate (SB) is not an entry point. The conversion of benzyl benzoate to salicortin may require the formation of salicyl benzoate; however, if SB cannot be introduced to the pathway SB as an intermediate may be held in an enzyme complex during salicortin biosynthesis. The presence of a metabolon is technically difficult to demonstrate due to labile and non-covalent protein-protein interactions. However, a tagged protein pull-down may provide an opportunity to enrich and proteomically sequence associated proteins (Mucha et al., 2019). Heterologous expression of all candidate salicinoid biosynthetic genes in *N. benthamiana*, with a fluorescently tagged UGT71L1 protein, followed by a protein extraction and pull-down assay may demonstrate these hypothesized associations (Huang et al., 2022). Understanding the complete biosynthesis from salicortin requires identification of further enzymatic steps. Several oxidation reactions are likely required for salicortin biosynthesis. Multiple CYP450s and an enone oxidizing reaction catalyzed by a Vein Patterning 1 (VEP1) gene identified from coexpression with SABT (Fellenberg et al., 2020) are under active investigation for their role in salicortin biosynthesis. Furthermore, UGT71L1 coexpression has presented additional genes that may be related to the

modification of core salicinoid structures. One potential gene is a serine carboxy peptidase like acyl transferase (SCPL) (Potri.001G312800). The SCPL gene has been disrupted using CRISPR/Cas9 genome editing. Analysis of lines with a disrupted SCPL gene demonstrate lack of cinnamoyl and acyl salicortin derivatives. Thus, indicating SCPL contribution to the transfer of a cinnamoyl group required for the formation of cinnamoyl salicortin (Keefover et al., 2014). Preliminary analysis and identification of cinnamoyl salicortin isomers is presented in Appendix 4.

The widespread perturbations in shikimate and amino acid metabolism observed in UGT71L1-KO lines were reminiscent of a similar but less pronounced phenotype in the MYB165 overexpressing (OE) plants (Ma et al., 2018; Gordon et al., 2022). The phenylpropanoid pathway is metabolically downstream to the shikimic acid pathway. The shikimic acid pathway is responsible for the formation of phenylalanine, tyrosine and Tryptophan. The phytohormone auxin and other indole derived regulators are also downstream products of the shikimic acid pathway (Maeda and Dudareva, 2012).
To conduct a cursory exploration into indole derived metabolic products we conducted an analysis of melatonin in MYB165 OE lines. It was found that MYB165 overexpressing lines accumulate nearly double the melatonin as wild-type *P. tremula x P. tremuloides*. The biological relevance of melatonin overaccumulation could be further explored and preliminary data is presented in Appendix 5.

Salicinoids are well established as defences against herbivores in poplar (Boeckler et al., 2011; Lindroth and St. Clair, 2013). However, the constitutive and whole plant production of salicinoids raises interesting questions regarding their evolutionary history. What pressures have led to such constitutive and developmentally important production of salicinoid specialized metabolites? Furthermore, tissue specific salicinoid profiles, such as the differences between roots and shoots observed in Chapter 2 implies a possible biological relevance of tissue specific distributions. For example, salicinoids providing resistance to leaf eating herbivory is well established (Boeckler et al., 2011). The roles of salicinoids as chemical defences against root herbivores is not known.

6.4 Concluding Remarks

As scientists we work on the precipice of mystery, often known only enough to generate reasonable hypotheses. With a single iteration of "why" or "how" we reach into the unknown. Excitingly, in the biological sciences we are beginning to have the tools necessary to answer almost any questions that arise. For the tenacious, answers – and the new questions that inevitably follow – are increasingly accessible. Through perseverance and good science many more connections, insights, and biochemical pathways will continue to be understood at an increasing rate.

This dissertation contributes to a growing understanding of plant biochemistry and biology. With modern accessible and available genomic, transcriptomic, and metabolomic techniques the study of plant chemistry is expanding from the biochemical reductionist approaches. No more are we limited to using purely recombinant proteins, model species, and heterologous expression. I am thankful to benefit from these growing advances and contribute in my own minor way. The exploration beyond conventional reductionist approaches led to the identification of intricate and unexpected connections in throughout salicinoid biosynthesis. Besides providing a name for my doctoral dissertation, my work reaffirms the need to continue to work on biochemical pathways in their respective plant hosts. It is in plant biochemical research, focused beyond reductionism, that scientists will continue to find more unexpected connections.

Scio me nihil scire

Bibliography

Ackerman D, Breen A. (2016) Infrastructure development accelerates range expansion of trembling aspen (*Populus tremuloides*, Salicaceae) into the arctic. Arctic. **69**: 130–136

Afendi FM, Okada T, Yamazaki M, Hirai-Morita A, Nakamura Y, Nakamura K, Ikeda S, Takahashi H, Altaf-Ul-Amin M, Darusman LK, Saito K, Kanaya S. (2012) KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. Plant and Cell Physiology. **53**: e1

Babst BA, Harding SA, Tsai CJ. (2010) Biosynthesis of phenolic glycosides from phenylpropanoid and benzenoid precursors in *Populus*. Journal of Chemical Ecology. **36**: 286–297

Babst BA, Chen HY, Wang HQ, Payyavula RS, Thomas TP, Harding SA, Tsai CJ. (2014) Stress-responsive hydroxycinnamate glycosyltransferase modulates phenylpropanoid metabolism in *Populus*. Journal of Experimental Botany. **65**: 4191–4200

Bai M, Yuan J, Kuang H, Gong P, Li S, Zhang Z, Liu B, Sun J, Yang M, Yang L, Wang D, Song S, Guan Y. (2020) Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean. Plant Biotechnology Journal. **18**: 721–731

Barker HL, Holeski LM, Lindroth RL. (2019) Independent and interactive effects of plant genotype and environment on plant traits and insect herbivore performance: A meta-analysis with Salicaceae. Functional Ecology. **33**: 422–435

Basey JM, Jenkins SH, Busher PE. (1988) Optimal central-place foraging by beavers: Tree-size selection in relation to defensive chemicals of quaking aspen. Oecologia. **76**: 278–282

Bewg WP, Harding SA, Engle NL, Vaidya BN, Zhou R, Reeves J, Horn TH, Joshee N, Jenkins JW, Shu S, Barry KW, Yoshinaga Y, Grimwood J, Schmitz RJ, Schmutz J, Tschaplinksi TJ, Tsai CJ. (2022) Multiplex knockout of trichome-regulating MYB duplicates in hybrid poplar using a single gRNA. Plant Physiology. **189**: 516–526

Bjarnholt N, Neilson EHJ, Crocoll C, Jørgensen K, Motawia MS, Olsen CE, Dixon DP, Edwards R, Møller BL. (2018) Glutathione transferases catalyze recycling of auto-toxic cyanogenic glucosides in sorghum. The Plant Journal. **94**: 1109–1125

Brinkman EK, Chen T, Amendola M, Van Steensel B. (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research. **42**: 1–8

Brunner AM, Busov VB, Strauss SH (2004) Poplar genome sequence: functional genomics in an ecologically dominant plant species. Trends in Plant Science. **9**: 49-56

Bryant JP, Clausen TP, Swihart RK, Landhäusser SM, Stevens MT, Hawkins CD, Carrière S, Kirilenko AP, Veitch AM, Popko RA, Cleland DT, Williams JH, Jakubas WJ, Carlson MR, Bodony KL, Cebrian M, Paragi TF, Picone PM, Moore JE, Packee EC, Malone T. (2009) Fire drives transcontinental variation in tree birch defense against browsing by snowshoe hares. The American Naturalist. **174**: 13-23

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

Boeckler GA, Gershenzon J, Unsicker SB (2013) Gypsy moth caterpillar feeding has only a marginal impact on phenolic compounds in old-growth black poplar. Journal of Chemical Ecology. **39**: 1301–1312

Boeckler GA, Paetz C, Feibicke P, Gershenzon J, Unsicker SB. (2016) Metabolism of poplar salicinoids by the generalist herbivore *Lymantria dispar* (lepidoptera). Insect Biochemistry and Molecular Biology. **78**: 39–49

Chedgy RJ (2015) The role of BAHD acyltransferases in poplar. PhD dissertation, University of Victoria.

Chedgy RJ, Köllner TG, Constabel CP. (2015) Functional characterization of two acyltransferases from *Populus trichocarpa* capable of synthesizing benzyl benzoate and salicyl benzoate, potential intermediates in salicinoid phenolic glycoside biosynthesis. Phytochemistry. **113**: 149–159

Choi SH, Ahn WS, Jie EY, Cho HS, Kim SW. (2022). Development of late-bolting plants by CRISPR/Cas9-mediated genome editing from mesophyll protoplasts of lettuce. Plant Cell Reports. **41**: 1627–1630

Chong J, Wishart D, Xia J (2019) Using metaboanalyst 4.0 for comprehensive and integrative metabolomics data analysis. Current Protocols in Bioinformatics. **68**: e86

Christenhusz MJM, Byng JW. (2016) The number of known plant species in the world and its annual increase. Phytotaxa. **261**: 201–217

Clifford MN, Jaganath IB, Ludwig IA, Crozier A. (2017) Chlorogenic acids and the acyl-quinic acids: discovery, biosynthesis, bioavailability and bioactivity. Natural Products Report. **34**: 1391–1421

Cole CT, Morrow CJ, Barker HL, Rubert-Nason KF, Riehl JFL, Köllner TG, Lackus ND, Lindroth RL. (2021) Growing up aspen: ontogeny and trade-offs shape growth, defence and reproduction in a foundation species. Annual reviews in Botany. **127:** 505-517

Cope OL, Keefover-Ring K, Kruger EL, Lindroth RL. (2021) Growth-defense trade-offs shape population genetic composition in an iconic forest tree species. Proceedings of the National Academy of Science USA. **118**: e2103162118

Cope OL, Lindroth RL, Helm A, Keefover-Ring K, Kruger EL. (2021) Trait plasticity and trade-offs shape intra-specific variation in competitive response in a foundation tree species. New Phytologist. **230**: 710-719

Coleman HD, Park JY, Nair R, Chapple C, Mansfield SD. (2008) RNAi-mediated suppression of *p*-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. Proceedings of the National Academy of Science USA. **105:** 4501-4506

Cordell GA. (2000) Biodiversity and drug discovery – a symbiotic relationship. Phytochemistry. **55**: 463–480

Costa MA, Marques JV, Dalisay DS, Herman B, Bedgar DL, Davin LB, Lewis NG. (2013) Transgenic hybrid poplar for sustainable and scalable production of the commodity/specialty chemical, 2-phenylethanol. PloS One. **8**: e83169

Curtis MD, Grossniklaus U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiology. **133**: 462-469

Dare AP, Yauk YK, Tomes S, McGhie TK, Rebstock RS, Cooney JM, Atkinson RG (2017) Silencing a phloretin-specific glycosyltransferase perturbs both general phenylpropanoid biosynthesis and plant development. The Plant Journal. **91**: 237–250

Dastmalchi M, Bernards MA, Dhaubhadel S. (2016) Twin anchors of the soybean isoflavonoid metabolon: Evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. The Plant Journal. **85**: 689–706

Datia RSS, Hammerlindl JK, Panchuk B, Pelcher LE, Keller W. (1992) Modified binary plant transformation vectors with the wild-type gene encoding NPTII. Gene. **122**: 383–384

D'Auria JC, Chen F, Pichersky E. (2002) Characterization of an acyltransferase capable of synthesizing benzylbenzoate and other volatile esters in flowers and damaged leaves of *Clarkia breweri*. Plant Physiology. **130**: 466–476

Dietze MC, Sala A, Carbone MS, Czimczik CI, Mantooth JA, Richardson AD, Vargas R. (2014) Nonstructural carbon in woody plants. Annual Reviews in Plant Biology. **65**: 667-687

Dean JV, Delaney SP. (2008) Metabolism of salicylic acid in wild-type, *ugt74f1* and *ugt74f2* glucosyltransferase mutants of *Arabidopsis thaliana*. Physiologia Plantarum. **132**: 417–425

Desborough MJR, Keeling DM. (2017) The aspirin story – from willow to wonder drug. British Journal of Haematology. **177**: 674–683

Do PT, Nguyen CX, Bui HT, Tran LTN, Stacey G, Gillman JD, Zhang ZJ, Stacey MG. (2019) Demonstration of highly efficient dual gRNA CRISPR/Cas9 editing of the homeologous GmFAD2-1A and GmFAD2-1B genes to yield a high oleic, low linoleic and α-linolenic acid phenotype in soybean. BMC Plant Biology. **19**: 1–14

Don D, Poovaiah C, Phillips L, Geddes B, Reeves C, Sorieul M, Thorlby G. (2021) Genome editing with CRISPR/Cas9 in *Pinus radiata* (D. Don). BMC Plant Biology. **21**: 363

Ding P, Ding Y. (2020) Stories of salicylic acid: a plant defense hormone. Trends in Plant Science. **25**: 549-656

Ding Y, Sun T, Ao K, Peng Y, Zhang Y, Li X, Zhang Y (2018) Opposite roles of salicylic acid receptors npr1 and npr3/npr4 in transcriptional regulation of plant immunity. Cell. **173**: 1454-1467

Donaldson JR, Stevens MT, Barnhill HR, Lindroth RL. (2006) Age-related shifts in leaf chemistry of clonal aspen (*Populus tremuloides*). Journal of Chemical Ecology. **32**: 1415–1429

Duran-Medina Y, Ruiz-Cortes BE, Guerrero-Largo H, Marsch-Martinez N. (2021) Specialized metabolism and develoment: an unexpected friendship. Current Opinions in Plant Biology. **64**: 102142

Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik AV, Meusel M, Dorrestein P, Rousu J, Böcker S. (2019) SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nature Methods. **16**: 299–302

Dührkop K, Shen H, Meusel M, Rousu J, Böcker S. (2015) Searching molecular structure databases with tandem mass spectra using CSI:FingerID. Proceedings of the National Academy of Science USA. **112**: 12580–12585

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

Erb M, Kliebenstein DJ. (2020) Plant Secondary Metabolites as Defenses, Regulators, and Primary Metabolites: The Blurred Functional Trichotomy. Plant Physiology. **184**: 39–52

Erland LAE, Giebelhaus RT, Victor JMR, Murch SJ, Saxena PK. (2020) The morphoregulatory role of thidiazuron: Metabolomics-guided hypothesis generation for mechanisms of activity. Biomolecules. **10**: 1253

Erland LAE, Turi CE, Saxena PK, Murch SJ. (2020) Metabolomics and hormonomics to crack the code of filbert growth. Metabolomics. **16**: 62

Fabisch T, Gershenzon J, Unsicker S. (2019) Specificity of herbivore defense responses in a woody plant, black poplar (*Populus nigra*). Journal of Chemical Ecology. **45**: 162-177

Feistel F, Paetz C, Lorenz S, Schneider B. (2015) The absolute configuration of salicortin, HCH-salicortin and tremulacin from *Populus trichocarpa x deltoides* Beaupré. Molecules. **20**: 5566–5573

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

Fernández-Milmanda GL, Crocco CD, Reichelt M, Mazza CA, Köllner TG, Zhang T, Cargnel MD, Lichy MZ, Fiorucci AS, Fankhauser C, Koo AJ, Austin AT, Gershenzon J, Ballaré CL (2020) A light-dependent molecular link between competition cues and defence responses in plants. Nature Plants. **6**: 223–230

Feys BJF, Benedetti CE, Penfold CN, Turner JG (1994) Arabidopsis mutants selected for resistance to the phytotoxin coronatine are male sterile, insensitive to methyl jasmonate, and resistant to a bacterial pathogen. The Plant Cell **6**: 751–759

Fischer S, Hanf S, Frosch T, Gleixner G, Popp J, Trumbore S, Hartmann H. (2015) *Pinus sylvestris* switches respiration substrates under shading but not during drought. New Phytologist. **207**: 542-550

Gallego-Giraldo L, Escamilla-Trevino L, Jackson LA, Dixon RA. (2011) Salicylic acid mediates the reduced growth of lignin down-regulated plants. Proceedings of the National Academy of Science USA. **108**: 20814–20819

Gaude E, Chignola F, Spiliotopoulos D, Mari S, Spitaleri A, Ghitti M. (2012) Muma: Metabolomics Univariate and Multivariate Analysis version 1.4 from CRAN https://rdrr.io/cran/muma/

Gibson DG, Young L, Chuang R, Venter JC, Hutchinson CA, Smith HO. (2010) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods. **6:** 343-345

Goetz M, Rabinovich M, Smith HM. (2021) The role of auxin and sugar signaling in dominance inhibition of inflorescence growth by fruit load. Plant Physiology. **187**: 1189-1202

Gordon H, Fellenberg C, Lackus ND, Archinuk F, Sproule A, Nakamura Y, Köllner TG, Gershenzon J, Overy DP, Constabel CP. (2022) CRISPR/Cas9 disruption of UGT71L1 in poplar connects salicinoid and salicylic acid metabolism and alters growth and morphology. The Plant Cell. **8:** 2925-2947

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

Günther J, Lackus ND, Schmidt A, Huber M, Stodtler HJ, Reichelt M, Gershenzon J, Köllner TG. (2019) Separate pathways contribute to the herbivore-induced formation of 2-phenylethanol in poplar. Plant Physiology. **180:** 767-782

Guo Q, Major IT, Howe GA. (2018) Resolution of growth-defense conflict: mechanistic insights from jasmonate signaling. Current Opinions in Plant Biology. **44:** 72-81

Han R, Meilan MC, Strauss SH. (2000) An *Agrobacterium tumefaciens* transformation protocol effective on a variety of cottonwood hybrids (genus *Populus*). Plant Cell Reports. **19**: 315-320

Hartmann H, Adams HD, Hammond WM, Hoch G, Landhäusser SM, Wiley E, Zaehle S. (2018) Identifying differences in carbohydrate dynamics of seedlings and mature trees to improve carbon allocation in models for trees and forests. Environmental and Experimental Botany. **152**: 7–18

Hartmann H, Trumbore S. (2016) Understanding the roles of nonstructural carbohydrates in forest trees – from what we can measure to what we want to know. New Phytologist. **211**: 386-403

Heldt HW, Piechulla B. (2021) *Plant Biochemistry* (5th ed.) London, UK: Elsevier

Howe GA, Major IT, Koo AJ. (2018) Modularity in jasmonate signaling for multistress resilience. Annual Reviews in Plant Biology. **69**: 387-415

Huang H, Liu B, Liu L, Song S. (2017) Jasmonate action in plant growth and development. Journal of Experimental Botany. **68**: 1349–1359

Huang XQ, Li R, Fu J, Dudareva N. (2022) A peroxisomal heterodimeric enzyme is involved in benzaldehyde synthesis in plants. Nature Communications. **13**: 1352

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

Innes R. (2018) The positives and negatives of NPR: a unifying model for salicylic acid signaling in plants. Cell. **173**: 1314–1315

Irmisch S, Jiang Y, Chen F, Gershenzon J, Köllner TG. (2014) Terpene synthases and their contribution to herbivore-induced volatile emission in western balsam poplar (*Populus trichocarpa*). BMC Plant Biology. **14**: 270

Jacobs TB, LaFayette PR, Schmitz RJ, Parrott WA. (2015) Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnology. **15**: 16

Jacobowitz JR, Weng JK. (2020) Exploring uncharted territories of plant specialized metabolism in the postgenomic era. Annual Reviews in Plant Biology. **71**: 631–658

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

Jamieson MA, Schwartzberg EG, Raffa KF, Reich PB, Lindroth RL. (2015) Experimental climate warming alters aspen and birch phytochemistry and performance traits for an outbreak insect herbivore. Global Change Biology. **21**: 2698–2710

Jander G, Norris SR, Joshi V, Fraga M, Rugg A, Yu S, Li L, Last RL. (2004) Application of a high-throughput HPLC-MS/MS assay to *Arabidopsis* mutant screening; evidence that threonine aldolase plays a role in seed nutritional quality. The Plant Journal. **39**: 465-475

Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. (2008) NCBI BLAST: a better web interface. Nucleic Acids Research. **36**: W5–W9

Jørgensen K, Bak S, Busk PK, Sørensen C, Olsen CE, Puonti-Kaerlas J, Møller BL. (2005) Cassava plants with a depleted cyanogenic glucoside content in leaves and tubers. Distribution of cyanogenic glucosides, their site of synthesis and transport, and blockage of the biosynthesis by RNA interference technology. Plant Physiology. **139**: 363–374

Julkunen-Tiitto R, Sorsa S. (2001) Testing the effects of drying methods on willow flavonoids, tannins, and salicylates. Journal of Chemical Ecology. **27**: 779–789

Kajikawa M, Sierro N, Kawaguchi H, Bakaher N, Ivanov NV, Hashimoto T, Shoji T. (2017) Genomic insights into the evolution of the nicotine biosynthesis pathway in tobacco. Plant Physiology. **174**: 999–1011

Kamatham S, Pallu R, Pasupulati AK, Singh SS, Gudipalli P. (2017) Benzoylsalicylic acid derivatives as defense activators in tobacco and *Arabidopsis*. Phytochemistry. **143**: 160–169

Katz E, Bagchi R, Jeschke V, Rasmussen ARM, Hopper A, Burow M, Estelle M, Kliebenstein DJ. (2020) Diverse allyl glucosinolate catabolites independently influence root growth and development. Plant Physiology. **183:** 1376-1390

Katz E, Nisani S, Yadav BS, Woldemariam MG, Shai B, Obolski U, Ehrlich M, Shani E, Jander G, Chamovitz DA. (2015) The glucosinolate breakdown product indole-3-carbinol acts as an auxin antagonist in roots of *Arabidopsis thaliana*. The Plant Journal. **82:** 547-555

Kavalier AR, Pitra NJ, Koelling JM, Coles MC, Kennelly EJ, Matthews PD. (2011) Increase in cone biomass and terpenophenolics in hops (*Humulus lupulus* L.) by treatment with prohexadione-calcium. Journal of Agricultural and Food Chemistry. **59**: 6720–6729

Kebrom TH. (2017) A growing stem inhibits bud outgrowth – the overlooked theory of apical dominance. Frontiers in Plant Science. **8**: 1874

Keefover-Ring K, Carlsson M, Albrectsen BR. (2014) 2′-(Z)-Cinnamoylsalicortin: a novel salicinoid isolated from *Populus tremula*. Phytochemistry Letters. **7**: 212–216

Keefover-Ring K, Ahnlund M, Abreu IN, Jansson S, Moritz T, Albrectsen BR. (2014) No evidence of geographical structure of salicinoid chemotypes within *Populus tremula*. PLoS ONE. **9:** e107189

Kessler A, Kalske A. (2018) Plant secondary metabolite diversity and species interactions. Annual Review of Ecology, Evolution, and Systematics. **49**: 115-138

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnology. **37**: 907–915

Kim JK, Bae YS. (2009) Grandidentatin isomer from bark of suwon poplar (*Populus alba L.*×*Populus glandulosa* Uyeki). Holzforschung. **63**: 315–318

Klempien A, Kaminaga Y, Qualley A, Nagegowda DA, Widhalm JR, Orlova I, Shasany AK, Taguchi G, Kish CM, Cooper BR, D'Auria JC, Rhodes D, Pichersky E, Dudareva N. (2012) Contribution of CoA ligases to benzenoid biosynthesis in petunia flowers. The Plant Cell. **24**: 2015–2030

Kruger EL, Keefover-Ring K, Holeski LM, Lindroth RL. (2020) To compete or defend: linking functional trait variation with life-history tradeoffs in a foundation tree species. Oecologia. **192**: 893-907

Kulasekaran S, Cerezo-Medina S, Harflett C, Lomax C, de Jong F, Rendour A, Ruvo G, Hanley S, Beale M, Ward JL. (2020) A willow UDP-glycosyltransferase involved in salicinoid biosynthesis. Journal of Experimental Botany. **72**: 1634-1648

Lackus ND, Müller A, Kröber TDU, Reichelt M, Schmidt A, Nakamura Y, Paetz C, Luck K, Lindroth R, Constabel CP, Unsicker S, Gershenzon J, Köllner TG. (2020) The occurrence of sulfated salicinoids in poplar and their formation by sulfotransferase. Plant Physiology. **183**: 137–151

Lackus ND, Schmidt A, Gershenzon J, Köllner TG. (2021) A peroxisomal β-oxidative pathway contributes to the formation of C6–C1 aromatic volatiles in poplar. Plant Physiology. **186**: 891-909

Landhäusser SM, Lieffers VJ. (2012) Defoliation increases risk of carbon starvation in root systems of mature aspen. Trees. **26**: 653–661

Landhäusser SM, Chow PS, Dickman LT, Furze ME, Kuhlman I, Schmid S, Wiesenbauer J, Wild B, Gleixner G, Hartmann H, Hoch G, Mcdowell NG, Richardson AD, Richter A, Adams HD. (2018) Standardized protocols and procedures can precisely and accurately quantify non-structural carbohydrates. Tree Physiology. **38**: 1764-1778

Lastra R, Kenkel N, Daayf F. (2017) Phenolic glycosides in *populus tremuloides* and their effects on long-term ungulate browsing. Journal of Chemical Ecology. **43**: 1023-1030

Laursen T, Borch J, Knudsen C, Bavishi K, Torta F, Martens HJ, Silvestro D, Hatzakis NS, Wenk MR, Dafforn TR, Olsen CE, Motawia MS, Hamberger B, Møller BL, Bassard J. (2016). Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum. Science. **354**: 890–893

Lee H, Leon J, Raskin I. (1995) Biosynthesis and metabolism of salicylic acid. Proceedings of the National Academy of Sciences USA. **92**: 493–496

León J, Shulaev V, Yalpani N, Lawton MA, Raskin I. (1995) Benzoic acid 2-hydroxylase, a soluble oxygenase from tobacco, catalyzes salicylic acid biosynthesis. Proceedings of the National Academy of Science USA. **92**: 10413–10417

Liechti R, Farmer EE. (2002) The jasmonate signal pathway. The Plant Cell. **296**: 1649–1650

Lindroth RL, St. Clair SB. (2013) Adaptations of quaking aspen (*Populus tremuloides* Michx.) for defense against herbivores. Forest Ecology Management. **299**: 14–21

Liu L, Sonbol FM, Huot B, Gu Y, Withers J, Mwimba M, Yao JHS, Dong X. (2016) Salicylic acid receptors activate jasmonic acid signalling through a non-canonical pathway to promote effector-triggered immunity. Nature Communications. **7**: 13099

Liu Q, Ding C, Chu Y, Chen J, Zhang W, Zhang B. (2016) PoplarGene: poplar gene network and resource for mining functional information for genes from woody plants. Scientific Reports. **6**: 31356

Liu Y, Sun T, Sun Y, Zhang Y, Radojičić A, Ding Y, Tian H, Huang X, Lan J, Chen S, Orduna A, Zhang K, Jetter R, Li X, Zhang Y. (2020) Diverse roles of the salicylic acid receptors NPR1 and NPR3/NPR4 in plant immunity. The Plant Cell. **32**: 4002-4016

Love MI, Huber W, Anders S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. **15**: 550

Lou YR, Pichersky E, Last RL. (2022) Deep roots and many branches: origins of plant-specialized metabolic enzymes in general metabolism. Current Opinion in Plant Biology. **66**: 102192

Luo J, Xia W, Cao P, Xiao Z, Zhang Y, Liu M, Zhan C, Wang N. (2019) Integrated transcriptome analysis reveals plant hormones jasmonic acid and salicylic acid coordinate growth and defense responses upon fungal infection in poplar. Biomolecules. **9**: 12

Ma D, Reichelt M, Yoshida K, Gershenzon J, Constabel CP. (2018) Two R2R3-MYB proteins are broad repressors of flavonoid and phenylpropanoid metabolism in poplar. The Plant Journal. **96**: 949–965

MacNeill GJ, Mehrpouyan S, Minow MAA, Patterson JA, Tetlow IJ, Emes MJ. (2017) Starch as a source, starch as a sink: The bifunctional role of starch in carbon allocation. Journal of Experimental Botany. **68**: 4433–4453

Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R. (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Research. **50**: W276–W279

Mader M, Le Paslier MC, Bounon R, Bérard A, Rampant PF, Fladung M, Leplé JC, Kersten B. (2016) Whole-genome draft assembly of *Populus tremula x P. alba* clone INRA 717-1B4. Silvae Genetica. **65**: 74–79

Madritch MD, Lindroth RL. (2015) Condensed tannins increase nitrogen recovery by trees following insect defoliation. New Phytologist. **208**: 410–420

Maeda H, Dudareva N. (2012) The shikimate pathway and aromatic amino acid biosynthesis in plants. Annual Reviews in Plant Biology. **63**: 73–105

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

Major IT, Guo G, Zhai J, Kapali G, Kramer DM, Howe GA. (2020) A phytochrome B-independent pathway restricts growth at high levels of jasmonate defense. Plant Physiology. **183**: 733-749

Malinovsky FG, Thomsen MLF, Nintemann SJ, Jagd LM, Bourgine B, Burow M, Kliebenstein DJ. (2017) An evolutionarily young defense metabolite influences the root growth of plants via the ancient TOR signaling pathway. Elife. **6**: e29353

Marchais M, Arseneault D, Bergeron Y. (2022) The rapid expansion of *Populus tremuloides* due to anthropogenic disturbances in eastern Canada. Canadian Journal of Forest Research. **52**: 991–1001

Martínez-Vilalta J, Sala A, Asensio D, Galiano L, Hoch G, Palacio S, Piper FI, Lloret F. (2016) Dynamics of non-structural carbohydrates in terrestrial plants: a global synthesis. Ecological Monographs. **86**: 495-516

Marín L, Gutiérrez-del-Río I, Entrialgo-Cadierno R, Claudio Villar J, Lombó F. (2018) De novo biosynthesis of myricetin, kaempferol and quercetin in *Streptomyces albus* and *Streptomyces coelicolor*. PLoS ONE. **13**: e0207278

Martin C, Smith A. (1995) Starch biosynthesis. The Plant Cell. **7**: 971-985

Maruri-López I, Aviles-Baltazar NY, Buchala A, Serrano M. (2019) Intra and extracellular journey of the phytohormone salicylic acid. Frontiers in Plant Science. **10**: 1–11.

Mason MG, Ross JJ, Babst BA, Wienclaw BN, Beveridge CA. (2014) Sugar demand, not auxin, is the initial regulator of apical dominance. Proceedings of the National Academy of Science USA. **111**: 6092-6097

Mellway RD, Tran LT, Prouse MB, Campbell MM, Constabel CP. (2009) The wound-, pathogen-, and ultraviolet B-Responsive MYB134 gene encodes an R2R3 MYB transcription factor that regulates proanthocyanidin synthesis in poplar. Plant Physiology. **150**: 924–941

Messiano GB, da Silva T, Nascimento IR, Lopes LMX. (2009) Biosynthesis of antimalarial lignans from *Holostylis reniformis*. Phytochemistry. **70**: 590–596

Montinari MR, Minelli S, De Caterina R. (2019) The first 3500 years of aspirin history from its roots – a concise summary. Vascular Pharmacology. **113**: 1–8

Moreira-Vilar FC, Siqueira-Soares RDC, Finger-Teixeira A, De Oliveira DM, Ferro AP, Da Rocha GJ, Ferrarese MDLL, Dos Santos WD, Ferrarese-Filho O. (2014) The acetyl bromide method is faster, simpler and presents best recovery of lignin in different herbaceous tissues than klason and thioglycolic acid methods. PLoS ONE. **9**: e110000

Morse AM, Tschaplinski TJ, Dervinis C, Pijut PM, Schmelz EA, Day W, Davis JM. (2007) Salicylate and catechol levels are maintained in *nahG* transgenic poplar. Phytochemistry. **68**: 2043–2052

Moritz F, Kaling M, Schnitzler JP, Schmitt-Kopplin P. (2017) Characterization of poplar metabotypes via mass difference enrichment analysis. Plant Cell and Environment. **40**: 1057–1073

Mucha S, Heinzlmeir S, Kriechbaumer V, Strickland B, Kirchhelle C, Choudhary M, Kowalski N, Eichmann R, Hückelhoven R, Grill E, Kuster B, Glawischnig E. (2019) The formation of a camalexin biosynthetic metabolon. The Plant Cell. **31**: 2697–2710

Mugford SG, Yoshimoto N, Reichelt M, Wirtz M, Hill L, Mugford ST, Nakazato Y, Noji M, Takahashi H, Kramell R, Gigolashvilli T, Flugge U, Wasternack C, Gershenzon J, Hell R, Saito K, Kopriva S. (2009) Disruption of adenosine-5′-phosphosulfate kinase in *Arabidopsis* reduces levels of sulfated secondary metabolites. The Plant Cell. **21**: 910–927

Neilson EH, Goodger JQD, Woodrow IE, Møller BL. (2013) Plant chemical defense: at what cost? Trends in Plant Science. **18**: 250-258

Noleto-Dias C, Ward JL, Bellisai A, Lomax C, Beale MH. (2018) Salicin-7-sulfate: A new salicinoid from willow and implications for herbal medicine. Fitoterapia. **127**: 166–172

Orlova I, Marshall-Colón A, Schnepp J, Wood B, Varbanova M, Fridman E, Blakeslee JJ, Peer WA, Murphy AS, Rhodes D, Pichersky E, Dudareva N. (2006) Reduction of benzenoid synthesis in petunia flowers reveals multiple pathways to benzoic acid and enhancement in auxin transport. The Plant Cell. **18**: 3458–3475

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

Osier TL, Hwang S, Lindroth RL. (2000) Within- and between-year variation in early season phytochemistry of quaking aspen (*Populus tremuloides* Michx.) clones. Biochemical Systematics and Ecolology. **28**: 197-208

Palo TR. (1984) Distribution of birch (*Betula* spp.), willow (*Salix* spp.) and poplar (*Populus* spp.) secondary metabolites and their potential role as chemical defense against herbivores. Journal of Chemical Ecology. **10**: 499-520

Park SB, Kim J, Han J, Ahn CH, Park EJ, Choi Y. (2017) Exploring genes involved in benzoic acid biosynthesis in the *Populus davidiana* transcriptome and their transcriptional activity upon methyl jasmonate treatment. Journal of Chemical Ecology. **43:** 1097-1108

Payyavula RS, Babst BA, Nelsen MP, Harding SA, Tsai CJ. (2009) Glycosylation-mediated phenylpropanoid partitioning in *Populus tremuloides* cell cultures. BMC Plant Biology. **9**: 151

Pearl IA, Darling SF, Justman O. (1962) Studies on the leaves of the family Salicaceae: I. populin from the leaves of *Populus grandidentata* and *Populus tremuloides*. Journal of Organic Chemistry. **27**: 2685–2687

Peer WA, Murphy AS. (2007) Flavonoids and auxin transport: modulators or regulators? Trends in Plant Science. **12:** 556-563

Pichersky E. Lewinsohn E. (2011) Convergent evolution in plant specialized metabolism. Annual Reviews in Plant Biology. **62**: 549–566

Piper FI, Fajardo A. (2016) Carbon dynamics of *Acer pseudoplatanus* seedlings under drought and complete darkness. Tree Physiology. **36**: 1400-1408

Pluskal T, Castillo S, Villar-Briones A, Oresic M. (2010) MZmine2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics. **11**: 395

Popko J, Hänsch R, Mendel RR, Polle A, Teichmann T. (2010) The role of abscisic acid and auxin in the response of poplar to abiotic stress. Plant Biology. **12**: 242–258

Raskin I. (1992) Salicylate, a new plant hormone. Plant Physiology. **99**: 799–803

Rekhter D, Lüdke D, Ding Y, Feussner K, Zienkiewicz K, Lipka V, Wiermer M, Zhang Y, Feussner I. (2019) Isochorismate-derived biosynthesis of the plant stress hormone salicylic acid. Science. **365**: 498–502

Rigsby CM, Body MJA, May A, Oppong A, Kostka A, Housman N, Savage S, Whitney ER, Kinahan IG, Deboef B, Orians CM, Appel HM, Schultz JC, Preisser EL. (2020) Impact of chronic stylet-feeder infestation on folivore-induced signaling and defenses in a conifer. Tree Physiology. **41**: 416–427

Rivas-San Vicente M, Plasencia J. (2011) Salicylic acid beyond defence: its role in plant growth and development. Journal of Experimental Botany. **62**: 3321-3338

Rivero-Cruz B, Rojas MA, Rodríguez-Sotres R, Cerda-García-Rojas CM, Mata R. (2005) Smooth muscle relaxant action of benzyl benzoates and salicylic acid derivatives from *Brickellia veronicaefolia* on isolated guinea-pig ileum. Planta Medica. **71**: 320–325

Rivero-Cruz B, Rivero-Cruz I, Rodríguez JM, Cerda-García-Rojas CM, Mata R. (2006) Qualitative and quantitative analysis of the active components of the essential oil from *Brickellia veronicaefolia* by nuclear magnetic resonance spectroscopy. Journal of Natural Products. **69**: 1172–1176

RStudio Team (2018) Rstudio: Integrated Development for R. Boston, MA: Rstudio Inc. URL http://www.rstudio.com/

R Core Team. (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ruan Y. (2014) Sucrose Metabolism: Gateway to Diverse Carbon Use and Sugar Signaling. Annual Reviews in Plant Biology. **65**: 33-67

Rogers PC, Pinno BD, Sebesta J, Albrectsen BR, Li G, Ivanova N, Kusbach A, Kuuluvainen T, Landhäusser SM, Liu H, Myking T, Pulkkinen P, Wen Z, Kulakowski D. (2020) A global view of aspen: Conservation science for widespread keystone systems. Global Ecology and Conservation. **21**: e00828

Rubert-Nason KF, Couture JJ, Major IT, Constabel CP, Lindroth RL. (2015) Influence of genotype, environment, and gypsy moth herbivory on local and systemic chemical defenses in trembling aspen (*Populus tremuloides*). Journal of Chemical Ecology. **41**: 651–661

Ruuhola T, Julkunen-Tiitto R. (2002) Salicylates of Intact *Salix myrsinifolia* plantlets do not undergo rapid metabolic turnover. Plant Physiology. **122**: 895-905

Ruuhola T, Julkunen-Tiitto R. (2003) Trade-off between synthesis of salicylates and growth of micropropogated *Salix pentandra*. Journal of Chemical Ecology. **29**: 1565-1588

Ruuhola T, Julkunen-Tiitto R, Vainiotalo P. (2003) *In vitro* degradation of willow salicylates. Journal of Chemical Ecology. **29**: 1083-1097

Schädel C, Blöchl A, Richter A, Hoch G. (2009) Short-term dynamics of nonstructural carbohydrates and hemicelluloses in young branches of temperate forest trees during bud break. Tree Physiology. **29**: 901-911

Schläpfer P, Zhang P, Wang C, Kim T, Banf M, Chae L, Dreher K, Chavali AK, Nilo-Poyanco R, Bernard T, Kahn D, Rhee SY. (2017) Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. Plant Physiology. **173**: 2041–2059

Schmelz EA, Leclere S, Carroll MJ, Alborn HT, Teal PEA. (2007) Cowpea chloroplastic ATP synthase is the source of multiple plant defense elicitors during insect herbivory. Plant Physiology. **144**: 793–805

Sevanto S, Mcdowell NG, Dickman LT, Pangle R, Pockman WT. (2014) How do trees die? A test of the hydraulic failure and carbon starvation hypotheses. Plant, Cell, and Environment. **37**: 153-161

Shen H, Strunks GD, Klemann BJPM, Hooykaas PJJ, de Pater S. (2017) CRISPR/Cas9-induced double-strand break repair in *Arabidopsis* nonhomologous end-joining mutants. G3: Genes Genomes Genetics. **7**: 193–202

Shine MB, Yang JW, El-Habbak M, Nagyabhyru P, Fu DQ, Navarre D, Ghabrial S, Kachroo P, Kachroo A. (2016) Cooperative functioning between phenylalanine ammonia lyase and isochorismate synthase activities contributes to salicylic acid biosynthesis in soybean. New Phytologist. **212**: 627–636

Shiraiwa, K. Yuan, S., Fujiyama, A., Matsuo, Y., Tanaka, T., Jiang, Z., Kouno, I. (2012) Benzyl benzoate glycoside and 3-deoxy-d-manno-2-octulosonic acid derivatives from *Solidago decurrens*. Journal of Natural Products. **75**: 88-92

Siritunga D, Sayre R. (2004) Engineering cyanogen synthesis and turnover in cassava (*Manihot esculenta*). Plant Molecular Biology. **56**: 661–669

Sjödin A, Street NR, Sandberg G, Gustafsson P, Jansson S. (2009) The Populus Genome Integrative Explorer (PopGenIE): A new resource for exploring the *Populus* genome. New Phytologist. **182**: 1013–1025

Soubeyrand E, Johnson TS, Latimer S, Block A, Kim J, Colquhoun TA, Butelli E, Martin C, Wilson MA, Basseta GJ. (2018) The peroxidative cleavage of kaempferol contributes to the biosynthesis of the benzenoid moiety of ubiquinone in plants. The Plant Cell. **30**: 2910–2921

Soubeyrand E, Latimer S, Bernert AC, Keene SA, Johnson TS, Shin D, Block AK, Colquhoun TA, Schäffner AR, Kim J, Basset GJ. (2021) 3-O-glycosylation of kaempferol restricts the supply of the benzenoid precursor of ubiquinone (Coenzyme Q) in *Arabidopsis thaliana*. Phytochemistry. **186**: 112738

Sperschneider J, Catanzariti AM, Deboer K, Petre B, Gardiner DM, Singh KB, Dodds PN, Taylor JM. (2017) LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. Scientific Reports. **7**: 44598

Stanton BJ, Bourque A, Coleman M, Eisenbies M, Emerson RM, Espinoza J, Gantz C, Himes A, Rodstrom A, Shuren R, Stonex R, Volk T, Zerpa J. (2021) The practice and economics of hybrid poplar biomass production for biofuels and bioproducts in the Pacific Northwest. Bioenergy Research. **14**: 543–560

Stark R, Grzelak M, Hadfield J. (2019) RNA sequencing: the teenage years. Nature Reviews Genetics. **20:** 631-656

Stein O, Granot D. (2019) An overview of sucrose synthases in plants. Frontiers in Plant Science. **10**: 95

Stevens M, Lindroth R. (2005) Induced resistance in the indeterminate growth of aspen (*Populus tremuloides*). Oecologia. **145**: 298-306

Steyn WJ, Wand SJE, Holcroft DM, Jacobs G. (2002) Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. New Phytologist. **155**: 349–361

Tahvanainen AJ, Helle E, Julkunen-Tiitto R, Lavola A. (1985) Phenolic compounds of willow bark as deterrents against feeding by mountain hare. Oecologia. **65**: 319–323

Thompson K, Hodgson JG, Smith RM, Warren PH, Gaston KJ. (2004) Urban domestic gardens (III): composition and diversity of lawn floras. Journal of Vegetation Science. **15**: 373–378

Torrens-Spence MP, Bobokalonova A, Carballo V, Glinkerman CM, Pluskal T, Shen A, Weng JK. (2019) PBS3 and EPS1 complete salicylic acid biosynthesis from isochorismate in *Arabidopsis*. Molecular Plant. **12**: 1577–1586

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg S, Rinn J, Pachter L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols. **7**: 562–578

Traw MB, Bergelson J. (2003) Interactive effects of jasmonic acid, salicylic acid, and gibberellin on induction of trichromes in Arabidopsis. Plant Physiology. **133**: 1367–1375

Tsai CJ, Harding SA, Tschaplinski TJ, Lindroth RL, Yuan Y. (2006) Genome-wide analysis of the structural genes regulating defense phenylpropanoid metabolism in *Populus*. New Phytologist. **172**: 47–62

Tsai CJ, Guo W, Babst B, Nyamdari B, Yuan Y, Payyavula R, Chen HY, Liangjiao X, Tay K, Michelizzi V, Harding S. (2011) Salicylate metabolism in *Populus*. BMC Proceedings. **5**: 1–2

Tuskan, GA., Difazio S., Jansson S., et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science. **313**: 1596–1604

Ullah C, Tsai CJ, Unsicker SB, Xue L, Reichelt M, Gershenzon J, Hammerbacher A. (2019) Salicylic acid activates poplar defense against the biotrophic rust fungus *Melampsora larici-populina* via increased biosynthesis of catechin and proanthocyanidins. New Phytologist. **221**: 960–975

van Butselaar T, Van den Ackerveken G. (2020) Salicylic acid steers the growth–immunity tradeoff. Trends in Plant Science. **25**: 566-576

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

Veillet F, Kermarrec MP, Chauvin L, Chauvin JE, Nogué F. (2020) CRISPR-induced indels and base editing using the *Staphylococcus aureus* Cas9 in potato. PloS One. **15**: e0235942

Vlot AC, Dempsey DA, Klessig DF. (2009) Salicylic acid, a multifaceted hormone to combat disease. Annual Review of Phytopathology. **47**: 177–206

Vogt T. (2010) Phenylpropanoid biosynthesis. Molecular Plant. **3**: 2–20

Wang B, Jin SH, Hu HQ, Sun YG, Wang YW, Han P, Hou BK. (2012) UGT87A2, an Arabidopsis glycosyltransferase, regulates flowering time via FLOWERING LOCUS C. New Phytologist. **194**: 666–675

Ward JL, Wu Y, Harflett C, Onafuye H, Corol D, Lomax C, Macalpine WJ, Cinatl J, Wass MN, Michaelis M, Beale MH. (2020) Miyabeacin: A new cyclodimer presents a potential role for willow in cancer therapy. Scientific Reports. **10**: 6477

Wasternack C, Hause B. (2013) Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. Annals of Botany. **111:** 1021-1058

Whitehill JGA, Yuen MMS, Henderson H, Madilao L, Kshatriya K, Bryan J, Jaquish B, Bohlmann J, (2018) Functions of stone cells and oleoresin terpenes in the conifer defense syndrome. New Phytologist. **221**: 1503–1517

Widhalm JR, Dudareva N (2015) A familiar ring to it: biosynthesis of plant benzoic acids. Molecular Plant. **8:** 83-97

Wildermuth MC, Dewdney J, Wu G, Ausubel FM. (2001) Isochorismate synthase is required to synthesize salicylic acid for plant defence. Nature. **417**: 562-565

Wiley E, Hoch G, Landhäusser SM. (2017) Dying piece by piece: carbohydrate dynamics in aspen (*Populus tremuloides*) seedlings under severe carbon stress. Journal of Experimental Botany. **68**: 5221-5232

Wiley E, King CM, Landhäusser SM. (2019) Identifying the relevant carbohydrate storage pools available for remobilization in aspen roots. Tree Physiology. **39**: 1109-1120

Winkel BSJ. (2004) Metabolic channeling in plants. Annual Reviews in Plant Biology. **55**: 85–107

Wooley SC, Walker S, Vernon J, Lindroth RL. (2008) Aspen decline, aspen chemistry, and elk herbivory: are they linked? Rangelands. **30**: 17–21

Wu J, Baldwin IT. (2009) Herbivory-induced signalling in plants: perception and action. Plant, Cell, and Environment. **32**: 1161–1174

Wullschleger SD, Jansson S, Taylor G. (2002) Genomics and forest biology: *Populus* emerges as the perennial favorite. The Plant Cell. **14**: 2651–2655

Xia W, Yu H, Cao P, Luo J, Wang N. (2017) Identification of TIFY family genes and analysis of their expression profiles in response to phytohormone treatments and *Melampsora larici-populina* infection in poplar. Frontiers in Plant Science. **8:** 493

Xiao S, Dai L, Liu F, Wang Z, Peng W, Xie D. (2004) COS1: An *Arabidopsis* coronatine insensitive1 suppressor essential for regulation of jasmonate-mediated plant defense and senescence. The Plant Cell. **16**: 1132–1142

Xue LJ, Alabady MS, Mohebbi M, Tsai CJ. (2015) Exploiting genome variation to improve next-generation sequencing data analysis and genome editing efficiency in *Populus tremula* × *alba* 717-1B4. Tree Genetics and Genomics. **11**: 82

Xue LJ, Guo W, Yuan Y, Anino, EO, Nyamdari B, Wilson MC, Frost C, Chen H, Babst B, Harding S, Tsai CJ. (2013) Constitutively elevated salicylic acid levels alter photosynthesis and oxidative state but not growth in transgenic *Populus*. The Plant Cell. **25**: 2714–2730

Yang DL, Yao J, Mei CS, Tong XH, Zeng LJ, Li Q, Xiao L, Sun T, Li J, Deng X, Lee C, Thomashow M, Yang Y, He Z, He SY. (2012) Plant hormone jasmonate prioritizes defense over growth by interfering with gibberellin signaling cascade. Proceedings of the National Academy of Science USA. **109**: E1192-E1200

Yazaki K, Arimura GI, Ohnishi T. (2017) "Hidden" terpenoids in plants: their biosynthesis, localization and ecological roles. Plant Cell Physiology. **58**: 1615–1621

Yokoyama R, de Oliveira MVV, Kleven B, Maeda HA. (2021) The entry reaction of the plant shikimate pathway is subjected to highly complex metabolite-mediated regulation. The Plant Cell. **33:** 671–696

Yonekura-sakakibara K, Saito K. (2009) Functional genomics for plant natural product biosynthesis. Natural Product Reports. **26**: 1466–1487

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 Physiology. **167**: 693–710

Zenk MH. (1967) Pathways of salicyl alcohol and salicin formation in *Salix purpurea* L. Phytochemistry. **6**: 245-252

Zhang K, Halitschke R, Yin C, Liu CJ, Gan SS. (2013) Salicylic acid 3-hydroxylase regulates *Arabidopsis* leaf longevity by mediating salicylic acid catabolism. Proceedings of the National Academy of Science USA. **110**: 14807–14812

Zhang S, Peng F, Xiao Y, Wang W, Wu X. (2020) Peach PpSnRK1 participates in sucrose-mediated root growth through auxin signaling. Frontiers in Plant Science. **11**: 409

Zhang XF, Thuong PT, Min BS, Ngoc TM, Hung TM, Lee IS, Na M, Seong Y, Song K, Bae KH. (2006) Phenolic glycosides with antioxidant activity from the stem bark of *Populus davidiana*. Journal of Natural Products. **69**: 1370–1373

Zhang Y, Turner, JG. (2008) Wound-induced endogenous jasmonates stunt plant growth by inhibiting mitosis. PLoS ONE. **3**: e3699

Zhang Y, Li X. (2019) Salicylic acid: biosynthesis, perception, and contributions to plant immunity. Current Opinion in Plant Biology. **50**: 29–36

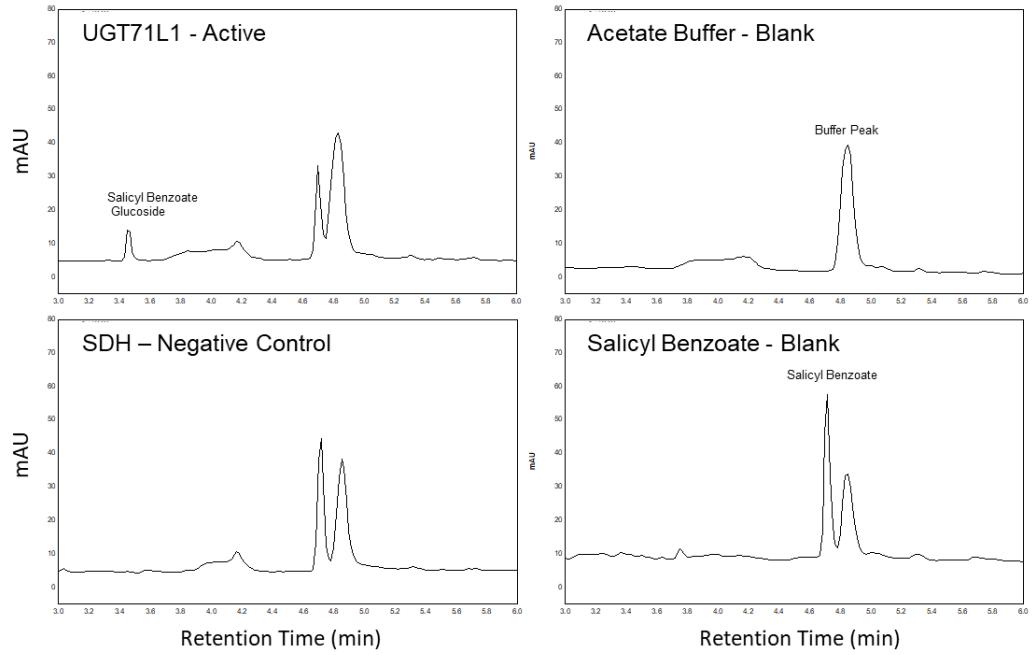
Zhang Y, Zhao L, Zhao J, Li Y, Wang J, Guo R, Gan S, Liu, CJ, Zhanga K. (2017) S5H/DMR6 encodes a salicylic acid 5-hydroxylase that fine-tunes salicylic acid homeostasis. Plant Physiology. **175**: 1082–1093

Zhao Q, Chen MY, Poisson T, Pannecoucke X, Bouillon JP, Besset T. (2018) Pd-Catalyzed trifluoromethylthiolation of unsaturated compounds: a general approach. European Journal of Organic Chemistry. **2018**: 6167-6175

Zhou K, Hu L, Li Y, Chen X, Zhang Z, Liu B, Li P, Gong X, Ma F. (2019) MdUGT88F1-mediated phloridzin biosynthesis regulates apple development and valsa canker resistance. Plant Physiology. **180**: 2290–2305

Appendices

Appendix 1. UGT71L1 salicyl benzoate enzyme assay control analyses



Appendix Figure A1.1 UGT71L1 salicyl benzoate enzyme assay control analyses. Top left, Chromatogram of recombinant UGT71L1 enzyme assay. Top right, sodium acetate buffer (50 mM, pH 5.1) blank. Bottom left, shikimate dehydrogenase (SDH) enzyme assay. Bottom right, salicyl benzoate substrate (100 µM) blank with no added protein. All absorbances (mAU) were measured at 280 nm with reactions separated by reverse phase chromatography using a Kinetex C18 column (50 x 4.6 mm, 5 µm; Phenomenex, Torrance, CA, USA) on an HPLC (System Gold, Beckman Coulter, Mississauga, ON, Canada). UGT71L1 and SDH reactions were run with 5 µg protein per 200 µL reaction. Enzyme reactions were run for three hours at 30 ºC with 100 µM salicyl benzoate, 5 mM UDP-glucose in a 50 mM sodium acetate buffer at pH 5.1 as described by Fellenberg et al. (2020).

Appendix 2. Analysis of UGT71m overexpressor poplar.

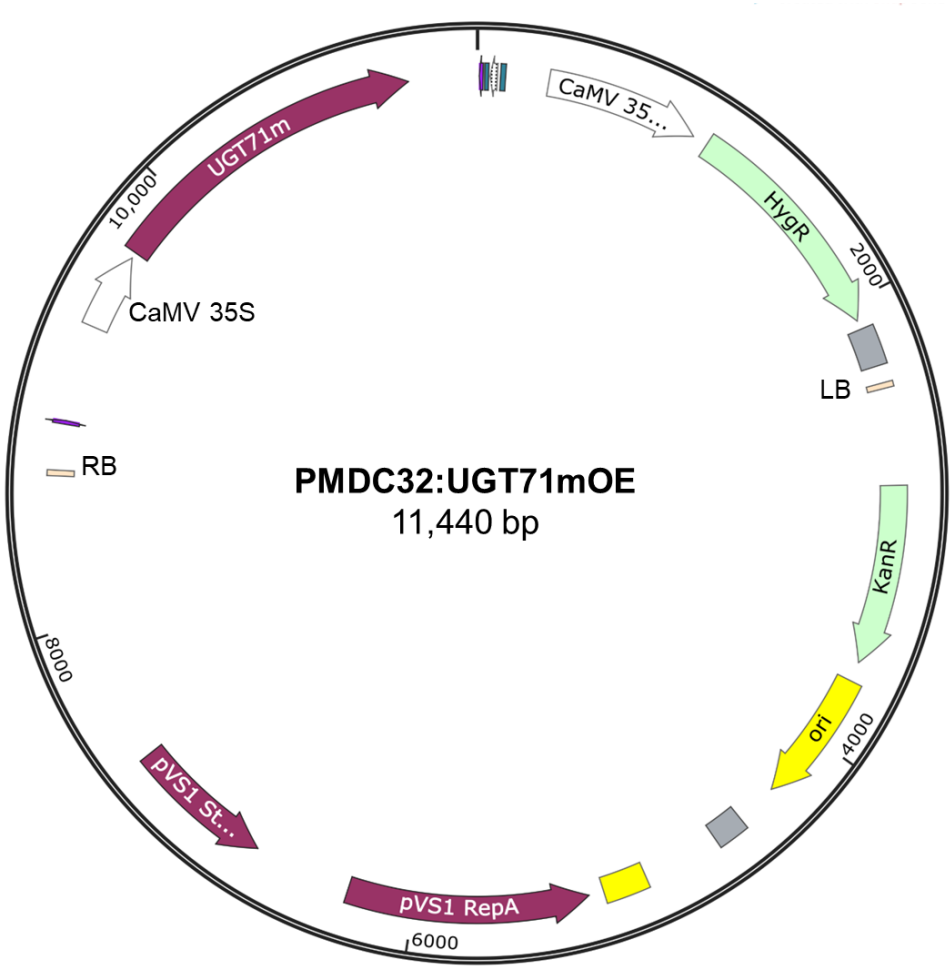


Figure A2.1. Vector map of UGT71m overexpressing construct. Construct map prepared by Snapgene (GSL Biotech, San Diego, CA, USA). PMDC32:UGT71m was constructed and transformants generated as per section 2.2.6. Abbreviations: HygR, hygromycin resistance; KanR, kanamycin resistance; CAMV, Cauliflower Mosaic Virus; LB/RB, left and right border; ori, origin of replication; bom, basis of mobility region; RepA, replication protein; StaA, stability protein; bp, base pairs.

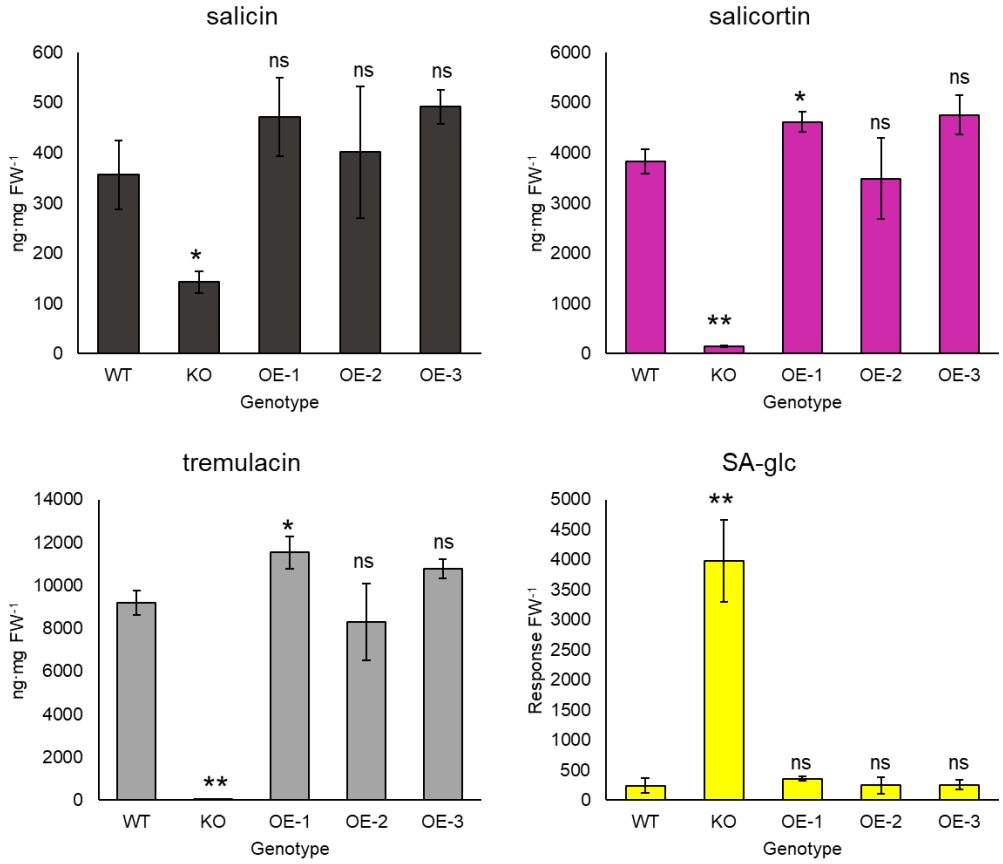


Figure A2.2. Salicinoid and salicylic acid glucoside concentrations in UGT71L1m overexpressor poplar plants. Panels show concentrations of major salicinoids and salicylic acid glucoside (SA-glc) in expanded leaves of 2-month-old greenhouse-grown poplars. Values shown are means ± SE. WT, wild-type with 5 biological replicates; KO, UGT71L1-KO with 4 biological replicates. Overexpressor (OE)-1, -2, and -3 represent independently transformed lines, with 4 - 5 biological replicates per line. Statistical significance was determined using a Welch's *t*-test. * Indicates significant ($P < 0.05$) differences from WT plants; ** indicates $P < 0.01$, ns = not-significant.

Appendix 3. Generation of UGT71L1 Promoter GUS expression construct

Appendix 3.1 Methods for the generation of GUS expression vector.

A GUS expression vector under the control if the UGT71L1 native promoter (NP) was prepared from the UGT71L1m Overexpressor construct presented in appendix 2. The NP was synthetically produced (Biomatik, Kitchener, ON, Canada) using the 1500 preceding bases of UGT71L1 of the *P. tremula* genome presented in AspenDB (Xue et al., 2015) . The NP was synthesized with a 5' HindIII site and a 3' KpnI site. HindIII and KpnI were used to excise the CAM35s promoter from PMDC32:UGT71L1mOE, and the NP was amplified using PCR and custom primers (F: GGGAAGCTTAGATTTGAACTTTGTACTTG, R: GGGGGTACCGTTTTAAGATATGGTTTTATG) and subsequently ligated into the plasmid. A PCR product of a GUSNosT construct was generated from a pRD410 plasmid (Datla et al., 1992) using custom primers (F: TATAGGTACCATGTTACGTCCTGTAG, R: GTCATCTATGTTACTAGATCACTAGTATAT) containing a 5' KpnI site and a 3' SpeI site. UGT71L1m was excised using KpnI and SpeI from the NP:UGT71L1m plasmid and the GUSNosT PCR product was ligated using the same restriction enzyme sites. Constructs were confirmed by Sanger sequencing.

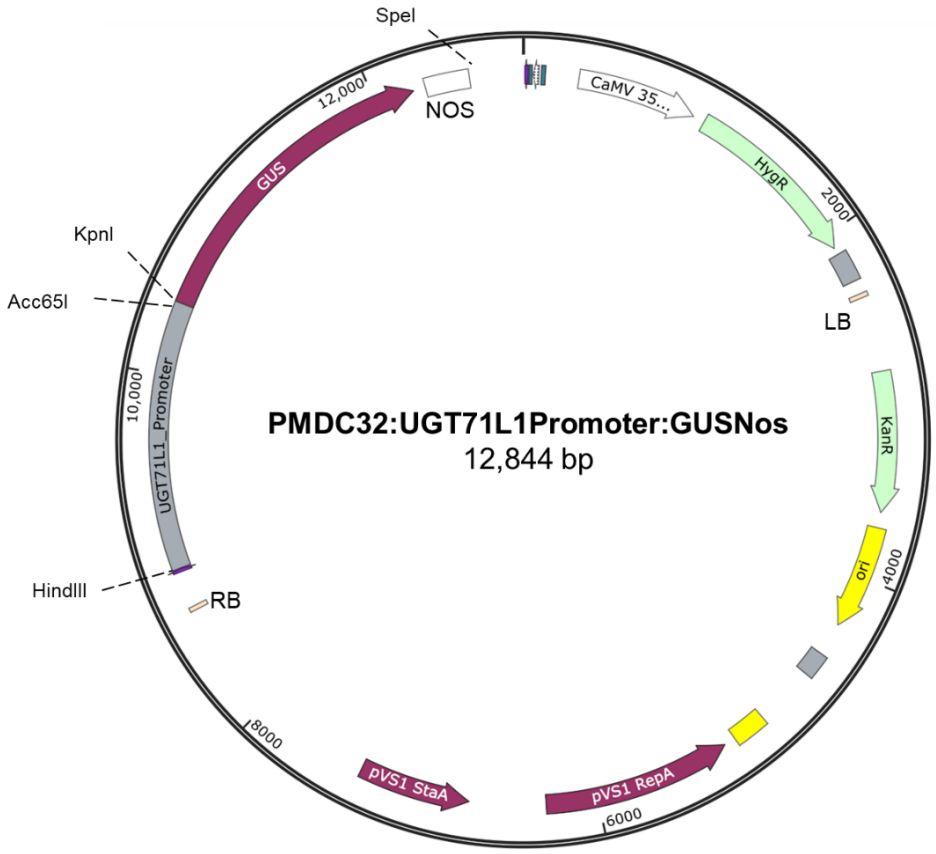


Figure A3.1. UGT71L1 Promoter : GUS expression construction in a PMDC32 background vector. Dashed lines indicate select restriction enzyme cut sites. Abbreviations: HygR, hygromycin resistance; KanR, kanamycin resistance; CAMV, Cauliflower Mosaic Virus; LB/RB, left and right border; ori, origin of replication; bom, basis of mobility region; RepA, replication protein; StaA, stability protein; bp, base pairs; GUS, β-glucuronidase; NOS, nopaline synthase terminator.

Appendix 4. Analysis of cinnamoyl salicortin isomers in SCPL knockout poplar.

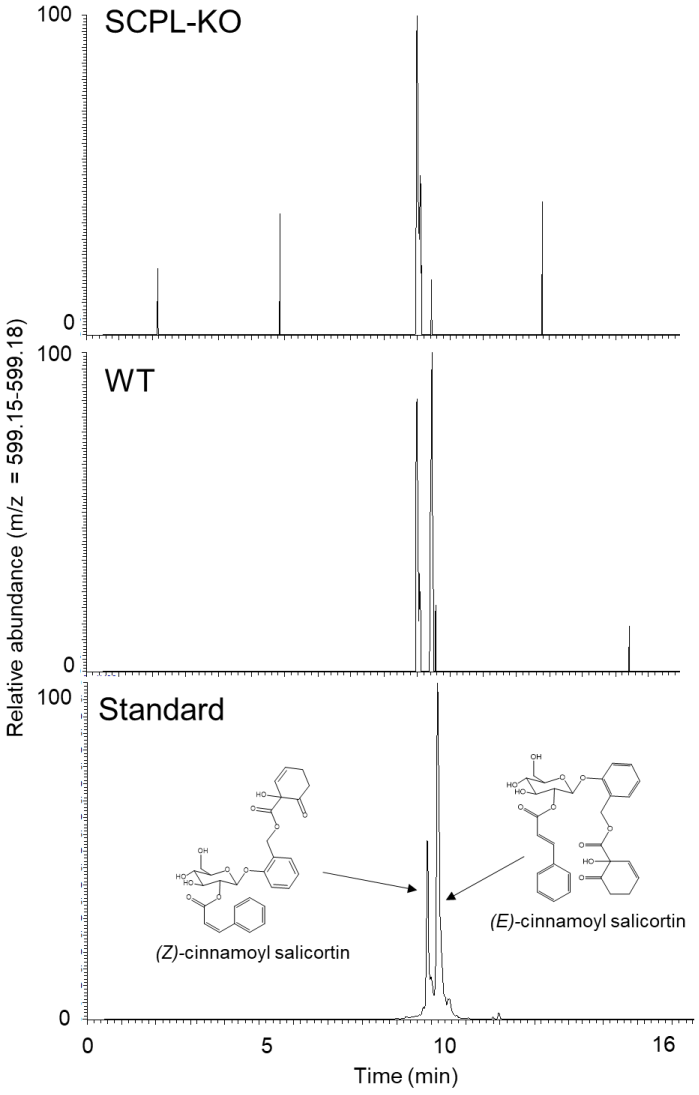


Figure A4.1. Extracted ion chromatograms (EIC) of cinnamoyl salicortin isomers in poplar leaf extracts. All chromatograms present a m/z of 599.15 - 599.18 in negative ionization mode. Top, representative EIC of young leaf tissue of a SCPL knockout (KO) line displaying a single cinnamoyl salicortin isomer. Middle, representative EIC of wild-type (WT) young leaves displaying two cinnamoyl salicortin isomers. Bottom, EIC of enriched (E)-cinnamoyl salicortin standard provided by Keefover et al. (2014). Peaks of both cis *(Z)* and trans *(E)* cinnamoyl salicortin are detected.

Appendix 5. Preliminary melatonin analysis in MYB165 overexpressing poplar

Table A5.1. Melatonin concentrations in Poplar lines.

Poplar Genotype	Melatonin concentration1 (ng · g FW^{-1})
MYB165 Overexpressor[2]	30.69
Wild-type 353[3]	18.67

[1] Analysis and quantification conducted through established methods by Erland et al. (2020a, 2020b). Data represents tissue from a single sample.

[2] Plants generated by Ma et a., (2018)

[3] *P. tremula x P. tremuloides* research hybrid 353

Appendix 6. Expanded substrate preferences for recombinant UGT78M1

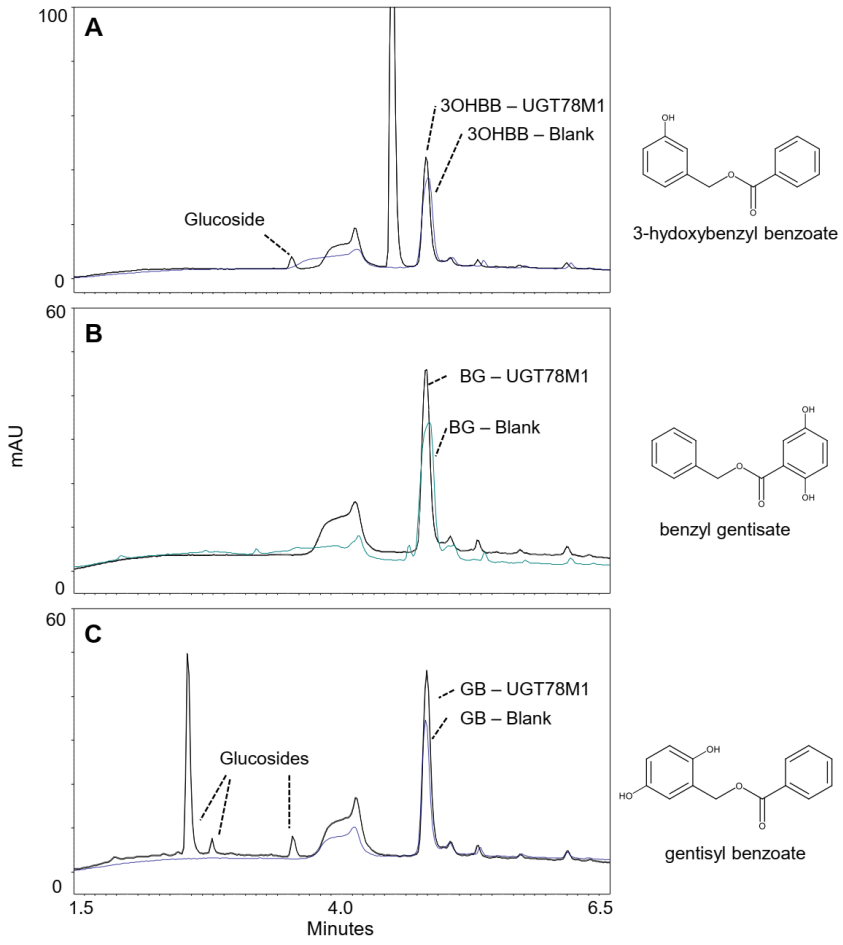


Figure A6.1. Chromatograms of UGT78M1 enzyme assays with dibenzenoid substrates. Chromatograms are HPLC-DAD traces observed at ʎ = 280 nm. Each is a reaction of a specified substrate (10 mM), with 5 mM UDP-glucose and 10 µg recombinant UGT78M1 in 50 mM of Sodium phosphate buffer (pH 7.4). The 200 µL reaction was incubated for 3 h at 30ºC, then halted with the addition of 20 µL 30% TCA as per Fellenberg et al (2020). A, trace of 3-hydroxybenzyl benzoate (3OHBB) incubated with UGT78M1 (black line) overlaid on a blank run containing no protein (blue dashed). The glucoside peak matches retention times consistent with monoglucosylated dibenzenoids (Gordon et al., 2022). B, benzyl gentisate (BG) reactions (black line)

overlaid with a protein free control (blue line). C, reaction of gentisyl benzoate (GB) incubated with UGT78M1 (black line) overlaid on a blank run containing no protein (blue dashed). The glucoside peaks match retention times consistent with monoglucosylated dibenzenoids, and monophenolic glycosides (Gordon et al., 2022). Structures of the aglycone substrate are presented to the right of the chromatograms.