THESIS

ANALYZING GENETIC RESPONSE MECHANISMS ASSOCIATED WITH COPPER HOMEOSTASIS IN POPULUS TRICHOCARPA USING A BIOINFORMATICS APPROACH

Submitted by

Eric Patterson

Department of Biology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2013

Master's Committee:

Advisor: Marinus Pilon

Patricia Bedinger
Courtney Jahn
Christina Walters

ABSTRACT

ANALYZING GENETIC RESPONSE MECHANISMS ASSOCIATED WITH COPPER HOMEOSTASIS IN POPULUS TRICHOCARPA USING A BIOINFORMATICS APPROACH

Copper is an essential micronutrient for plants and plays an important role in photosynthesis, respiration, hormone signaling, cell wall structure and wound healing. Copper deficiency can cause chlorosis, leaf curling, and weakened stems. It is proposed that under copper deficient conditions plants down regulate genes whose proteins use copper as a cofactor but also play an "unessential" role for the plants survival, thereby preserving copper for more "essential" proteins like plastocyanin or cytochrome-C oxidase. Down-regulation of "unessential" genes is performed by the copper microRNAs *miR307, miR398,* and *miR408*. This thesis increases our understanding of copper homeostasis in plants by analyzing the transcriptomic response of *Populus trichocarpa* to copper deficiency in four vegetative organs and applies this knowledge to the study of multi-copper oxidases. Organs have drastically different responses to copper deficiency with few genes being systemically differentially expressed and most genes that are differential expressed only are in one organ. Our data also show that not all genes are regulated to the same extent. Genes that are already highly expressed (>50 RPKM) under copper-sufficient conditions are only up-regulated 1- to 4-fold, while low expressed genes can be up-regulated as much as 8-fold. We go on to describe 25 unannotated genes as laccases based on their sequence similarity with known laccases from *Arabidopsis* and *Populus*. The laccases break up into seven phylogenetically distinct groups. Each of the seven groups have a distinct expression pattern across the four organs in response to copper deficiency that seems to be mediated by Cu-miRNAs *miR397* and *miR408*.

TABLE OF CONTENTS

CHAPTER 1

AN INTRODUCTION TO COPPER AND ITS ROLE IN PLANT BIOLOGY

Copper is the twenty-ninth most abundant element in the earth's crust and is an essential micronutrient for life (Linder and Goode, 1991). Copper has the ability to perform one-electron reduction/oxidation (Redox) reactions, making it indispensable for many of life's metabolic processes (Linder and Goode, 1991; Lippard and Berg, 1994). In plants, copper functions as a cofactor in a variety of metabolic proteins found in many organelles (Cohu and Pilon, 2010). In the chloroplast, copper is most important in the protein plastocyanin (PC), where it serves as an electron carrier in the electron transport chain (Katoh *et al.*, 1960). *Arabidopsis thaliana*, like most plants, has two PC isoforms, each of which requires a substantial amount of copper (Weigel *et al.*, 2003; Schubert *et al.*, 2002; Kieselbach *et al.*, 1998). In the mitochondrion, copper is required for activity of the cytochrome-c oxidase complex (COX), which reduces molecular oxygen to water (Carr and Winge, 2003). Copper also functions as a cofactor in a variety of plastidic, cytosolic, apoplastic and vacuolar enzymes, including copper/zinc superoxide dismutases (Cu/Zn SOD), polyphenol oxidases (PPO), amine oxidases, ascorbate oxidases, ethylene receptors and laccases (Lac) (Marschner, 2011) (Figure 1).

Copper's numerous functions in the plant cell make managing its abundance crucial. Copper deficiency can result in changes in root, stem, and leaf morphology as well as decreased photosynthetic activity (Marschner, 2011; Epstein and Bloom, 2005). The dangers of copper deficiency have led to the evolution of a complex copper uptake system as well as a system to allow for optimal copper distribution within the cell (see Burkhead *et al.*, 2009 for a review). Conversely, copper's reactive properties force plants to control the amount and distribution of copper to prevent the formation of Reactive Oxygen Species (ROS) (Halliwell and Gutteridge,

1984). Copper toxicity leads to inhibition of photosystem II as well as inhibition of chlorophyll synthesis, both of which lead to a decrease in photosynthetic efficiency and an increase in ROS (Yruela *et al.,* 1996; Bernal *et al.,* 2004). These dangers are avoided by a chaperone system for copper as well as storage of extra copper in organelles such as the vacuole.

Copper ions in soil exist primarily as Cu(II) but in plants they are primarily transported as Cu(I) (Marschner, 2011). Ions in the soil are thought to be reduced to Cu(I) by a cell surface ferric reductase (FRO), making the ions available for transport (Welch *et al.,* 1993)[1]. Copper uptake into the roots of the plant is performed by a copper transporter (COPT1) on the surface of root cells (Kampfenkel *et al.,* 1995; Sancenón *et al.,* 2003). The COPT are a multi-gene family (COPT1-6) present in both the plasma membrane and the tonoplast. COPT6 is primarily found in the vascular tissue plasma membranes and it interacts with COPT1. When COPT1 is knocked out COPT6 is up-regulated most likely to compensate for lack of COPT1 function (Jung *et al.,* 2012). COPT5 is important for copper efflux to the vacule (Garcia-Molina *et al.,* 2011). However, COPT4 is inactive because of a deletion of an essential methionine residue (Sancenón *et al.,* 2003). COPT3's complete expression pattern and subcellular localization have not yet been fully investigated (see Burkhead *et al.,* 2009; Ravet *et al.,* 2011; Pilon, 2011 for a review). In addition to the COPT transporters, ATP-independent ZIP (ZRT, IRT-like Protein) transporters have also exhibited the ability to transport copper ions across the plasma membrane (Wintz *et al.,* 2003).

Copper ions are transported to the shoots apoplastically through the xylem. The ions are transported out of the root symplast, probably as Cu(I), by a heavy metal P-Type ATPase (HMA) (Andrés-Colás *et al.,* 2006). Once in the xylem, ions are thought to be bound to a metal chelator, such as nicotianamine, for long distance transport (Briat, Curie, and Gaymard, 2007; Pich and

[1] This use of FRO to reduce Cu suggests a link between iron and copper homeostasis.

Scholz, 1996). Copper is then taken into the symplast of the shoots by COPT transporters (Sancenón *et al.*, 2003; Wintz *et al.*, 2003). Once inside the photosynthetic cells, copper can be transported to specific subcellular compartments via a variety of mechanisms, most importantly by the P-type ATPases which mediate transport into organelles (S. Abdel-Ghany *et al.*, 2005; Andrés-Colás et al., 2006; Puig et al., 2007).

The demand for copper in green tissues is in constant flux, depending on light availability, water availability, the plant's developmental stage, and availability of copper in the soil (Burkhead *et al.*, 2009). Several strategies have arisen to adjust the supply of copper to meet the fluctuating demands. These strategies are controlled by the transcription factor SPL7. SPL7 shares sequence similarity to the Cu Response Regulator in *Chlamydomonas*, CRR1, which regulates gene targets with an abundance of the sequence GTAC in their promoter region (Yamasaki *et al.*, 2009). This GTAC sequence is known as the Cu Response Element (CuRE). It is highly abundant in genes that are important for copper uptake and mobilization of copper. Several lines of evidence support the idea that CRR1 is able to sense copper abundance in *Chlamydomonas* (Kropat *et al.*, 2005).

Genetic evidence suggests that SPL7 is important for a plant's ability to respond to variations in copper supply (Cardon *et al.*, 1999; Yamasaki *et al.*, 2009). In plants, under copper depleted conditions, genes with a higher than average number of CuRE motifs are up-regulated by SPL7 (Yamasaki *et al.*, 2009). These genes include, but are not limited to, COPT1 and 2, FSD1[2], ZIP2, YSL2[3], FRO3, and the miRNAs *397*, *398*, and *408* (Yamasaki *et al.*, 2009). The miRNAs *397*, *398*, and *408* are known as the Cu-miRNAs because they are transcriptionally regulated by SPL7 in response to copper deficiency (Yamasaki *et al.*, 2007; Abdel-Ghany and

[2]Iron (ferrous) Superoxide Dismutase (FSD)
[3]Yellow Stripe Like (YSL)

Pilon, 2008; Burkhead *et al.*, 2009). They also have sequence complementarity to the mRNA of genes that encode copper-containing proteins, and are able to target those transcripts for degradation by the RNA-induced silencing complex (RISC). In *Arabidopsis* the Cu-miRNAs' targets include CSD1 and 2, plantacyanin, and many members of the laccase family (LAC2, 3, 4, 7, 12, 13, and 17) (S. E. Abdel-Ghany and Pilon, 2008).

The miRNA-mediated down-regulation of certain copper genes, combined with the up-regulation of copper transporters by SPL7, have led to the development of a fairly complex yet elegant model for plant copper homeostasis. When copper is sufficient within a cell, all copper-containing and copper-regulating proteins are expressed at normal levels. When copper is deficient, however, SPL7 is active, which in turn up-regulates copper uptake and transport genes like COPT1 and 2, FRO3 and ZIP2 (see Burkhead *et al.*, 2009 for a review). SPL7 also activates miRNAs *397*, *398*, and *408*, which down-regulate a subset of copper genes. This response is thought to increase the amount of copper being obtained by the plant, limit the number of proteins that demand copper, and conserve available copper for essential proteins such as plastocyanin and cytochrome-*c* oxidase (Figure 2).

Until now, work on copper homeostasis in plants has primarily been performed in *Arabidopsis thaliana.* While this has been a good model organism for discovering the fundamental pathways of copper homeostasis, we have begun new studies in *Populus trichocarpa* (black poplar). Poplar is a perennial woody dicot that grows relatively quickly and produces large quantities of biomass (including secondary cell walls and wood), making it an attractive model organism for studies related to biofuel production as well as paper production (Tuskan *et al.*, 2006).

The *Populus* genome was sequenced in 2003, and then partially annotated using software to predict open reading frames and protein products (Tuskan *et al.*, 2006). The genome is approximately 480 million base pairs, spanning 19 chromosomes. Modeling programs have reported 45,555 gene models using homology with known plant proteins and *ab initio* gene prediction. The size of the poplar genome is the result of two genome duplication events since its divergence from the *Arabidopsis* lineage 100-120 million years ago (Tuskan *et al.*, 2006). This increase in genome size has led to greater gene diversity than in *Arabidopsis*, as well as the presence of multiple poplar homologs for a single *Arabidopsis* gene. Genome duplication creates a two-fold problem: firstly, gene annotation cannot be taken wholesale from the *Arabidopsis* genome and applied to the *Populus* genome, and secondly, discovering the function of every gene is much more difficult due to possible functional redundancies. Fortunately, new tools and better processing power mean that this problem can be solved partly by bioinformatic techniques and partly by molecular techniques.

RNA-SEQ provides a high-throughput, transcriptome-wide, quantitative method for measuring gene expression, which can be applied in poorly annotated organisms like *Populus trichocarpa* (Wang, Gerstein, and Snyder, 2009). RNA-SEQ has become possible because of the recent development of next-generation sequencing techniques such as Illumina, SOLiD, and 454. RNA-SEQ uses massive parallel sequencing techniques to sequence 100-200 bp fragments of RNA collected from a starting material, which are then assembled into entire transcriptomes (Figure 3). RNA-SEQ is an excellent tool for studying the complete transcriptomic response to stimuli like copper deficiency or pathogen infection, as well as for uncovering previously unexamined genes that are differentially expressed in response copper deficiency.

Scope of the Thesis

This thesis describes our new research into copper homeostasis in *Populus trichocarpa*. As a way to discover novel mechanisms of copper regulation, we set out to get a more complete understanding of poplar's transcriptomic response to copper deficiency in the vegetative organs by using RNA-SEQ. Chapter 2 describes how data from an RNA-SEQ experiment were analyzed, including plant-wide patterns and trends of differentially expressed genes, a functional analysis of differentially expressed genes using MapMan, and differences in the transcriptome between four vegetative organs. Chapter 3 describes a bioinformatics approach to understanding the laccase gene family in poplar. In this chapter we report the discovery of previously unannotated laccase genes and discuss Cu-miRNA down-regulation of this important group of copper-containing enzymes.

Copper-containing Proteins			
Name	Subcellular Location	Function (in *Arabidopsis*)	Remarks
Plastocyanin	Chloroplast (thylakoid)	Electron transport chain, photosynthesis	
Cytochrome C oxidase	Mitochondrion (inner membrane)	Electron transport chain, respiration	
Cu/Zn SOD	Cytosol (CSD1), stroma (CSD2), peroxisome (CSD3)	Superoxide dismutation	
Laccase	Apoplast	Cell wall modeling, polyphenol synthesis	Function is shown *in vitro*
Ethylene Receptors	Endoplasmic Reticulum	Ethylene sensing	
Ascorbate Oxidase	Apoplast		Function unclear, possibly salt tolerance
Amine Oxidase	Apoplast	Wound healing, pathogen response, cell wall differentiation	
Plantacyanin	Apoplast		Function unclear, possibly reproduction
Polyphenol oxidase	Chloroplast (Lumen)	Diphenol synthesis	

Copper Regulatory Proteins			
Name	Organ/Subcellular location	Function (in *Arabidopsis*)	Remarks
COPT	Plasma membrane, roots (COPT1); Plasma membrane, shoots (COPT2); Vacuolar or organellar (COPT3 and 5)	Copper uptake in the roots, copper uptake in the shoot symplast, transport of copper in vacuole	Subcellular localization is unclear for COPT3 and 5, but they are predicted to be vascular
FRO	Cell surface, roots (FRO2 and 3)	Reduction of soil copper from Cu(II) to Cu(I)	Suggests a possible link with Fe Homeostasis
ZIP	Roots (ZIP2), leaves (ZIP4)	Possible redundancy with COPT	ZIP transporters show altered expression under low copper and can complement crr1 knockouts in yeast
HMA	Roots and flowers (HMA5)	Export Cu(I) into the stem for long-distance transport	HMA5 and COPT1 are thought to transport in opposite directions across the plasma membrane
RAN1	Endoplasmic Reticulum	Delivery of copper to ethylene receptors	Also called HMA7
PAA	Chloroplast inner membrane (PAA1), Chloroplast thylakoid (PAA2)	Delivery of copper to plastocyanin	
CCS	Cytosol, chloroplast lumen	Copper chaperone for Cu/Zn SODs	
SPL7	Nuclear	Transcription factor, recognizes CuRE motifs, detects cellular copper abundance	Homolog of *Chlamydomonas* .CRR1 protein
YSL	Plasma membrane	Transport of iron and other metals associated with nicotianamine.	In rice, YSL transporters uptake iron associated with phytosiderophores

Figure 1: Functions and cellular locations of important copper proteins in *Arabidopsis*

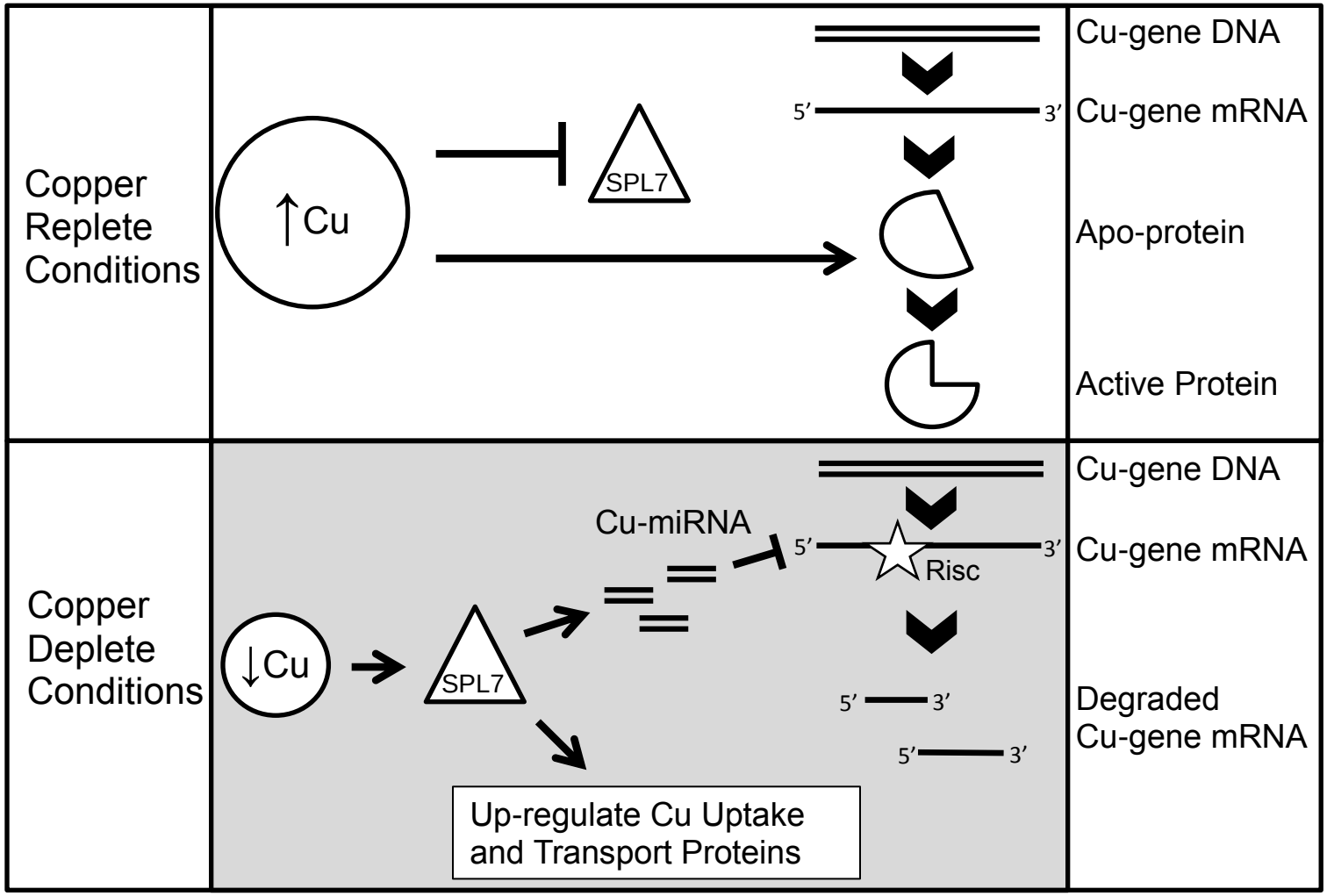

Figure 2: A model for Copper (Cu)-miRNA mediated regulation of copper-containing proteins. SPL7; Squamosa Promoter Binding Like transcription factor. Risc; RNA-Induced Silencing Complex

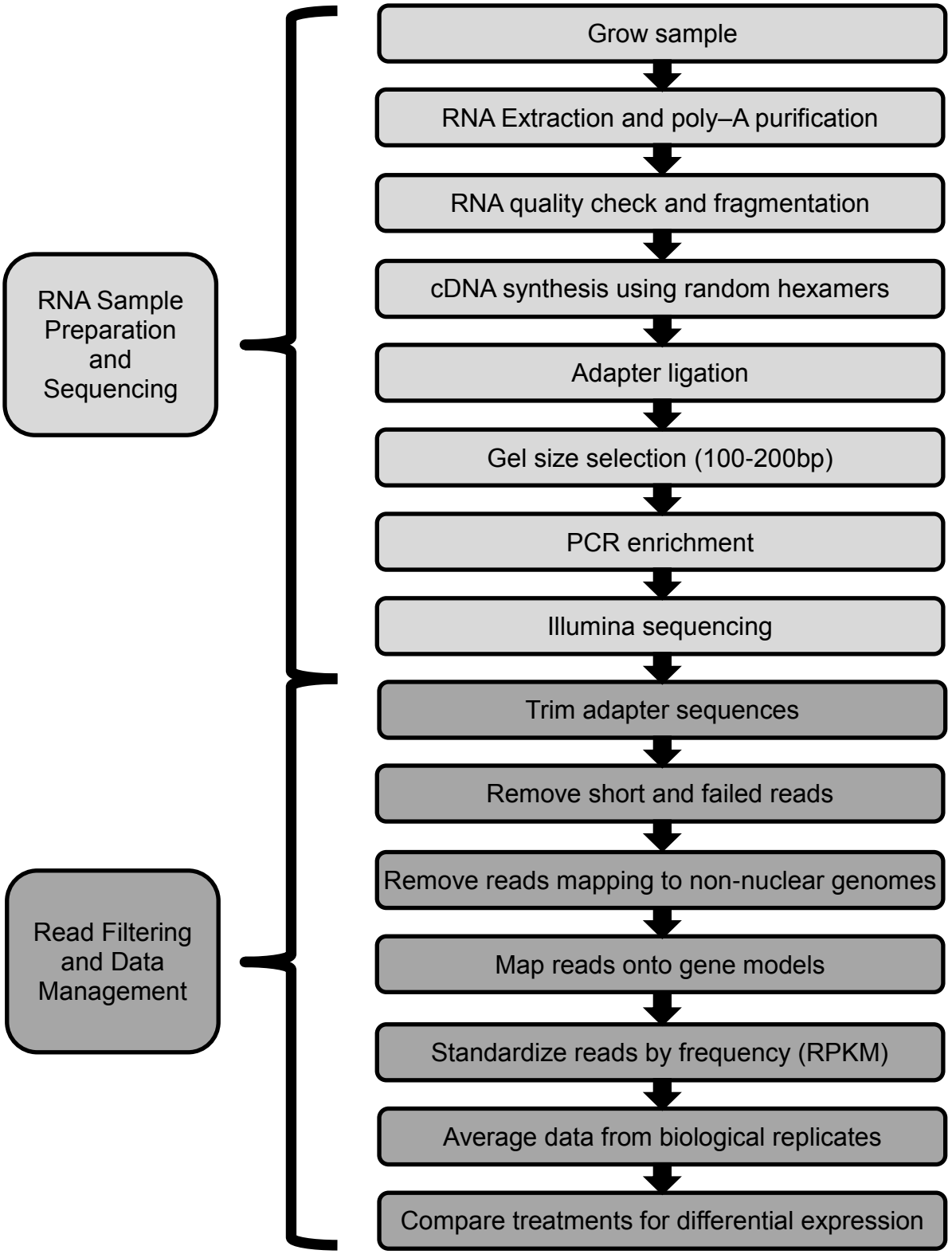

Figure 3: A flow chart of RNA-SEQ, RNA preparation and post-sequencing read handling.

CHAPTER 2

AN ANALYSIS OF THE EFFECT OF COPPER DEFICIENCY ON THE TRANSCRIPTOME OF POPULUS TRICHOCARPA

SUMMARY

A majority of what is known about copper homeostasis has been discovered in *Arabidopsis thaliana*. While *Arabidopsis* has been an excellent model organism for the discovery of basic mechanisms of copper homeostasis, *Populus trichocarpa* offers a new, and possibly more complicated, understanding. Under copper-deficient conditions, poplar shows good spatiotemporal separation of symptoms that allows for high resolution in organ-specific experiments. In this chapter we analyze the data of an RNA-SEQ experiment concerning the effects of copper deficiency on four vegetative organs: young leaves, old leaves, stems and roots. This experiment uncovers general trends in transcriptome expression in copper-deficient conditions and indicates hitherto undiscovered candidate genes that may play a large role in copper homeostasis and copper deficiency stress response. This chapter also explains organ-specific responses to copper deficiency and expands our understanding of how copper homeostasis is regulated at the organ level. This work will guide future molecular and genetic experiments in discovering copper deficiency response mechanisms.

INTRODUCTION

Trees are an important natural resource for humans, providing lumber, fiber and fuel, and they create unique habitats for a diverse group of organisms. Forests cover nearly 4 billion hectares of earth and are shrinking every year due to increased demand for tree products and farmland (Food and Agriculture Orginization, 2007). Tree physiology is interesting for plant biologists because of their extensive secondary growth, their ability to transport water and

nutrients over a relatively large distance, and the immense amount of resources required to grow and maintain the size of the organism (Tuskan *et al.*, 2006). Tree physiology can be greatly impacted by copper deficiency, with symptoms ranging from reduced and distorted growth to decreased and chlorotic foliage (Marschner, 2011; Ruiter, 1969). This study aims to describe the transcriptomic response to copper deficiency of a model tree, *Populus trichocarpa.*

The sequencing of *Populus trichocarpa* (Black Poplar) has allowed it to become the most important model organism for tree biology. *Populus* is a woody perennial dicot that grows relatively quickly (Tuskan *et al.*, 2006). Poplar has a number of advantages for the study of copper homeostasis due to its size and tractability. First, copper can be removed and resupplied to poplar in a controlled manner because it grows well in hydroponics systems. Hydroponics also allows for the control of all nutrients, to ensure copper is the only limiting nutrient. Second, when copper deprived, poplar exhibits very good spatiotemporal resolution of copper deficiency symptoms between organs (Ravet *et al.*, 2011).

The *Populus* genome contains approximately 41,000 individual genes with 45,000 gene models. The size of the genome is due to two genome duplications that occurred since its divergence from the herbaceous dicots (which include *Arabidopsis thaliana*) 100-120 million years ago. These duplications led to expanded gene families containing large numbers of paralogous genes that may have redundant functions (Tuskan *et al.*, 2006). Functional redundancy and difficulties with transforming poplar make forward and reverse genetic experiments onerous, and determining a single gene's function conclusively is often not possible. With the development of next-generation sequencing, whole-transcriptomic experiments are becoming more manageable, making a more targeted approach to gene-by-gene studies possible.

RNA-SEQ is the one of the newest and most robust kinds of whole-transcriptomic experiment. RNA-SEQ quantitatively determines the relative abundance of all transcripts within a sample and allows for the quantitative comparison of transcript abundance between samples. RNA-SEQ is an attractive research tool for studying copper homeostasis because it can provide information on the composition and quantity of the transcripts in the transcriptome of a sample, and about how gene expression responds to copper deficiency (Wang *et al.*, 2009).

The aim of this portion of my thesis is to expand our knowledge of copper homeostasis by analyzing a transcriptomic experiment in *Populus trichocarpa*. I will begin by discussing the experimental design and the initial experiments to confirm copper depletion in the hydroponics system, followed by the RNA preparation and sequencing, and lastly the RNA-SEQ data annotation and analysis.

METHODS

Experimental Design

This RNA-SEQ experiment was performed using *Populus trichocarpa* (cultivar Nisqually-1), grown in a hydroponic system (Ravet et al., 2011). Poplar plants were propagated by cutting the last 3-4 inches of mature stems of soil-grown plants with a razor blade. Excess leaf material was removed until only ~1 cm^2 of leaf area remained on a cutting, to minimize water loss by transpiration. The cuts were made below a node, where adventitious roots easily arise. The nodes of each cutting were then dipped in the rooting compound Clonex®, which contains 3 g/L of the rooting hormone Indole-3-butyric acid (IBA), for 1 hour, during which the cuttings were kept in the dark in 100% humidity to minimize transpiration during hormone application. The cuttings were then transferred to a medium consisting of coarse-grained vermiculite saturated with double distilled H_2O to minimize Cu absorption. The cuttings were

covered (100% humidity) and placed under low light (~150 μE) until adventitious roots formed (about 2 weeks). The cuttings were slowly introduced to ambient humidity by punching small holes in the coverings. After the roots were sufficiently developed (the first root hairs had begun to form), the cuttings were transferred to a hydroponics system.

The hydroponics system was a series of 5 gallon buckets filled with $1/10^{th}$ strength Hoagland's solution, modified to include varying amounts of copper sulfate. Plants grown in sufficient copper received 50 nM (final concentration) of copper sulfate in their media, while plants grown in copper deficient conditions were grown without the addition of any copper sulfate to the Hoagland's solution. Since copper is essential, plants could not live without some copper. However, all the cuttings contained some copper before they were put in the hydroponics system because they came from plants grown in soil. It is also likely that the copper-deficient Hoagland's solution still has a minuscule quantity of copper ions present, since it is nearly impossible to remove all copper ions without a chelator. The plants in both the sufficient and deficient conditions were grown for 5 weeks in ~200 μE of light on a long-day cycle until harvest. To ensure statistical strength, at least three unique but genetically identical cuttings were grown in each copper condition.

After 5 weeks of growth in hydroponics, samples were harvested from the plants and immediately frozen in liquid nitrogen. From each plant four organs were harvested: roots, stems, old leaves and young leaves. Old leaves were defined as the three oldest leaves, while the young leaves were the three newest leaves. A differentiation was made between the old and young leaves because copper is fairly immobile, and leaves that were formed when the plant was a cutting (i.e., mature leaves) still had locally sufficient amounts of copper, but leaves formed after the cuttings were placed in hydroponics showed strong copper-deficiency symptoms (leaves

were slightly chlorotic and curled). After being frozen in liquid nitrogen, samples were ground to a fine powder using a pre-chilled mortar and pestle.

The ground-up organ samples were then divided into three pools: one used in Inductively Coupled Plasma – Atomic Emission Spectroscopy (ICP-AES), to determine the number of copper ions in the organs; one used in Western blots, confirming copper regulation of known copper proteins; and the last for RNA extraction and sequencing.

RNA-SEQ Validation

ICP-AES was used to determine the relative abundance of copper and other ions in the various samples. Ravet *et al.*, 2011, performed this work on the samples that were sent for sequencing, and their methodology is described here. Samples were washed for 10 minutes in double distilled (dd) H_2O, 20 minutes in 40 mM EDTA, and then rinsed again for 10 minutes in ddH_2O. These samples were then dried for 3 days in a 55°C oven. One hundred milligrams of each sample was weighed into separate test tubes and mixed with 1 mL of 70% nitric acid. The samples were incubated for 2 hours at 60°C, followed by 6 hours at 130°C, with a glass funnel on top to prevent the evaporation of nitric acid. Once all organic material dissolved in the acid, the samples were diluted 10x with ddH_2O. Ion type and abundance were then determined using an ICP-AES calibrated with iron and copper ion standards.

Poplar plants grown in copper-deficient conditions for 5 weeks in hydroponics began showing the classical symptoms of copper deficiency at 3 weeks, such as stem bending, leaf curling, and chlorosis between the veins of the leaf. Copper deficiency in the plants was quantified after 5 weeks by ICP-AES by Ravet *et al.* in 2011, to ensure that there were decreased amounts of copper in all four organs compared to plants grown in copper-sufficient conditions and that the copper deficiency threshold was achieved ($<5\ \mu g\ g^{-1}$) (Marschner, 2011). In old

leaves, young leaves, and stems, copper levels were decreased in low copper conditions by nearly half (~10 µg g^{-1} to ~5 µg g^{-1}). Copper deficiency is described as anything below 5 µg g^{-1} copper, because most plants typically begin showing the symptoms of copper deficiency at or below this concentration. Old leaves did not show chlorosis or leaf curling to nearly the extent of young leaves. This is most likely because copper is not remobilized from the leaves. The old leaves developed before the plant was put in copper-deficient conditions, therefore, there was copper inherent in the old leaves. Roots were not nearly as affected as the leaves and stem, maintaining a high amount of copper (>5 µg g^{-1}) in their tissues under copper-deficient conditions; however, they did show decreased copper content compared to the copper-sufficient conditions (from 18 µg g^{-1} to 8 µg g^{-1}). Although the apparent Cu ion concentration in the roots is sufficient, it is likely that many Cu ions are apoplastic in the roots, and much of the copper detected in ICP-AES is not inside the plant cells. Iron, nickel, and magnesium were also measured and showed no change in their concentration under copper deficiency, as expected. Zinc and manganese, however, did show elevated concentrations in the leaves under copper deficient conditions. See Ravet *et al.*, 2011 for ICP-AES showing down regulation of Cu SODs and PC.

RNA-Extraction

RNA was purified from 100 mg (fresh weight) of ground tissue using the Invitrogen RNeasy® Plant Mini Kit (Invitrogen RNeasy Mini Handbook, 2012). RNA purity and concentration were checked by Nanodrop spectrophotometer readings. RNA samples of five □g each were stored overnight in a -80°C freezer and shipped directly to the University of Missouri on dry ice for Illumina sequencing.

RNA Sequencing

All RNA sequencing was performed at the University of Missouri DNA Core Facility (http://biotech.missouri.edu/dnacore/). RNA integrity was checked using the Agilent Bioanalyzer 2100, a fluorescence-based electrophoresis system that determines the length, quantity, and quality of the samples. The mRNA was purified from other RNAs by oligo-dT purification beads, ethanol washing and magnetic separation. First and second strand synthesis were performed using a random hexamer mix. The double-stranded cDNA was then fragmented and overhangs resulting from fragmentation were converted into blunt ends using an End Repair Mix®. The 3' to 5' exonuclease activity of this mix removes the 3' overhangs and the polymerase activity fills in the 5' overhangs. A single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent the fragments from ligating to one another during the adapter ligation reaction. Unique adapters (8 different sets for multiplexing, 1 set per sample, 24 sets total) were ligated to the double-stranded cDNA fragments. Fragments with adapters attached were enriched using PCR and primers that recognize the adapter sequence. Fragments were again checked for quality and quantified using the Agilent Bioanalyzer 2100 system. Eight cDNA libraries at a time were pooled, normalized, and run on the sequencer in a single lane[4].

Sequencing of the cDNA libraries was performed using the Illumina HiSeq 2000 ultra-high-throughput DNA sequencing platform. This platform used solid flow cells with oligos ligated to their surfaces. The cDNA fragments were washed over the surface and bound to the flow cell oligos. Each bound fragment was then amplified by PCR so that clusters of identical transcripts were generated on the solid cell. These clusters were large enough to emit a

[4] (Low Sample RNA-SEQ protocol, starting on page 39 of the TruSEQ RNA preparation guide, 2012)

detectable signal during sequencing. The reverse strands generated during the PCR were cleaved and washed away so only forward strands remained. The ends of the forward strands were blocked to prevent degradation and a sequencing primer was then hybridized to the base of the forward strands. All clusters were then washed with fluorescently labeled single nucleotides and polymerase. The nucleotides were bound to a blocking group to prevent polymerase from adding more than one nucleotide at a time. Between nucleotide washes, the fluorescent labels were excited, read by a spectrophotometer and then removed for the next cycle. All clusters were read simultaneously, with one cluster being the equivalent of one read.

RNA-SEQ Read Cleaning

An entire cluster of identical cDNAs attached to the solid matrix corresponds to a single "read". Reads were cleaned at the University of Missouri, Informatics Research Core Facility (http://ircf.rnet.missouri.edu:8000/) through a series of computational steps. Reads with ends ending in unknown nucleotides (NNNNNNN….) were trimmed to the last reliably determined nucleotide. All short reads (<50 bp) and failed reads (reads with high interior 'N' counts) were immediately removed. The adapter sequence was trimmed from all reads, and reads that were now less than 50 bp were removed. The reads were then aligned to the *Populus* genome, mitochondria libraries, plastid libraries and rRNA libraries by massively parallel BLAST searches. Reads blasting to the non-nuclear genomes were removed from the final read pool. Reads blasting to the *Populus* nuclear genome were annotated with their ENSEMBL gene ID for differential expression analysis.

Calculating Differential Expression

Reads were quantified as the number of Reads Per Kilobase of exon per Million reads in the sample (RPKM). This method standardizes transcript abundance by two things: 1) the length

of the mature mRNA and 2) the total number of reads generated from a sample. Firstly, long mRNAs will have more reads then a short mRNA because a single transcript will be fractionated into more 100bp fragments. As such, genes with long mRNAs are over-represented unless they are standardized. Secondly, not all samples generate the same number of total reads. Samples with a higher number of total reads will over-represent all transcripts when two samples are compared, and therefore each gene needs to be standardized by the total number of reads in the sample.

An RPKM was calculated for every gene in the *Populus* transcriptome for each sample, and this value represents the relative transcript abundance for that gene. RPKM data from the three replicates were averaged and these averages were compared with all other organ types and treatments using a statistical T-test (28 pairwise comparisons for each gene). Significantly differentially expressed genes (threshold p-value = 0.02) were then compiled and annotated for further bioinformatics experiments.

Data Annotation

All RPKM data were annotated using FASTA word files and Excel. Two Excel spreadsheet files were returned to us from the University of Missouri Informatics Research Core Facility (IRCF). The first spreadsheet contained every predicted gene in the *Populus trichocarpa* genome (v 2.2) and its RPKM in every sample (4 organs X 2 copper conditions X 3 biological replicates = 24 RPKMs per gene). The second spreadsheet contained only genes with at least one significant change in expression in a pairwise comparison of RPKMs between samples (p-value < 0.02). For the most part these were superfluous comparisons, and we removed all comparisons that were not between an RPKM in the sufficient condition and the RPKM in the deficient condition in the same organ.

A master spreadsheet was generated. For every predicted gene in the *Populus* genome, we averaged the RPKM values for biological replicates. We then calculated the fold-change in RPKM value for each gene between the copper sufficient and copper deficient conditions in each organ (fold-change = $\log_2 (RPKM_{def} / RPKM_{suf})$). Finally, if the gene had any significant p-values (< 0.02) for any comparisons between the conditions, the p-value was included in the master spreadsheet.

Genes were then annotated using a variety of databases. Partial gene annotation was performed by IRCF with Ensembl Plant database gene IDs and common names (http://plants.ensembl.org/) and InterPro database protein domain information (http://www.ebi.ac.uk/interpro/). *Populus* genes were further annotated using *Arabidopsis* homologs, as previously determined by the Plant Genomic Database, in the '*Populus* Annotation' file (http://www.plantgdb.org). The complete CDS and protein sequences were obtained from the 'Transcript' and 'Peptide' files from the Plant Genomic Database. These three files, 'Annotation', 'Transcript', and 'Peptide', were searched using the Ensembl gene id for every gene found in the master spreadsheet, and the *Arabidopsis* homolog, CDS, and predicted protein sequence were added to the master spreadsheet. Finally, all genes were annotated with their probe-set ID from a spreadsheet available from Affymetrix (http://www.affymetrix.com/) for MapMan analysis.

MapMan

Gene transcripts were functionally annotated using MapMan's ontology tool (Thimm *et al.*, 2004). To generate maps, MapMan requires three files: the experiment, the map, and the pathway. In the experiment file, gene IDs were replaced with their Affymetrix probe-set ID. If a gene did not have an Affymetrix probe-set ID, then the AGI number of the *Arabidopsis* homolog

was used. Some genes had neither an Affymetrix probe-set ID nor an *Arabidopsis* homolog. These gene IDs were combined with their respective fold-change information for each organ to complete the experiment file.

MapMan divides genes into 36 functional groups called BINs. The BINs are broad categories that contain from tens to hundreds of genes. Map files can be downloaded from MapMan (http://mapman.gabipd.org) to translate gene IDs (either an Affymetrix probe-set ID or an AGI number) into a BIN number for mapping. Genes that did not have a probe-set ID or an AGI number were included in MapMan's 36th BIN, "Miscellaneous". We combined the Affymetrix *Populus* map file with the *Arabidopsis thaliana* map file to be able to map genes with Affymetrix probe-set IDs and genes with AGI numbers in the same map. The combined list of gene IDs along with their respective BIN numbers formed the complete map file.

All pathway files were downloaded directly from the MapMan website and no modifications were necessary. These three files, experiment, map, and pathway, were imported into the MapMan comparison software, available from the MapMan website, to make a graphical representation of the fold-change for each significantly differentially expressed gene overlaid onto a diagram of a plant cell's pathways.

RESULTS AND CONCLUSIONS

Patterns of Differential Expression

The *Populus* genome has 40,183 predicted genes, not including splice variants. Splice variants were ignored for this analysis because of difficulties in aligning a read to a particular splice variant. Our experiment consisted of 24 samples: 2 conditions (50 mM and 0 mM $CuSO_4$), 4 organs (young leaves, old leaves, stems, and roots), and three biological replicates. A read library was generated from each sample using Illumina sequencing. The resulting libraries

ranged from 23.4 million reads to 42.4 million reads before read cleaning. After trimming, filtering, and cleaning, the read libraries contained between 14.7 million reads and 33.7 million reads (Figure 4). The average RPKM of a gene is 18.5. Of the 40,183 genes, 6909 (17.2%) showed significant (p-value < 0.02) differential expression in at least one organ, leaving 33,274 (82.8%) with no significant differential expression under copper deficiency.

In this experiment, there are eighty-one expression patterns a gene can exhibit between copper-sufficient and copper-deficient conditions: increased, decreased, or unchanged expression levels in each of the four organs we looked at (young leaves, old leaves, stem, and roots). These eighty-one patterns were each assigned a number, and the number of genes that exhibited each pattern was plotted (Figure 5). Genes with unchanged expression in all organs (pattern 41) were not included. Seven of the top eight patterns involve organ-specific regulation with no significant differential expression in any other organ. Interestingly, pattern 14 (genes down-regulated in roots in copper-deficient conditions with no change in any other organ) had the greatest number of genes. Nearly 25.3% of all significantly differentially expressed genes (1746 genes) show this pattern of expression. The reciprocal, pattern 68 (genes up-regulated in roots in copper-deficient conditions with no change in any other organ), had the second highest number of genes, with another 10.8% of the differentially expressed genes (746 genes). These two patterns constitute 36.1% of all differentially expressed genes. Furthermore, very few genes are differentially expressed in the same way in all four organs. Only 40 genes are down-regulated in every organ (pattern 1), while 77 genes are up-regulated in every organ (pattern 81). This observation is biased because in some organs, certain genes are not expressed even under copper sufficiency, so they cannot possibly be down-regulated in those organs. This bias means there is

only a small systemic plant transcriptome response to copper deficiency, but an individual organ's transcriptome response can be quite large as well as unique.

Plant-Wide Trends in Differential Expression

Every significantly differentially expressed gene in the *Populus* transcriptome was plotted to compare its log fold-change in expression level from the copper-sufficient to the copper-deficient condition (Figures 6 and 7). There is a trend that highly expressed genes (>50 RPKM), such as light harvesting complexes, have small changes in expression, while low-expressed genes, like SULTR3/5, seem to have large changes in expression between sufficient and deficient conditions. A large change in expression for a low-expressed gene can be a matter of going from 2 RPKM to 8 RPKM, which is a four-fold change. Although it is tempting to discount this change as due to variability in the expression data, the variability is accounted for by the standardization of the number of reads into RPKM, as well as by averaging the 3 biological replicates, which are both taken into account in our confidence measurements (p-values). Therefore, even though the actual increase in the number of reads for a low-expressed gene is small, the statistical analysis ensures that this is a statistically significant change in expression. However, statistical significance is not equivalent to biological significance, giving rise to the following speculation: a large fold-change in a low-expressed gene is not as biologically significant as a small fold-change in a high-expressed gene.

The most highly expressed gene that is up-regulated in copper-deficient conditions is annotated as a Mercury or Heavy Metal scavenger (HMA) (Figure 6). Both annotations are used because the mercury scavengers do not always target mercury specifically, and many have an affinity for all heavy metals (Fe, Cu, Hg, etc.) or a specific heavy metal other than mercury (Dykema *et al.*, 1999). An HMA gene is highly expressed in all the organs, but is only

significantly up-regulated under copper-deficient conditions in the roots, stems, and old leaves. The closest *Arabidopsis* homolog of this gene is ATX-1, a copper chaperone that delivers Cu to PAA1 (Puig *et al.*, 2007). The HMAs are metal-binding proteins that function in metal transport and metal detoxification. Most likely these Hg-scavengers are functioning as metal-transporting proteins regulating the amount of copper and iron ions in the cell. The Hg-scavenger protein family is large, especially in *Populus*, and only a few of these genes respond to copper deficiency. Given their general proposed role in metal homeostasis, these genes are of interest as candidate genes for further study because of their potential role in copper homeostasis.

The highest fold-change in expression is a sulfur transporter in the young leaves (nearly 8 fold). Sulfur assimilation is important for the production of glutathione. Glutathione is a small tri-peptide created from cysteine, glutamate, and glycine. Glutathione functions as an antioxidant, scavenging free radicals and peroxides and reducing the potential for ROS damage to the cell's macromolecules. ROS may be generated under copper-deficient and copper-toxic conditions, primarily from the disruption of important processes such as photosynthesis and respiration. The link between sulfur assimilation and copper abundance primarily hinges on glutathione production. The relationship between sulfur and copper homeostasis remains largely unexplored and our results here suggest there may be greater interaction between the two homeostasis pathways then has been previously suggested.

Genes down-regulated under copper deficiency show a similar pattern to genes up-regulated under copper deficiency (Figure 7). For the most part, highly expressed genes do not exhibit a large fold-change in expression. In the roots, however, a large number of the highly expressed (50-5000 RPKM) genes have dramatically decreased expression (fold-change of -2.5) in copper-deficient conditions (179 genes, see box in figure 7). Not surprisingly, 59 of these

genes are involved with photosynthesis, including the chlorophyll synthesis genes, photosystem I and II subunits, and many members of the thylakoid electron transport chain. Without this organ-specific, transcriptome-wide experiment, this root-specific phenomenon would have gone unnoticed.

Functional Analysis of Differentially Expressed Genes Based on MapMan BINs

In almost every MapMan BIN, roots have a significantly higher number of down-regulated genes compared to the other organs, but a nearly equivalent number of up-regulated genes (Figures 8 and 9). The most significant example of this trend is the photosynthesis BIN. Down-regulation of photosynthetic genes under copper deficiency is explained by the copper homeostasis model. These genes seem to be "non-essential" in roots, and therefore, under copper deficiency, decreasing their expression saves important copper resources for other "essential" copper-containing proteins.

It is counterintuitive that photosynthetic genes should be expressed in the roots at all, but there are many possible reasons for this. In *Arabidopsis*, photosynthetic genes are expressed in newly formed root tissue, especially at the root tip or in roots close to the soil surface (Sawchuk *et al*., 2008). The hypothesis is that these roots may be prepared to differentiate into new shoot tissue when exposed to light. The containers used in our hydroponics did not filter out 100% of the light, so the roots were exposed to low intensity light, which may have encouraged expression of photosynthetic genes on the root surface. Although the light also allowed for some growth of algae, since the read libraries were cleaned to include only reads mapping to the *Populus* genome, it is unlikely that these photosynthetic genes are from algae or cyanobacteria. Therefore, poplar roots most likely do expressing photosynthetic genes but these genes are highly down-regulated under copper deficient conditions.

Analysis of differential expression in each BIN that is based on the sheer number of differentially expressed genes can be misleading, since each BIN is composed of a different number of genes. To correct for this, the data were transformed by calculating what percentage of the total number of genes are differentially expressed in each BIN (Figure 9). For example, under copper-deficient conditions, between 18% and 25% (depending on the organ) of the sulfur assimilation genes are up-regulated. The "Sulfur Assimilation" BIN only has 16 total genes total, so a few differentially expressed genes can have a large impact in a small BIN. We see a similar trend in the "Metal Handling" BIN, with a comparatively large percentage (~10%) of the genes being up-regulated in this relatively small BIN (96 genes).

Unexpectedly, the "Stress" and "Redox" BINs do not show an overabundance of genes with differential expression when compared to the other MapMan BINs. The plants grown in copper-deficient conditions showed a decrease in chlorophyll amount, photosynthetic activity, and non-photochemical quenching (Ravet *et al.*, 2011). These conditions could lead to cellular stress and a concomitant increase in the expression of stress- and redox-related genes. While it is true that there was some differential expression of genes in these BINs, it was not more than the average number of genes differentially expressed across the genome (about 5-10%). There are several ways to explain this unexpected result. First, many stress and redox genes may not be annotated, and are therefore in the "Miscellaneous" BIN. Second, although copper deficiency does stress the plant, the stress may not have been great enough to cause differential expression of many "Stress" and "Redox" genes. Third, the "Stress" and "Redox" BINs are loosely defined, so the genes they contain may respond to many different stressors besides copper deficiency. Ravet *et al.* 2011 showed that poplar grown for only 5 weeks under copper-deficient conditions could recover (almost completely) if resupplied with sufficient copper. It is hypothesized that if

the plant could recover from the copper-deficient treatment, then it never became so copper starved that large secondary (stress) responses had occurred.

We see that ~17.2% of the *Populus* transcriptome is differentially expressed in copper-deficient conditions. Interestingly, we observe that plants respond to copper deficiency in an organ-specific manner. Few genes are systemically down-regulated and fewer still are systemically up-regulated. Ravet *et al.* 2011 showed that roots have the highest concentration of Cu ions in copper-deficient conditions, yet roots seem to be most sensitive to copper deficiency. Unfortunately, it is unclear whether copper is primarily apoplastic or symplastic when it is in the root tissue, however, the roots are thoroughly washed before ICP-AES so it is assumed that most apoplastic copper is removed. It could be that roots are the most effect by copper deficiency because they are the only organ importing copper and are regulating copper uptake and transport for the rest of the plant.

Differential Expression in the Organs

As we have seen, the different organs are regulating their transcriptomes for the most part independently. With this in mind, the various organs and their transcriptomes need to be individually analyzed. The questions I will attempt to answer are: what are the top genes that are the most differentially expressed under copper-deficient conditions in each organ, and can this tell us anything about how that organ is responding to copper deficiency? To answer the first question, the top ten genes were looked at in two ways. Firstly, a table was made for each organ of the ten most up-regulated and down-regulated genes, strictly determined by fold-change. Second, a table was made for each organ of the top ten genes with a fold-change of at least two, sorted by their RPKM in copper-deficient (up-regulated) or copper-sufficient conditions (down-regulated). This second table highlights genes that may have large biological impact because

they are highly expressed. As previously discussed, low expressed genes may have high changes in expression, but may not have a large biological impact. To compensate for this possibility, both tables will be taken into consideration for this study. To answer the second question, all differentially expressed genes were mapped onto a MapMan metabolic pathway figure in an attempt to take entire pathways into consideration.

Young leaves

Young leaves up-regulate 1637 genes and down-regulate 743 genes under copper deficiency. For both top ten lists, the young leaves down-regulate copper/zinc SODs, tyrosinases, a laccase, and a plantacyanin (Figures 10 and 11). These genes are all down-regulated 2-3 fold (1/4-1/8 the number of transcripts) in copper-deficient conditions. In *Arabidopsis*, plantacyanin and Cu/Zn SODs are heavily down-regulated under copper-deficient conditions, and both have been shown to be targets for Cu-miRNAs (Yamasaki *et al.*, 2007). This regulation has also been validated in *Populus* (Ravet *et al.*, 2011). The tyrosinase genes are polyphenol oxidases (see Mayer, 2006 for a review). These genes transcripts have also been shown to be down-regulated by Cu-miRNAs in poplar because their gene products have copper as a cofactor, and they are most likely involved in "non-essential" cell defense responses and pigmentation (Ravet *et al.*, 2011).

However, the up-regulated genes may prove more interesting than the down-regulated genes. None of the 10 most up-regulated genes in either list of differentially expressed genes has been identified as copper regulated, making all of them targets for investigation. The only confirmed mechanism for gene up-regulation under copper deficiency is the binding of SPL7 to cis-acting CuRE motifs in the promoter. If these genes are regulated in an SPL7-dependent manner, then we expect to find an overabundance of these motifs in the promoters of these

genes. However, there may be novel mechanisms of gene regulation at work. The questions still remain: why are these genes over-expressed in copper-deficient conditions, and how is their expression related to copper homeostasis?

Mapping all significantly differentially expressed genes from young leaves onto the various metabolic pathways allows us to visualize the RNA-SEQ data more completely (Figure 12). We see that most pathways have some up-regulated genes, with the exception of the light reactions and tetrapyrrole synthesis. For these two pathways, we see a nearly ubiquitous (although not very dramatic) down-regulation of genes. The light reaction genes consist primarily of plastocyanin, photosystem I and photosystem II subunits, and genes involved in the thylakoid electron transport chain. Tetrapyrrole synthesis primarily involves chlorophyll synthesis and light harvesting complex assembly. This is a small BIN, consisting of only 62 genes. Down-regulation of tetrapyrrole synthesis genes in young leaves helps to explain the chlorosis in copper starved plants.

Old leaves

Old leaves up-regulate 1071 genes and down-regulate 769 genes under copper deficiency. The top ten most differentially expressed genes in old leaves are slightly different than what appears in young leaves, although two of the genes in old leaves are also down-regulated in young leaves (Figures 13 and 14). The top ten down-regulated genes in the old leaves are, similarly to the young leaves, down-regulated 2-4 fold. Plantacyanin and many of the polyphenol oxidases are not as highly down-regulated in old leaves when compared to young leaves, although they are still significantly differentially expressed. Surprisingly, most of the down-regulated genes in both lists for old leaves have not been predicted as targets of Cu-mediated miRNA down-regulation, the exceptions being laccase, Cu/Zn SODs and tyrosinases.

The top ten most up-regulated genes for old leaves under copper deficiency in both lists are also distinct from the genes in young leaves. Many peroxidases and stress response genes are up-regulated, as shown in figures 13 and 14, most likely to mitigate damage from ROS. None of these genes have been indicated as potential SPL7 targets, and the method of their up-regulation remains a mystery.

When the differentially expressed genes in the old leaves are mapped onto the metabolic pathways, we see a slightly different pattern than in the young leaves (Figure 15). Very few genes in the light reaction and tetrapyrrole synthesis pathways are down-regulated. Instead, many genes involved in cell wall synthesis and phenol production are down-regulated. One possible explanation is that mature leaves have secondary cell walls that are in the process of lignification and thickening. Both polyphenol oxidases and laccases are enzymes that require copper as a cofactor. Polyphenol oxidases are herbivory defense genes, while laccases are implicated in the lignification of cell walls (Claus, 2004; Mayer, 2006). Both families contain known targets of copper-mediated miRNA degradation. To conserve copper, cell wall macromolecule synthesis (especially of lignin) may be down-regulated so that other "essential" genes can be expressed and their proteins can mature with the now available copper.

Stems

Stems up-regulate 1899 genes and down-regulate 1026 genes under copper deficiency. Five of the top ten most down-regulated genes by fold-change alone are either laccases or polyphenol oxidases (Figure 16). Figure 17 shows the same trend as Figure 16, but with a more diverse group of copper-containing proteins, including cupredoxins and Cu/Zn SODs. Other important enzymes in the lists include plastocyanin as well as a laccase 17 homolog. Plastocyanin is the most abundant copper-containing enzyme in the plant and absolutely essential

for photosynthesis (Yamasaki *et al.*, 2009). It is remarkable that this gene should be down-regulated. Most likely, PC is not being down-regulated by miRNAs, but instead directly by the copper availability in the cell.

The LAC17 homolog is also one of the top ten down-regulated genes in both the young leaves and the old leaves. This is interesting because not only is this gene down-regulated in all the organs, it is also among the top ten most down-regulated for the green tissues. It is not a very highly expressed gene, with RPKM values ranging from 0.5 to 6. It is possible that one of the Cu miRNAs has a high affinity for this gene's mRNA target site. An investigation of this mechanism can give insight into predicting other Cu miRNA targets, as well as why this regulation seems to be so tightly controlled. Besides the LAC17 homolog, there are also homologs for LAC2, LAC5, and LAC11 in the top ten most down-regulated genes in the stem (Figure 16). The strength of the regulation of laccases is unique to the stem and will be looked at in more depth in the next chapter.

The genes up-regulated in stems are unique. The fact that the COPT1 and FRO4 homologs are among the top ten up-regulated genes in the stem supports the idea that both the stem and the roots are involved in copper import into the symplast; they are, however, still low-expressed. The combination of these two proteins is thought to facilitate the majority of copper uptake into cells. FRO4 was recently characterized as a copper deficiency response gene, which suggests a link between copper and iron homeostasis in *Arabidopsis* (Bernal *et al.*, 2012). The FRO enzymes may act to reduce copper for transport by COPT1, which is the major importer of copper ions into the symplast.

More interesting still are the genes up-regulated with high expression (Figure 17). The top four of these genes are all nucleoside phosphorylases. In *Arabidopsis* these genes have been

assigned a variety of functions, from facilitating plant growth to pathogen defense (Ascencio-Ibáñez *et al.*, 2008; Irshad *et al.*, 2008). Most likely these genes are up-regulated and highly expressed to help manage oxidative stress (Charron *et al.*, 2008).

Mapping the significantly differentially expressed genes of the stems onto the metabolic pathways reveals the same pattern as in the old leaves (Figure 18). There is down-regulation of a significant portion of the genes implicated for phenol synthesis, as well as cell wall biogenesis. The stem-bending symptom under copper-deficient conditions may be explained by the down-regulation of these pathways, which seems likely to decrease the strength of the cell walls. The down-regulation of these pathways also helps to provide more support for the involvement of laccases and polyphenol oxidases in the generation of phenolic compounds, as well as implicating these phenolic compounds as an important structural component in cell walls. I will examine these hypotheses in more detail in the next chapter.

Roots

Roots up-regulate 1353 genes and down-regulate 3336 genes under copper deficiency. The number of down-regulated genes in roots is more than in the other three organs combined. In roots, the top ten down-regulated genes are quite different than in the other organs, and none of them have been predicted as Cu-miRNA targets (Figures 19 and 20). Most of the genes that are down-regulated and highly expressed are photosynthesis genes. This supports what we have seen in Figures 5, 6, and 7.

The top ten up-regulated genes are also unique, except for two FRO4 homologs that are also up-regulated in the stem. Since FRO4 seems to be involved in copper uptake, this up-regulation in the roots makes sense. We see co-regulation with another COPT1 homolog in Figure 19. The only other up-regulated genes of interest are three Myb-like transcription factors.

These transcription factors are worth further investigation because they may be new regulators of copper response genes.

Mapping significantly differentially expressed genes in the roots onto the metabolic pathways gives a unique result (Figure 21). We see an almost universal pattern of down-regulation of genes in every pathway. This observation supports the data shown in Figures 6 and 7. Roots seem to be the most sensitive to copper deficiency. This is nearly double the number of genes we see differentially expressed in the other organs. In the other organs, more genes are up-regulated than down-regulated in copper deficiency, but in the roots this trend is reversed. Explaining why the roots respond to copper deficiency so differently compared to the other organs will require more data. However, a possible explanation may lie in the fact that roots are the only non-photosynthetic organ, and they are the direct means for nutrient import. Since roots regulate nutrient uptake for the rest of the plant, it makes sense that they have the most dramatic response to a nutrient-related stress.

Conclusions

Given our understanding of the mechanisms of copper regulation of gene expression, the organ-specific differences in the patterns of gene regulation are surprising. We expected that the most important genes would have plant-wide changes in transcript abundance and that those genes would be among the top most differentially expressed genes in each organ. It is clear that this is not the case, and there are no genes that can necessarily be considered the most important. Instead, our data show that organs respond differently to copper deficiency based on local copper abundance (i.e. between young leaves and old leaves, or roots versus shoots) and the composition of the transcriptome before copper starvation.

This data set provides us with many new genes of interest. There are almost no predicted copper responsive genes among the most up-regulated genes under copper-deficient conditions. The only known mechanism of gene up-regulation in response to copper deficiency is the transcription factor SPL7. Up-regulation by SPL7 can be predicted by looking for copper responsive elements (CuRE) in the genes' promoters. We would like to look at the promoters of all up-regulated genes. However, this is 4179 genes. In *Populus*, defining and obtaining the promoter sequences for this many genes is not a trivial task, and is left for future work.

Genes down-regulated under copper-deficient conditions show many familiar genes and a few new candidate genes important for copper homeostasis. Genes in the CSD, LAC, PPO, and PC families are commonly among the most down-regulated genes in the green tissues. Most of these genes have already been shown to be targets of Cu-mediated miRNA down-regulation in *Arabidopsis* and *Populus* under copper-deficient conditions (Abdel-Ghany and Pilon, 2008; Ravet *et al.*, 2011; Yamasaki *et al*., 2007). Roots also show down-regulation of these predicted targets, but a large number of highly expressed photosynthesis genes are also being down-regulated. How these genes are down-regulated in the roots and the roots alone is not abundantly clear. Examining this phenomenon in the future may result in a novel mechanism of root-specific down-regulation.

Our data also show that not all genes are regulated to the same extent. Genes that are already highly expressed (>50 RPKM) under copper-sufficient conditions are only up-regulated 1- to 4-fold, while low expressed genes can be up-regulated as much as 8-fold. This trend also holds true for down-regulated genes, and to an even greater extent. The varying degree of regulation of genes should be examined more closely to determine if regulation by SPL7 and Cu-

miRNAs can be modulated based on target transcript abundance, or if there is some limiting factor in the amount of regulation that high expressed genes can undergo.

Populus is an emerging model organism; so many bioinformatics tools that facilitate RNA-SEQ data analysis are still being developed. Consequently, to perform the analyses shown in this thesis, the data had to be properly annotated using a variety of sources and then manipulated into the proper format required by each tool. This is a time-consuming process and future experiments will benefit from the curation done here.

The discoveries about copper homeostasis made here serve to emphasize the need to develop more complex homeostasis models that take into consideration local copper abundance, organ function, and the composition of the transcriptome in the various organs in copper sufficient conditions. This work leads directly to more detailed studies characterizing molecular and genetic response mechanisms to copper deficiency and opens hitherto unexamined paths of discovery in the forest of copper homeostasis.

Sample ID	Raw Reads	After Short Read Removal	After Low Quality Read Removal	After Adapter Trim, Followed by Short Read Removal	After Plastid Read Removal (clean reads)	Percent of Original Remaining	Mappable Reads	Percent of Clean Reads Mapped	Average reads per gene
YL 50_1	30,787,102	29,253,683	28,262,453	27,415,294	27,229,285	88.4%	26,360,735	96.8%	648
YL 50_2	32,685,504	31,113,338	30,080,779	29,208,635	29,025,300	88.8%	28,151,641	97.0%	692
YL 50_3	29,816,857	28,355,206	27,388,309	26,616,192	26,436,551	88.7%	25,594,946	96.8%	629
YL 0_1	28,469,995	26,628,892	25,673,251	24,953,323	15,507,582	54.5%	14,777,579	95.3%	363
YL 0_2	36,077,662	34,402,728	33,256,110	32,300,568	32,155,597	89.1%	30,995,202	96.4%	762
YL 0_3	27,736,780	26,262,026	25,312,604	24,519,399	21,593,701	77.9%	20,805,172	96.3%	512
OL 50_1	38,203,898	35,253,419	33,367,343	31,698,092	31,111,256	81.4%	29,940,725	96.2%	736
OL 50_2	31,331,888	28,955,948	27,415,987	26,013,859	25,855,222	82.5%	24,879,391	96.2%	612
OL 50_3	37,049,028	34,041,193	32,266,458	30,685,476	30,480,982	82.3%	29,477,234	96.7%	725
OL 0_1	37,808,498	34,882,898	33,109,001	31,501,519	30,975,946	81.9%	29,831,927	96.3%	733
OL 0_2	34,796,140	32,135,042	30,525,041	29,042,852	28,882,179	83.0%	27,806,325	96.3%	684
OL 0_3	32,599,057	29,590,925	28,013,182	26,580,227	26,432,454	81.1%	25,546,275	96.6%	628
S 50_1	33,980,609	31,078,902	29,379,203	27,832,615	27,763,655	81.7%	26,034,916	93.8%	640
S 50_2	32,772,714	30,385,323	28,821,283	27,342,407	27,263,022	83.2%	25,938,100	95.1%	638
S 50_3	38,022,499	34,878,580	33,046,814	31,355,964	31,210,861	82.1%	29,976,475	96.0%	737
S 0_1	34,998,296	31,977,412	30,194,335	28,581,454	28,481,397	81.4%	26,965,006	94.7%	663
S 0_2	41,061,127	36,776,542	34,665,022	32,808,078	32,719,032	79.7%	31,195,413	95.3%	767
S 0_3	32,100,457	29,499,869	27,886,203	26,367,460	26,251,895	81.8%	25,141,742	95.8%	618
R 50_1	38,572,074	36,002,993	34,447,528	33,177,860	33,045,889	85.7%	30,385,759	92.0%	747
R 50_2	23,416,623	22,031,628	21,096,669	20,329,174	20,245,882	86.5%	18,988,757	93.8%	467
R 50_3	42,440,859	39,787,072	38,102,404	36,695,598	36,501,098	86.0%	33,732,820	92.4%	829
R 0_1	38,113,949	35,587,522	34,055,516	32,767,923	32,647,639	85.7%	29,546,963	90.5%	726
R 0_2	27,197,217	25,147,661	24,035,912	23,114,007	22,921,057	84.3%	20,818,670	90.8%	512
R 0_3	31,057,608	29,048,836	27,795,909	26,731,924	26,293,169	84.7%	23,813,313	90.6%	585

Figure 4: A table depicting the how the RNA-SEQ sample read libraries was cleaned and the number of reads filtered at each step of the process

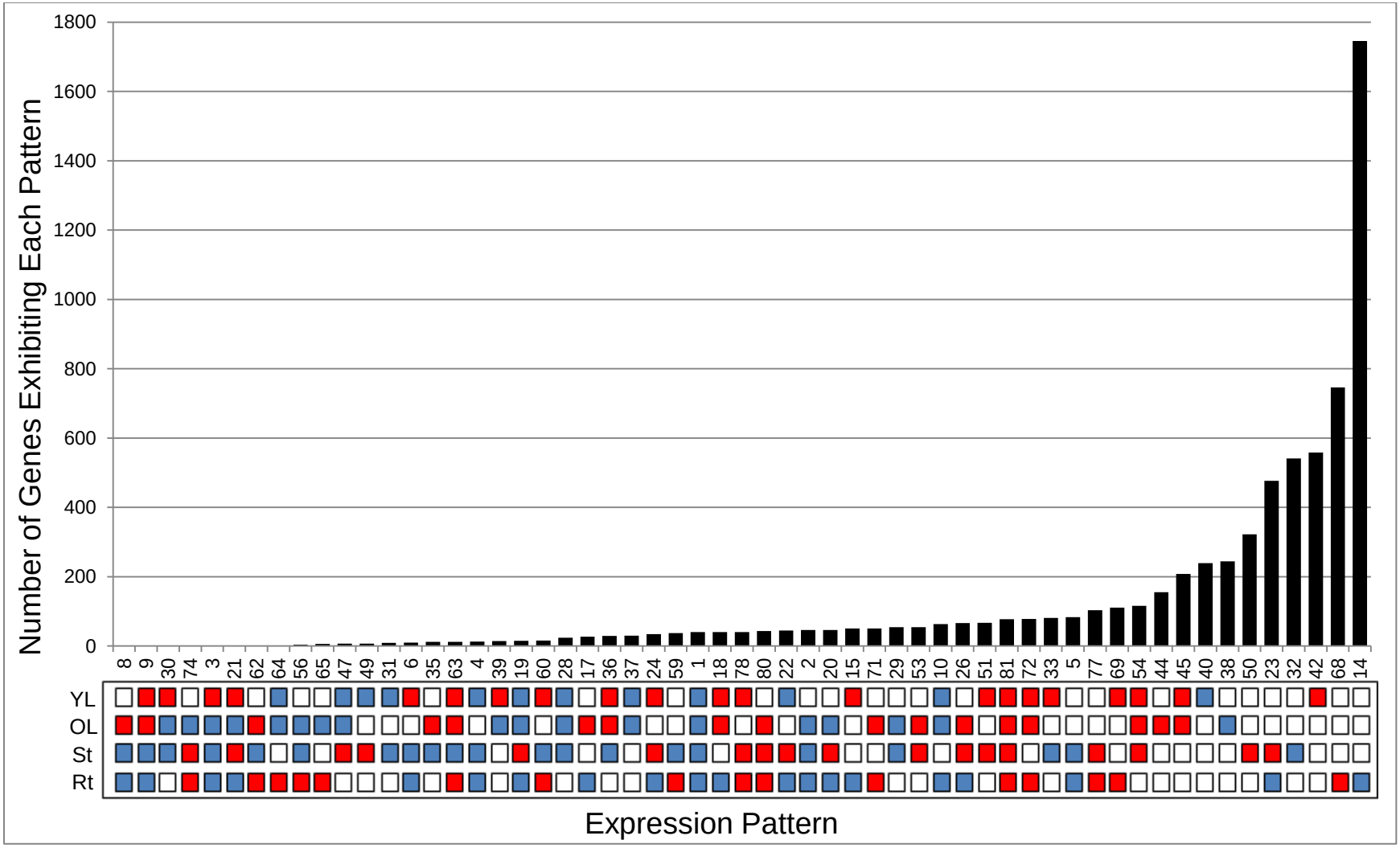

Figure 5: The number of genes that exhibit each of the 81 possible patterns of expression under copper-deficient conditions. Red squares represent up-regulation, blue squares represent down-regulation, and white squares represent no change. The squares correspond to the organs shown at the left of the key. YL: Young Leaves; OL: Old Leaves; St: Stem; Rt: Root.

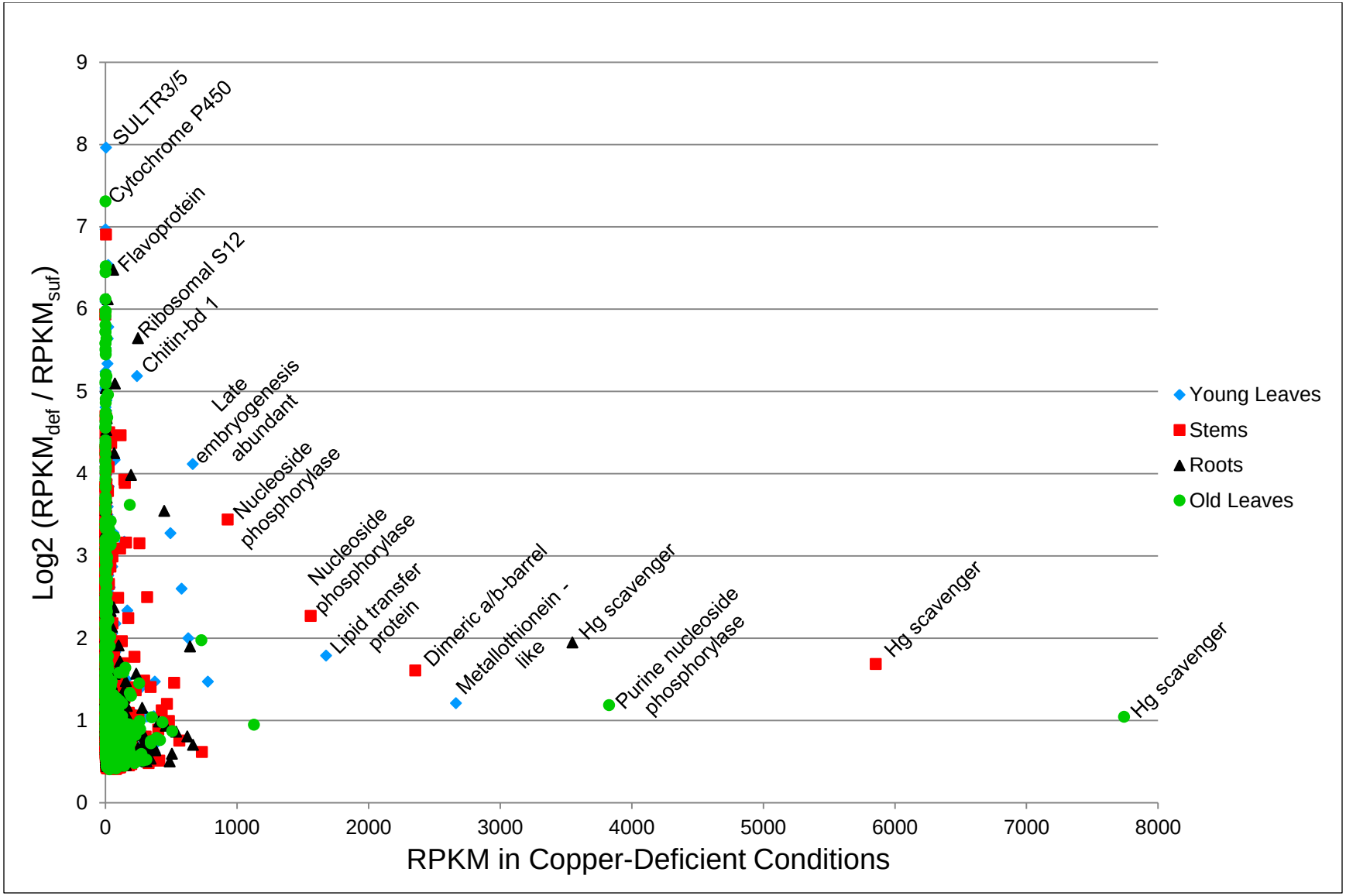

Figure 6: All genes significantly up-regulated under copper-deficient conditions.

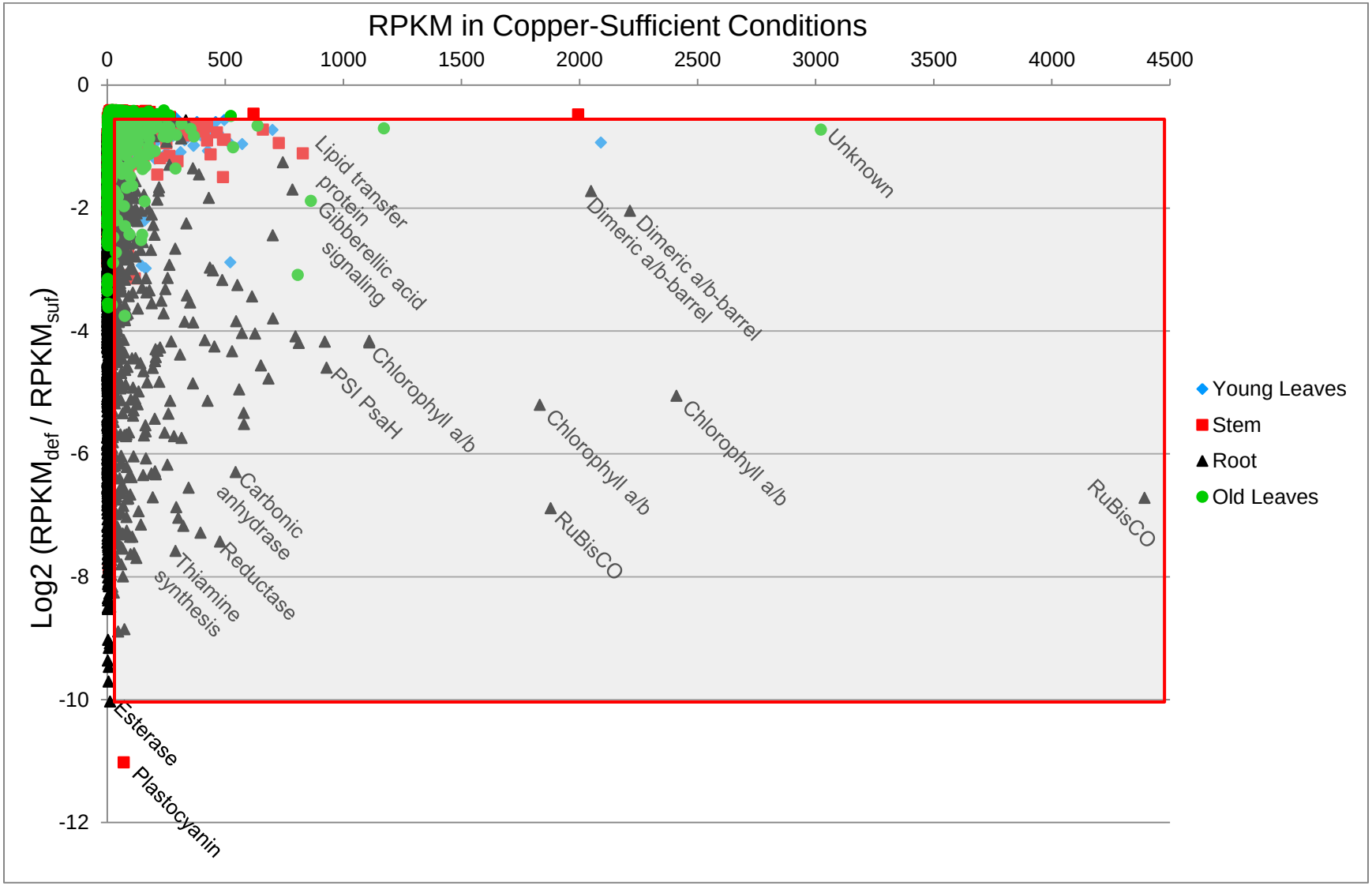

Figure 7: All genes significantly down-regulated under copper-deficient conditions.

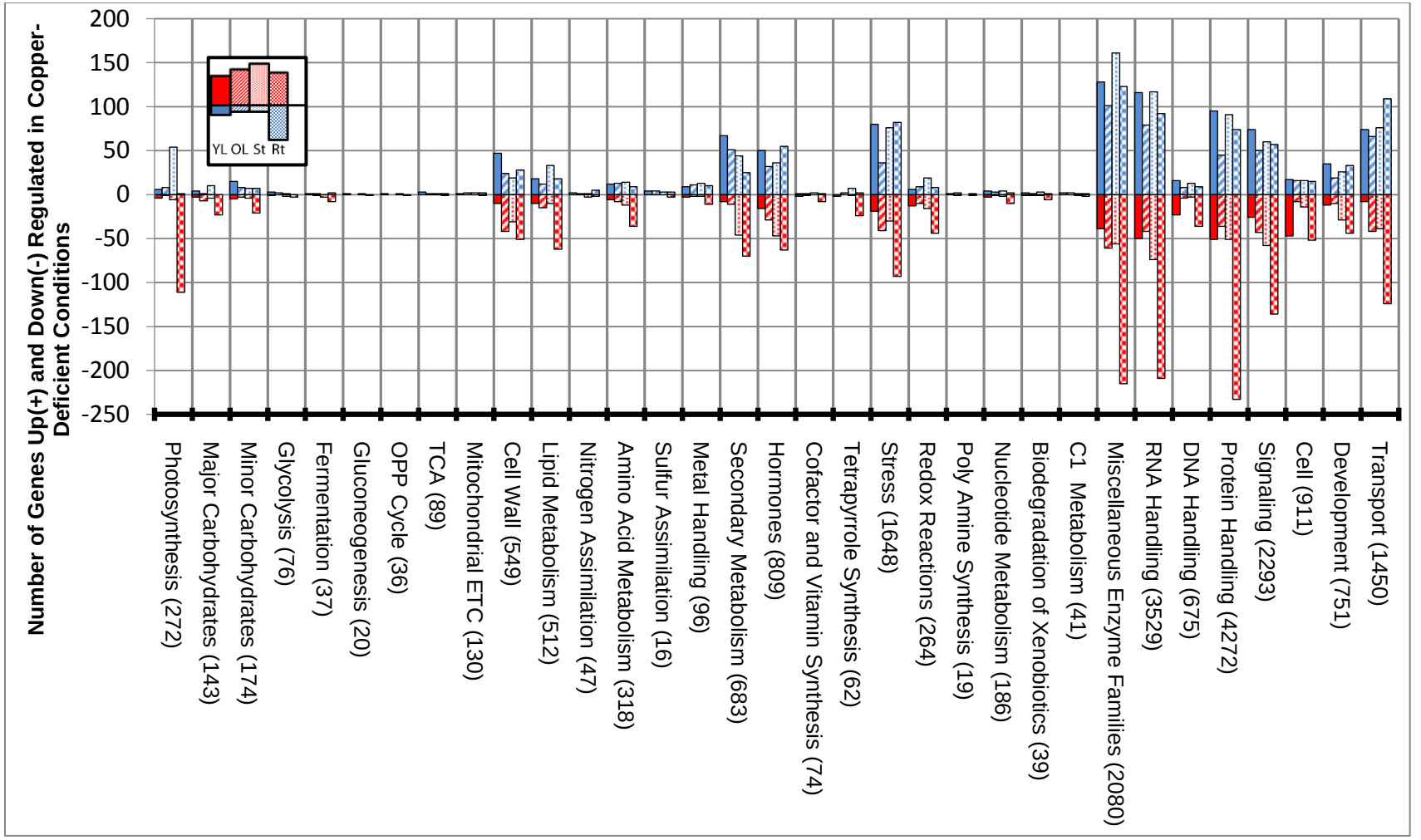

Figure 8: The absolute number of genes differentially expressed in 33 of the MapMan BINs. The miRNA BIN is excluded because it contained no significantly differentially expressed genes, and the miscellaneous BIN is excluded because of the large number of genes up- and down-regulated in each organ. Blue bars represent the number of up-regulated genes, while red bars represent the number of down-regulated genes. In each category, the organs read, from left to right: young leaves, old leaves, stem, and roots. The total number of genes in each category is in parentheses after the BIN name.

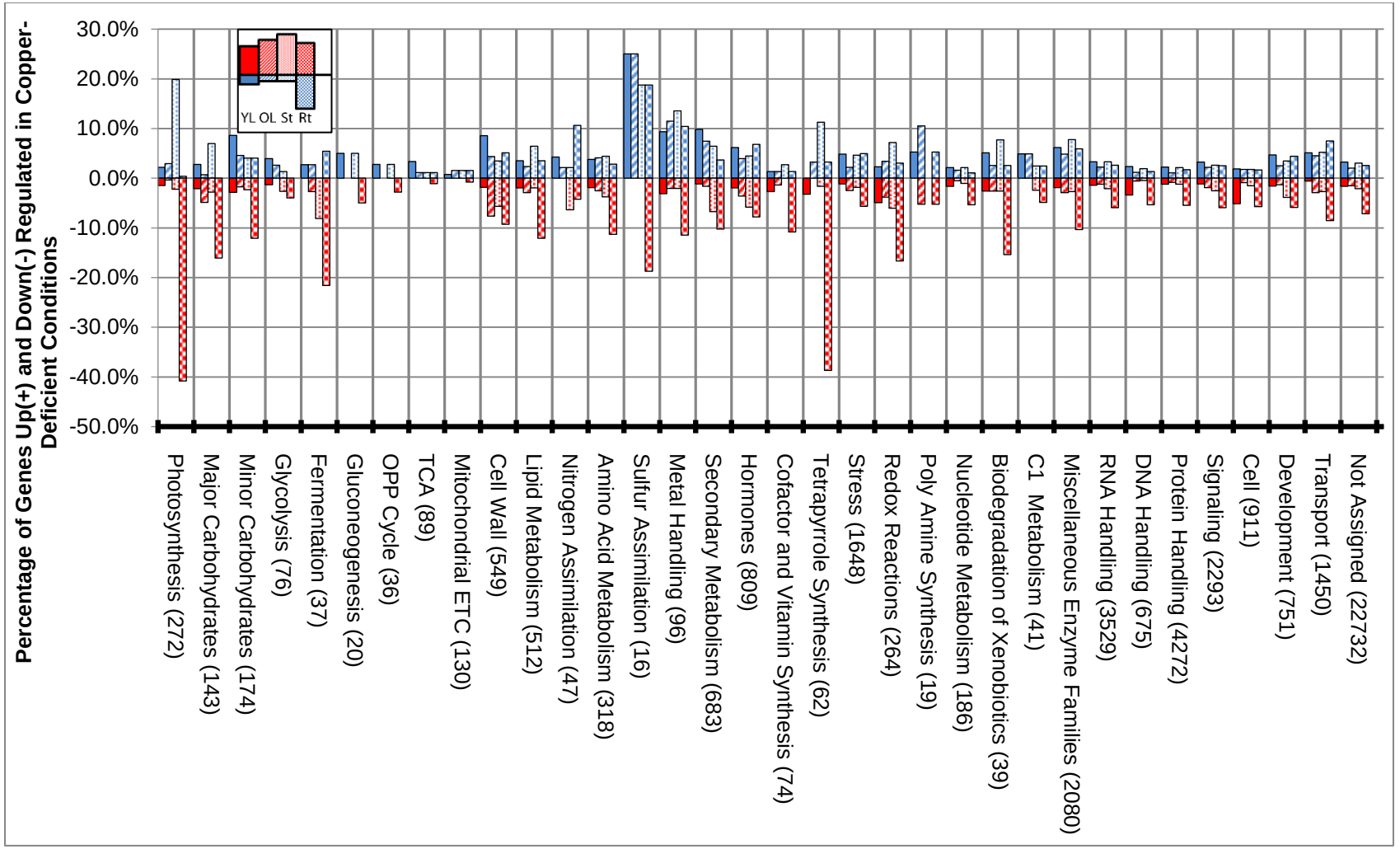

Figure 9: The percentage of genes significantly differentially expressed in 34 of the MapMan BINs. The miRNA BIN is excluded because it contained no significantly differentially expressed genes. Blue bars represent the percentage of up-regulated genes, while red bars represent the percentage of down-regulated genes. In each category, the organs read, from left to right: young leaves, old leaves, stem, and roots. The total number of genes in each category is in parentheses after the BIN name.

Down-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0011s01280	CSD2, CZSOD2	Superoxide dismutase, copper/zinc binding	163.0	20.6	-2.98	~0
POPTR_0004s22620	CSD2, CZSOD2	Superoxide dismutase, copper/zinc binding	146.0	19.1	-2.94	~0
POPTR_0001s39680	No homolog	Tyrosinase, PPO4	2.3	0.3	-2.92	3.08E-07
POPTR_0001s39950	No homolog	Tyrosinase	522.0	70.7	-2.88	~0
POPTR_0001s39630	No homolog	Tyrosinase, PPO2	19.5	3.0	-2.72	~0
POPTR_0001s39650	No homolog	Tyrosinase	0.9	0.1	-2.70	0.00047
POPTR_0004s22630	No homolog		90.5	14.6	-2.63	~0
POPTR_0001s21660	ARPN	Cupredoxin, Plantacyanin	24.5	5.3	-2.21	3.20E-09
POPTR_0001s14010	ATLAC17, LAC17	Multicopper oxidase, Laccase	1.8	0.4	-2.21	2.92E-05
POPTR_0009s01050	CSD2, CZSOD2	Superoxide dismutase, copper/zinc binding	157.5	34.2	-2.20	~0

Up-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0006s16150	SULTR3/5	Sulfate transporter	~0.0	5.9	7.96	5.45E-08
POPTR_0006s12430	No common name	Major facilitator superfamily, general substrate transporter	~0.0	0.8	6.97	0.00747083
POPTR_0016s04570	LEA4-5	Late embryogenesis abundant protein, group 1	0.2	21.7	6.53	~0
POPTR_0016s04670	LEA4-5	Late embryogenesis abundant protein, group 1	0.3	22.6	6.46	~0
POPTR_0001s42200	No Homolog		0.1	5.4	6.45	1.46E-05
POPTR_0001s31500	No common name	FAD/NAD(P)-binding oxidoreductase family protein	~0.0	0.7	6.09	0.0159292
POPTR_0006s04060	No common name	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	~0.0	1.7	5.87	0.0011056
POPTR_0015s05760	No common name	Alpha/beta-Hydrolases superfamily protein	0.4	21.7	5.78	~0
POPTR_0001s09360	No common name		0.4	18.3	5.64	0.00174916
POPTR_0015s02070	HAI3	Protein phosphatase 2C, manganese/magnesium aspartate binding site	0.1	2.7	5.52	1.64E-06

Figure 10: The top ten up- and down-regulated genes (fold-change) in young leaves in copper-deficient conditions.

Down-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)
POPTR_0001s39950	Tyrosinase	522.1	70.7	-2.9
POPTR_0011s01280	Superoxide dismutase, copper/zinc binding	163.0	20.6	-3.0
POPTR_0009s01050	Superoxide dismutase, copper/zinc binding	157.5	34.2	-2.2
POPTR_0004s22620	Superoxide dismutase, copper/zinc binding	146.0	19.1	-2.9
POPTR_0004s22630		90.5	14.6	-2.6
POPTR_0001s21660	Cupredoxin, Plantacyanin	24.5	5.3	-2.2
POPTR_0001s39630	Tyrosinase, PPO2	19.5	3.0	-2.7
POPTR_0011s04710	Tyrosinase, PPO3	7.3	1.6	-2.2
POPTR_0009s10550	Multicopper oxidase, Laccase	5.1	1.2	-2.1
POPTR_0001s39680	Tyrosinase, PPO4	2.4	0.3	-2.9
Up-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)
POPTR_0010s01590	Late embryogenesis abundant protein, Drought Induced	38.4	664.9	4.1
POPTR_1545s00200	Plant lipid transfer protein	95.5	578.9	2.6
POPTR_0013s10350	Phosphorylase	51.1	494.5	3.3
POPTR_0004s18880	Chitinase	6.6	240.8	5.2
POPTR_0001s27540		32.9	166.3	2.3
POPTR_0018s12320		16.1	145.1	3.2
POPTR_0009s11760		17.0	76.8	2.2
POPTR_0012s14200	TonB box, Embryo-specific protein 3	4.0	71.8	4.2
POPTR_0015s13050	PAR1	6.4	62.0	3.3
POPTR_0013s10380	Phosphorylase	7.3	52.9	2.9

Figure 11: Genes in young leaves with a $\log_2$ (fold-change) of 2 or greater and the highest RPKM in copper-sufficient (down-regulated) and copper-deficient conditions (up-regulated) are shown

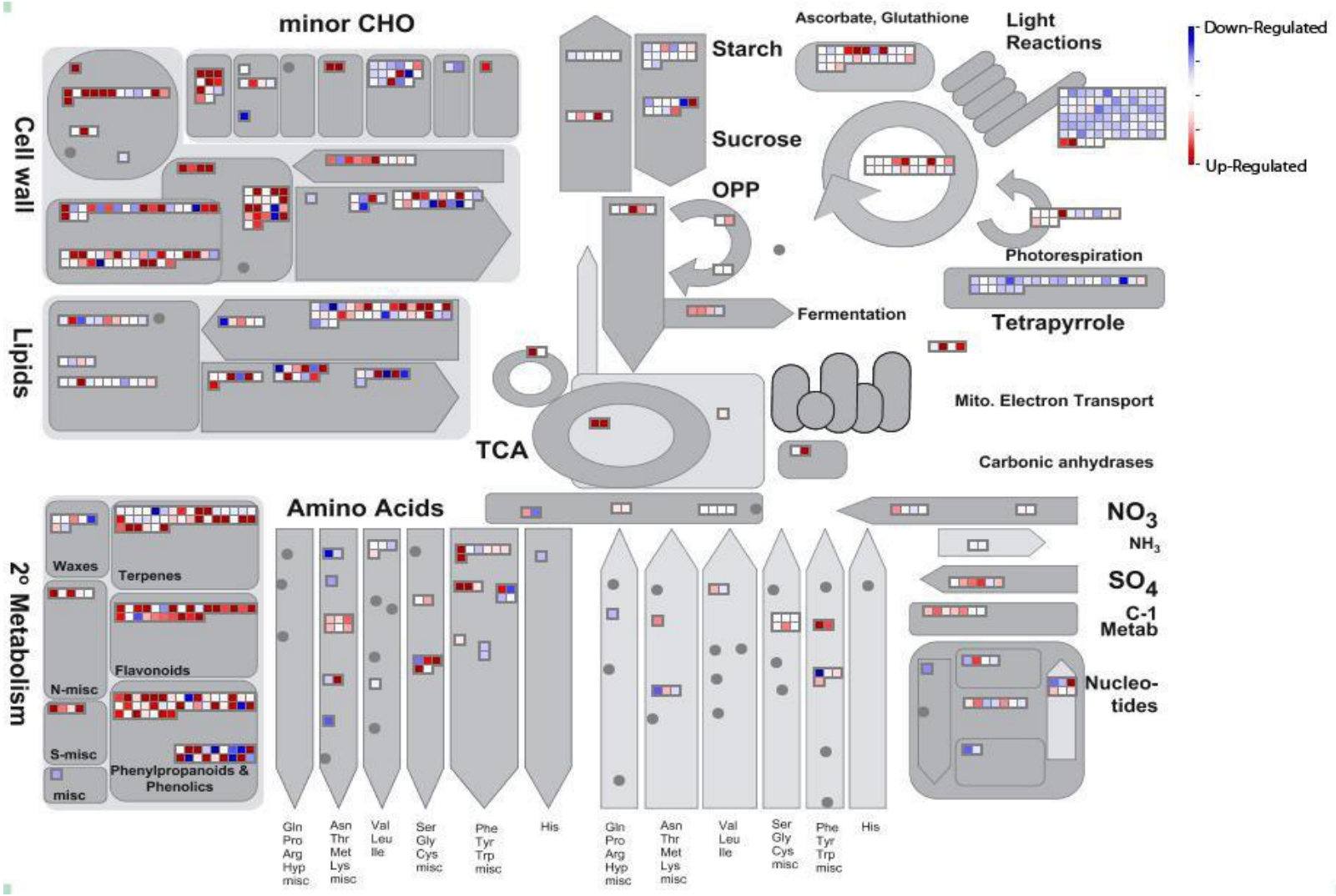

Figure 12: A MapMan metabolic pathway diagram showing all genes significantly differentially expressed in young leaves. Genes indicated with a red box are up-regulated, genes indicated with a blue box are down-regulated, and genes indicated with a white box do not have a large $\log_2$ change in expression.

Down-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0009s13250	ATPRP2, PRP2	Pistil-specific extensin-like protein	74.8	5.5	-3.75	~0
POPTR_0006s22290	ABA2	Glucose/ribitol dehydrogenase	5.2	0.4	-3.62	5.400E-08
POPTR_0018s12630	No common name	Pistil-specific extensin-like protein	20.0	1.7	-3.58	~0
POPTR_0010s07800	No common name	Mycolic acid synthase	2.5	0.2	-3.55	2.300E-06
POPTR_0007s03480	No homolog	Pistil-specific extensin-like protein	0.7	0.1	-3.35	0.012
POPTR_0001s14010	ATLAC17, LAC17	Multicopper oxidase, Laccase	0.5	0.1	-3.24	0.002
POPTR_0008s16940	No homolog	Prion protein	2.6	0.3	-3.16	~0
POPTR_0001s39950	No homolog	Tyrosinase	807.5	94.9	-3.09	~0
POPTR_0002s21820	No common name	Acylhydrolase superfamily protein	25.0	3.4	-2.89	~0
POPTR_0473s00200	No common name	CAP superfamily protein	37.1	5.6	-2.72	4.440E-16
Up-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0001s06250	CYP76C4	Cytochrome P450, E-class, group I	~0.0	1.5	7.31	0.006
POPTR_0003s13420	No common name	Plant disease resistance response protein	~0.0	2.5	6.52	0.006
POPTR_0013s13350	No common name	Protein of unknown function DUF506	~0.0	0.8	6.45	0.016
POPTR_0018s09710	No common name	Plant peroxidase	~0.0	1.0	6.12	0.010
POPTR_0015s00590	RCI3, RCI3A	Plant peroxidase	~0.0	2.5	5.97	9.510E-05
POPTR_0001s31950	AtUGT85A2, UGT85A2	UDP-glucosyltransferase	~0.0	1.1	5.97	0.002
POPTR_0143s00200	ATNRT2.4, NRT2.4	Major facilitator superfamily, general substrate transporter	~0.0	1.0	5.91	0.003
POPTR_0014s12180	No common name	Pectin esterase	~0.0	0.9	5.80	0.002
POPTR_0005s12070	No common name	Plant peroxidase	~0.0	2.1	5.72	6.660E-05
POPTR_0011s00280	No common name	Plant peroxidase	~0.0	1.6	5.72	0.002

Figure 13: The top ten up- and down-regulated genes (fold-change) in old leaves in copper-deficient conditions.

Down-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	log_2 (FC)
POPTR_0001s39950	Tyrosinase	807.5	94.9	-3.1
POPTR_0011s01280	Superoxide dismutase, copper/zinc binding	148.2	27.4	-2.4
POPTR_0004s22620	Superoxide dismutase, copper/zinc binding	143.9	25.1	-2.5
POPTR_0001s19220	Esterase, SGNH hydrolase-type	95.3	17.7	-2.4
POPTR_0004s22630		75.8	15.4	-2.3
POPTR_0009s13250	Pistil-specific extensin-like protein	74.8	5.5	-3.8
POPTR_0002s26160	RAD-like	42.1	9.2	-2.2
POPTR_0001s08330	Superoxide dismutase, copper/zinc binding	39.3	8.6	-2.2
POPTR_0473s00200	CAP superfamily protein	37.1	5.6	-2.7
POPTR_0008s18250	Pectate lyase family protein	29.0	6.3	-2.2
Up-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	log_2 (FC)
POPTR_0019s13150	Plant metallothionein, family 15	15.2	187.0	3.6
POPTR_0007s13420	Plant peroxidase	7.4	69.6	3.2
POPTR_0453s00230	Myb-like DNA-binding domain, SHAQKYF class	3.9	41.8	3.4
POPTR_0008s06210	Plant lipid transfer protein	4.6	40.7	3.1
POPTR_0005s17200	Phosphate-induced protein	9.2	37.2	2.0
POPTR_0012s01630	Alternative oxidase	3.1	30.6	3.3
POPTR_0015s05760	Lipase	5.3	24.5	2.2
POPTR_0012s14200	TonB box, conserved site	5.9	23.7	2.0
POPTR_0001s16640	Tify protein	5.6	22.6	2.0
POPTR_0008s13040	Lipid transport superfamily protein	0.7	20.4	5.0

Figure 14: Genes in old leaves with a log_2 (fold-change) of 2 or greater and the highest RPKM in copper-sufficient (down-regulated) and copper-deficient conditions (up-regulated) are shown.

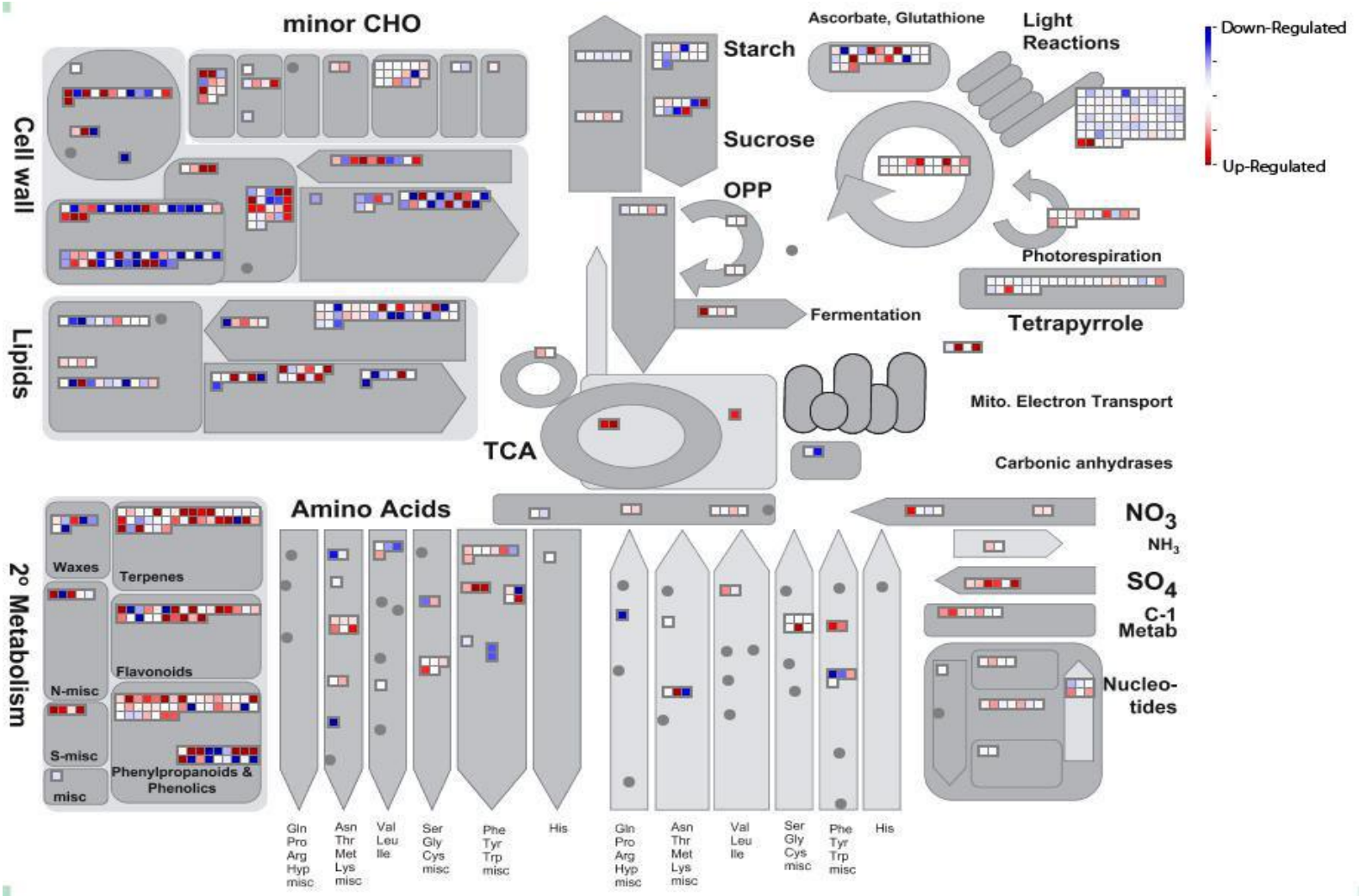

Figure 15: A MapMan metabolic pathway diagram showing all genes significantly differentially expressed in old leaves. Genes indicated with a red box are up-regulated, genes indicated with a blue box are down-regulated, and genes indicated with a white box do not have a large $\log_2$ change in expression.

Down-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0002s01740	PETE1	Plastocyanin	69.8	~0.0	-11.00	1.460E-09
POPTR_0002s03670	No homolog		6.7	~0.0	-7.86	1.920E-07
POPTR_0002s18910	No homolog		18.7	0.3	-6.05	~0
POPTR_0001s18500	ATLAC2, LAC2	Multicopper oxidase, Laccase	4.0	0.2	-4.05	1.210E-12
POPTR_0001s14010	ATLAC17, LAC17	Multicopper oxidase, Laccase	6.2	0.4	-3.98	2.220E-16
POPTR_0002s10170	Cupredoxin	Cupredoxin	117.2	13.2	-3.15	~0
POPTR_0008s07370	LAC5	Multicopper oxidase, Laccase	67.8	7.9	-3.09	~0
POPTR_0001s39660	No homolog		26.3	3.35	-2.97	~0
POPTR_0006s15600	No common name		0.4	0.1	-2.87	~0.011
POPTR_0009s10550	LAC11	Multicopper oxidase, Laccase	2.57455	0.3	-2.84	5.720E-07

Up-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0001s25290	COPT1	Copper transporter	~0.0	4.4	6.90	0.009
POPTR_0006s22510	No common name	Uncharacterized protein family UPF0497, trans-membrane plant subgroup	0.1	1.5	4.69	0.010
POPTR_0001s23570	No common name	myb-like HTH transcriptional regulator family protein	0.3	6.8	4.67	0.012
POPTR_0009s03950	ZIP2	Zinc/iron permease	1.3	29.3	4.51	~0
POPTR_0008s11950	LBD42	Lateral organ boundaries	0.1	1.9	4.49	0.001
POPTR_0012s01630	AOX1A, ATAOX1A	Alternative oxidase	5.3	116.6	4.47	~0
POPTR_0011s04090	No homolog		2.2	45.3	4.38	~0
POPTR_0096s00200	ATFRO4, FRO4	Ferric Reductase	0.1	2.1	4.30	0.001
POPTR_0017s00660	No common name	Zinc ion binding; nucleic acid binding	0.1	1.4	4.28	~0
POPTR_0003s10600	DYAD, SWI1	SWITCH1	0.6	11.8	4.20	~0

Figure 16: The top ten up- and down-regulated genes (fold-change) in stems in copper-deficient conditions.

Down-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)
POPTR_0002s10170	Cupredoxin	117.2	13.2	-3.1
POPTR_0002s10150	Cupredoxin	89.4	13.8	-2.7
POPTR_1040s00200	Cupredoxin	82.4	11.8	-2.8
POPTR_0002s01740	Plastocyanin	69.8	0.0	-11.0
POPTR_0008s07370	Multicopper oxidase, Laccase	67.8	7.9	-3.1
POPTR_0001s21660	Cupredoxin	66.9	11.0	-2.6
POPTR_0001s33960	Cupredoxin	55.3	12.0	-2.2
POPTR_0016s11950	Multicopper oxidase, Laccase	53.1	12.7	-2.1
POPTR_0011s01280	Superoxide dismutase, copper/zinc binding	51.4	11.0	-2.2
POPTR_0009s04720	Multicopper oxidase, Laccase	47.6	10.7	-2.1
Up-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)
POPTR_0013s10350	Nucleoside phosphorylase	323.7	1560.8	2.3
POPTR_0013s10380	Nucleoside phosphorylase	85.6	929.8	3.4
POPTR_0019s07690	Nucleoside phosphorylase	56.4	318.6	2.5
POPTR_0013s10370	Nucleoside phosphorylase	29.3	260.3	3.2
POPTR_0012s03180	Oligopeptide transporter OPT superfamily	36.9	174.4	2.2
POPTR_0001s28030	Las1-like	17.6	157.9	3.2
POPTR_0453s00230	Myb-like DNA-binding domain, SHAQKYF class	10.1	148.6	3.9
POPTR_0007s13420	Plant peroxidase	9.5	145.0	3.9
POPTR_0012s01630	Alternative oxidase	5.3	116.6	4.5
POPTR_0011s07830	Yellow Stripe like	13.2	112.4	3.1

Figure 17: Genes in stems with a $\log_2$ (fold-change) of 2 or greater and the highest RPKM in copper-sufficient (down-regulated) and copper-deficient conditions (up-regulated) are shown.

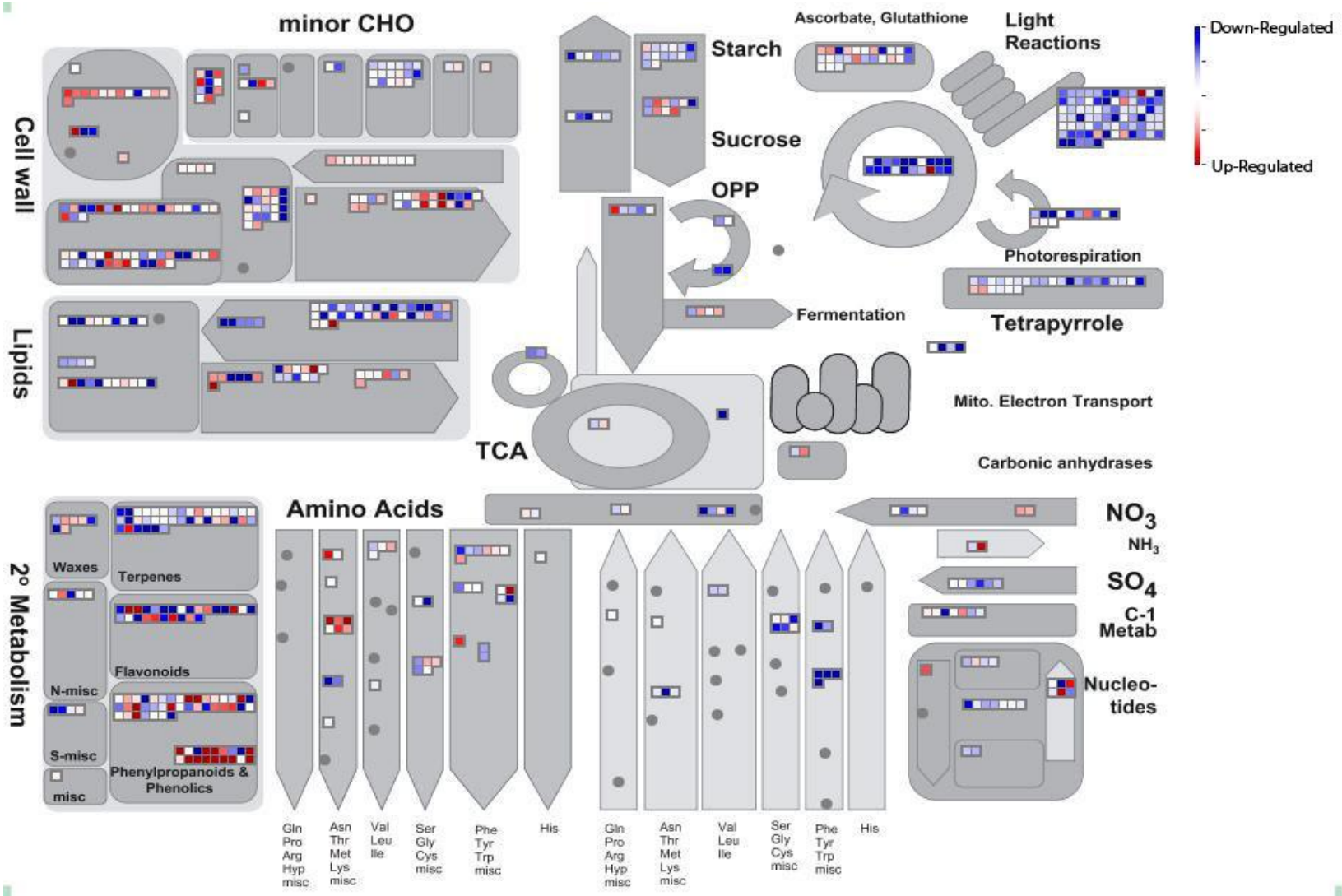

Figure 18: A MapMan metabolic pathway diagram showing all genes significantly differentially expressed in stems. Genes indicated with a red box are up-regulated, genes indicated with a blue box are down-regulated, and genes indicated with a white box do not have a large $\log_2$ change in expression.

Down-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0078s00210	PLA IIA, PLA2A, PLP2	Acyl transferase	5.8	~0.0	-9.71	0.004
POPTR_0002s12610	No common name	Subtilisin-like serine endopeptidase	7.1	~0.0	-9.47	2.880E-07
POPTR_0007s06310	CYP75B1, D501, TT7	Cytochrome P450I	2.8	~0.0	-9.36	0.007
POPTR_0019s01130	ATFLS1, FLS, FLS1	Isopenicillin N synthase	7.1	~0.0	-9.16	~0
POPTR_0003s09200	bHLH071	Helix-loop-helix DNA-binding domain	11.9	~0.0	-9.08	1.850E-05
POPTR_0001s20800	EDA17, HTH	Amine oxidase	3.6	~0.0	-9.03	~0
POPTR_0006s06090	No common name		47.3	0.1	-8.89	1.120E-10
POPTR_0019s12370	ATCHITIV, ATEP3, CHIV, EP3	Glycoside hydrolase	73.1	0.2	-8.86	1.390E-12
POPTR_0005s25350	BLH11	Lambda-like repressor	3.3	~0.0	-8.53	~0
POPTR_0019s03110	PLA IIA, PLA2A, PLP2	Patatin/Phospholipase A2-related	2.8	~0.0	-8.50	0.009

Up-regulated						
Ensembl Gene ID	Ara. Homolog	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)	P-value
POPTR_0096s00200	ATFRO4, FRO4	Ferric reductase	0.7	60.2	6.48	~0
POPTR_0011s04090	No homolog		0.3	19.0	6.12	1.290E-07
POPTR_0017s01690	ATFRO4, FRO4	Ferric reduction oxidase	4.9	247.4	5.64	~0
POPTR_0012s01630	AOX1A, ATAOX1A	Alternative oxidase	2.1	73.2	5.10	~0
POPTR_0003s10600	DYAD, SWI1	SWITCH1	0.1	3.1	5.04	~0
POPTR_0016s04840	No common name	Myb-like DNA-binding domain	0.4	10.8	4.67	1.600E-11
POPTR_0001s23570	No common name	Myb-like HTH transcriptional regulator	0.3	7.7	4.53	0.006
POPTR_0453s00230	No common name	Myb-like DNA-binding domain	3.6	68.0	4.25	~0
POPTR_0001s28030	No common name	Las1-like	12.3	194.8	3.98	~0
POPTR_0010s11030	No homolog		0.8	10.2	3.70	~0

Figure 19: The top ten up- and down-regulated genes (fold-change) in roots in copper-deficient conditions.

Down-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)
POPTR_0005s15660	Ribulose bisphosphate carboxylase, small chain	4392.8	41.8	-6.7
POPTR_0005s26080	Chlorophyll a/b binding protein domain	2410.5	72.6	-5.1
POPTR_0010s16030	Heat stable protein	2214.3	536.8	-2.0
POPTR_0004s09910	Ribulose bisphosphate carboxylase, small chain	1878.2	15.9	-6.9
POPTR_0011s02770	Chlorophyll a/b binding protein domain	1831.8	49.7	-5.2
POPTR_0002s22220	Chlorophyll a/b binding protein domain	1110.2	61.0	-4.2
POPTR_0019s09140	Light harvesting complex of photosystem II	1109.6	62.0	-4.2
POPTR_0003s05110	Photosystem I PsaH, reaction center subunit VI	929.4	38.3	-4.6
POPTR_0001s11600	Photosystem I PsaF, reaction center subunit III	922.0	51.1	-4.2
POPTR_0014s17070	Chlorophyll a/b binding protein domain	811.3	44.3	-4.2
Up-regulated				
Ensembl Gene ID	Short Description	RPKM 50 nM Cu	RPKM 0 nM Cu	$\log_2$ (FC)
POPTR_0010s24290	Hg scavenger	921.1	3549.1	1.9
POPTR_0006s23580	Copper transporter, COPT5	172.9	644.4	1.9
POPTR_0009s03950	Zinc/iron permease	38.2	447.0	3.5
POPTR_0017s01690	Ribosomal protein S12/S23	4.9	247.4	5.6
POPTR_0001s28030	Las1-like	12.3	194.8	4.0
POPTR_0012s03180	Oligopeptide transporter OPT superfamily	26.9	101.2	1.9
POPTR_0012s01630	Alternative oxidase	2.1	73.2	5.1
POPTR_0453s00230	Myb-like DNA-binding domain, SHAQKYF class	3.6	68.0	4.2
POPTR_0011s07830	Yellow Stripe-Like	12.2	63.4	2.4
POPTR_0096s00200	Ferric Reductase	0.7	60.2	6.5

Figure 20: Genes in roots with a $\log_2$ (fold-change) of 2 or greater and the highest RPKM in copper-sufficient (down-regulated) and copper-deficient conditions (up-regulated) are shown.

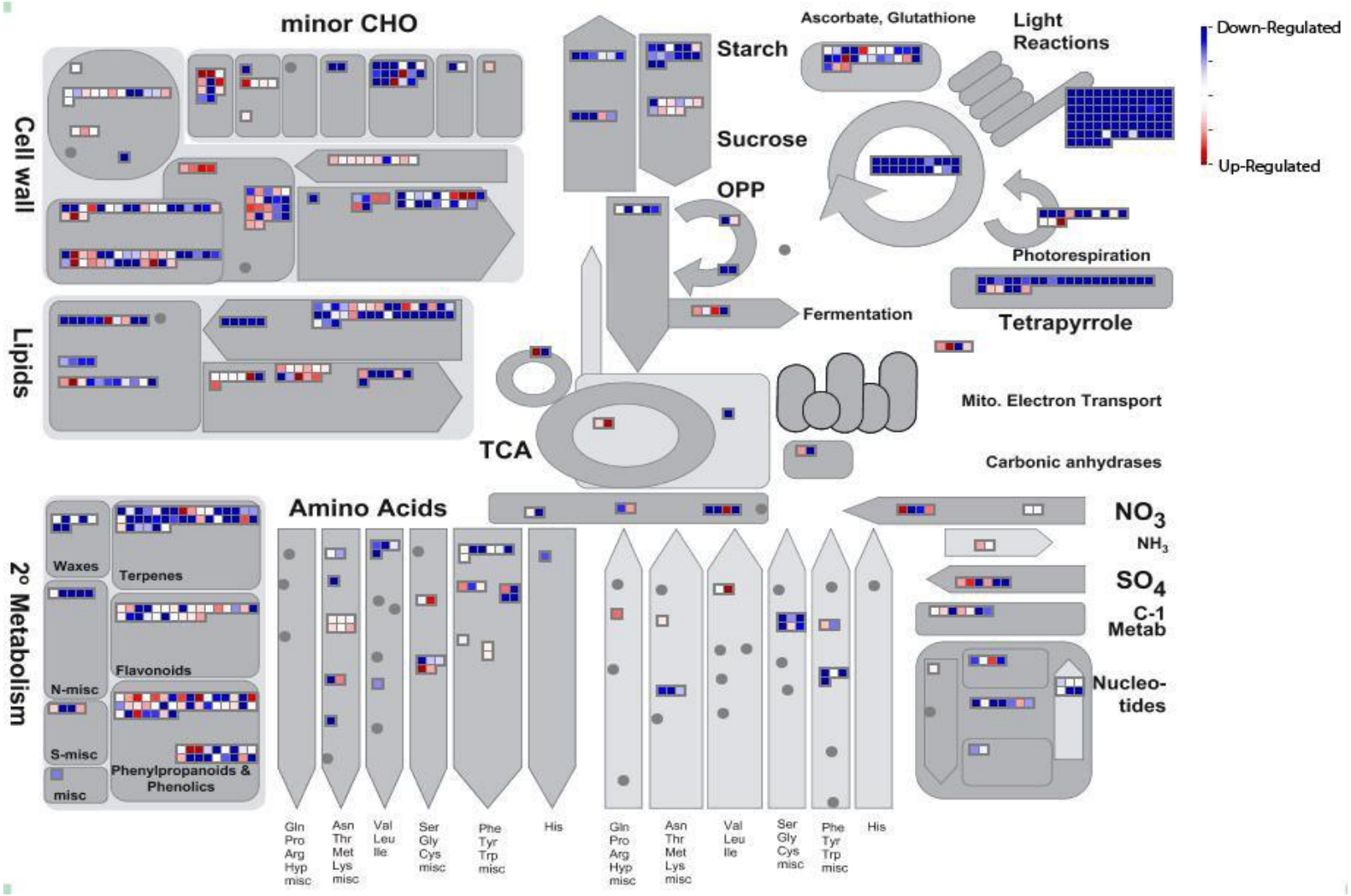

Figure 21: A MapMan metabolic pathway diagram showing all genes significantly differentially expressed in roots. Genes indicated with a red box are up-regulated, genes indicated with a blue box are down-regulated, and genes indicated with a white box do not have a large $\log_2$ change in expression.

CHAPTER 3

COPPER HOMEOSTASIS OF THE POPULUS TRICHOCARPA LACCASE GENE FAMILY

SUMMARY

The laccases are a family of copper-containing proteins that have been implicated in the formation of lignin. Laccases are found in many organisms and make up large gene families in the higher plants. In this chapter we looked at the bioinformatics of laccases in *Populus trichocarpa* as an initial step toward understanding these proteins. We describe twenty-five new laccase genes in poplar, based on their sequence similarity with the seventeen laccases from *Arabidopsis* and thirteen previously annotated laccases from *Populus*. We proceeded to classify these twenty-five genes and the thirteen already described genes into seven categories based on their protein sequence similarity. The RNA-SEQ data of these laccases under copper-sufficient conditions in *Populus* show many of the laccases as highly expressed in the stem tissue or ubiquitously throughout the vegetative organs. Twenty-one of these thirty-eight laccase genes have a conserved target site for miRNA *397* and four genes have a target site for miRNA *408*. All but two high-expressed laccases have a predicted targeting site for a Cu-miRNA and show down-regulation under copper-deficient conditions. These data support the copper homeostasis model described in *Arabidopsis* and further studies may implicate laccases as lignin synthesizing enzymes.

INTRODUCTION

In plants, copper deficiency leads to a variety of symptoms including reduced photosynthetic activity, decreased growth, changes in cell wall formation, and impeded wound healing (Epstein and Bloom, 2005; Marschner, 2011). Copper deficiency severely affects wood production in forests and can be a major contributor to biomass loss for woody plants (Ruiter,

1969). Modifications in cell wall structure related to copper deficiency include a reduction in lignin and overall cell wall thickness (Marschner, 2011). Reduced lignin and cell wall thickness lead to a decrease in cell wall strength and an inability to resist the negative pressure of transpiration, resulting in xylem vessel collapse (Berthet *et al.*, 2011). Conversely, copper toxicity also has a large impact on plant growth and photosynthetic activity (Bernal *et al.*, 2004; Yruela *et al.*, 1996). The primary symptom of copper toxicity is chlorosis in the green, vegetative tissue due to increased intercellular oxidative stress and the down regulation of photosynthetic genes (Patsikka *et al.*, 2002).

It may be that the impact of copper deficiency on wood formation is due to the lack of activity of copper-containing enzymes known as laccases (Bao *et al.*, 1993; Berthet *et al.*, 2011; Davin *et al.*, 1992; Driouich *et al.*, 1992; McDougall and Morrison, 1996; Sterjiades, Dean, and Eriksson, 1992). Laccases are secreted multicopper oxidases that can catalyze one-electron oxidations of phenolic compounds (Reinhammar and Malmstroem, 1981). The oxidation of four phenolic compounds occurs simultaneously with the reduction of one O_2 molecule into two H_2O molecules by the four-copper catalytic center of laccases. Superoxides are not released in this process because O_2 is reduced directly to H_2O in the catalytic center (Claus, 2004; Solomon, Sundaram, and Machonkin, 1996). It has been proposed that oxidized phenolics can then undergo subsequent radical-radical coupling, polymerizing with each other to form complex polyphenolics, such as lignin (Turlapati *et al.*, 2011).

The catalytic center of laccases is unique and contains four copper atoms in three different conformations and two coordination centers (Figure 22). The first coordination center is responsible for catalyzing 4 one-electron oxidations of the substrate and contains one type I (T1) copper atom (Solomon *et al.*, 1996). As electrons are stripped from their substrate

molecules they are passed to the second coordination center (Ducros *et al.*, 1998). The second coordination center is a trinuclear cluster containing one type II (T2) copper atom and two type III (T3a and T3b) copper atoms. The trinuclear cluster is responsible for reducing oxygen and generating water (Solomon *et al.*, 1996).

Since the discovery of laccases in the lacquer tree (Yoshida, 1883) they have been shown to be broadly distributed across many phyla, including bacteria, fungi, insects and plants (Claus, 2003; Mayer and Staples, 2002; McCaig, Meagher, and Dean, 2005; Silva *et al.*, 1995). In fungi, laccases have been shown to degrade lignin (Eggert, Temp, and Eriksson, 1997), produce pigments (Aramayo and Timberlake, 1990), and even protect the fungus from pathogens (Salas *et al*., 1996). In plants, however, laccase function has remained elusive. *In vitro* evidence has indicated that they play a role as a biocatalyst in the generation or degradation of phenolic polymers (Mayer and Staples, 2002), leading us to believe that they are involved in lignin synthesis or degradation or both (Bao *et al.*, 1993; Berthet *et al.*, 2011; Freudenberg, 1959; Liu *et al*., 1994; O'Malley *et al*., 1993; Richardson, Duncan, and McDougall, 2000; Sterjiades *et al.*, 1992), pigment production (Pourcel *et al.*, 2005), and sap hardening during wound healing (Butt, 1980; Malmstrom, Andreasson, and Reinhammar, 1975). Genetic approaches to showing the precise function of laccases in plants have been obscured by the size and complexity of their gene family in which gene function redundancy is likely, while biochemical approaches have also been inconclusive because the genes are apoplastic.

In the model organism *Arabidopsis thaliana*, there are seventeen putative laccase genes (AtLac1-17) (McCaig *et al.*, 2005). Double knockouts of genes AtLAC4 and 17 have shown decreases in total lignin, implying that these laccases have a role in lignin synthesis (Berthet *et al.*, 2011). AtLAC15 single mutants have a decrease in the browning of the *Arabidopsis* seed

testa caused by the reduction in the oxidation of flavonoids (Pourcel *et al.*, 2005). To date, these are the only three laccases in *Arabidopsis* that have exhibited a phenotype from knockout assays. Other single knockout assays of laccases have been unsuccessful in generating a defined phenotype, meaning more complicated double and triple knockout/knockdown methods will most likely be necessary to ascribe detailed functions to each gene in the family.

The copper economy model predicts that under copper-deficient conditions, copper ions are conserved for proteins essential for the plant's survival, such as plastocyanin and cytochrome-c oxidase (Burkhead *et al.*, 2009). In order to accomplish this, "non-essential" copper-containing proteins such as Cu/Zn SODs, laccases and plantacyanin are down-regulated by miRNA-mediated degradation of their transcripts (S. E. Abdel-Ghany and Pilon, 2008). These miRNAs are regulated in turn by the SQUAMOSA promoter binding-like transcription factor 7 (SPL7), which is inhibited by abundant copper in the cell (Bernal *et al.*, 2012; Yamasaki *et al.*, 2009). This system allows for down-regulation of "non-essential" copper-containing proteins so that the majority of the free copper pool is conserved for "essential" copper proteins when copper is limiting. In *Arabidopsis*, the primary miRNAs that are directly up-regulated by the SPL7 transcription factor are miRNAs *397*, *398* and *408* (Yamasaki *et al.*, 2009). The miRNA *397* targets laccases 2, 4 and 7; miRNA *398* targets Cu/Zn SODs and a subunit of cytochrome c oxidase; and miRNA *408* targets laccases 3, 12 and 13 as well as plantacyanin (S. E. Abdel-Ghany and Pilon, 2008). In *Populus trichocarpa*, a novel miR1444 has been shown to target another copper-containing protein, polyphenol oxidase, under copper-deficient conditions (Ravet *et al.*, 2011).

The laccase gene family is a large and complex family with a proposed function in cell wall lignification. The gene family has been difficult to study in *Arabidopsis* because of a lack

of lignified tissues and weak phenotypes. *Populus* is a perennial that develops thick, lignified, and woody xylem tissues, which gives it an advantage over *Arabidopsis* for the study of the laccase gene family. However, the *Populus* genome is much larger than that of *Arabidopsis*, meaning that the laccase gene family also has the potential of being larger with more redundancy. Before detailed molecular studies can begin on individual laccases in *Populus* we need to understand the size and scope of the entire laccase gene family. In this chapter we will identify, classify and compare what is known from identified laccases genes as well as identifying and classifying new un-annotated laccases to give a starting point for future molecular studies.

MATERIALS AND METHODS

Sequence Collection

A BLASTn search of the NCBI *Populus* database (http://blast.ncbi.nlm.nih.gov/) was performed using the 17 described laccase genes (LAC1-17) from *A. thaliana* (http://www.arabidopsis.org/) and 13 laccases (LAC1a-d, 2, 3, 90 a-d, 110a-c) from *P. trichocarpa* (http://blast.ncbi.nlm.nih.gov/) as queries. The top fifty best hits for each query were compared and the redundant hits were removed. Hits with non-laccase annotations were removed from the list and only hits that were either un-annotated or known laccases were retained. This list was further refined by ClustalX alignment (Settings-Protein Weight Matrix: Gonnet; Gap open: 10; Gap Extension: 0.20; Gap distances: 5; No end gaps) to include only sequences that contained the one methionine, one cysteine, and ten histidine residues necessary for binding the four copper atoms required for laccase function (Claus, 2004; Kumar, Phale, Durani, and Wangikar, 2003; Solomon *et al.*, 1996). All *Populus* laccase genes were renamed using a standard nomenclature to simplify this discussion and are listed in Figure 23.

Signal sequence and subcellular targeting prediction was performed for all laccase genes using SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/). The program was run with the following parameters: organism group: eukaryote; D-cutoff values: Default; minimum signal sequence length: 10.

A chromosome map was generated using the Popgenie Chromosome Diagram Tool (http://popgenie.org/gp). Chromosome size is to scale. Genes were placed on the diagram according to their approximate location on the chromosome (Figure 24).

Tree Generation

The amino acid sequences of the laccases were aligned with previously identified laccase sequences from *Arabidopsis* using ClustalW (http://www.clustal.org/). A phylogenetic analysis was performed to determine relationships. The sequence alignment, containing 574 amino acid characters, was input into Mr. Bayes (http://mrbayes.sourceforge.net/) to generate a phylogenetic tree using a Bayesian Markov chain Monte Carlo approach. A mixed amino acid model was used, four chains were run for 1.5E7 generations, and the run was repeated to establish convergence. A 50% majority rule consensus tree was used to display results.

miRNA Target Prediction

We used psRNATarget (http://plantgrn.noble.org/psRNATarget/) to predict targeting of miRNAs *408* and *397* against the *Populus* laccases because these are the two cu miRNAs that target laccases in *Arabidopsis*. Maximum exception tolerated for targeting was set to a score of 5, with a complete mismatch counting as 1 and a wobble pair (G-U) counting as 0.5.

Laccase Expression

Laccase expression in four vegetative organs (young leaves, old leaves, stem, and roots) and their regulation under copper-deficient conditions were determined by RNA-SEQ. The RNA-SEQ experiment is described in detail in chapter 2.

RESULTS AND DISCUSSION

New laccase genes in *Populus trichocarpa* were uncovered by a BLASTp search of the NCBI *Populus* database using known laccases from *A. thaliana* and *P. trichocarpa* as the query sequences (McCaig *et al.*, 2005; Ranocha *et al.*, 1999). Twenty-seven previously un-annotated protein sequences were recovered and aligned to previously annotated laccase sequences. Of these twenty-seven, two sequences (POPTR_0001s18500 and POPTR_0011s06880) did not contain the appropriate number or arrangement of histidine and cysteine residues to bind the four copper atoms essential for laccase function (Figures 22 and 25). These two predicted genes were not included in further analysis. As seen in *Arabidopsis*, the *Populus* laccase genes are highly variable at the beginning of their sequence because of their subcellular targeting domain, and in between the second and third copper binding motifs (McCaig *et al.*, 2005). Almost all laccase genes end in an N-terminal cysteine residue (Figure 25).

The final list of poplar laccases contains twenty-five un-annotated proteins and thirteen previously annotated proteins. These thirty-eight proteins range between 518 and 582 amino acids in length. When put into SignalP (http://www.cbs.dtu.dk/services /SignalP/) to predict their cellular target location and signal sequence, thirty-two are predicted to be secreted (apoplastic) with the signal peptide consisting of 14-34 amino acids (Figure 23).

Mapping the putative popLAC genes onto a diagram of the chromosomes shows no obvious pattern (Figure 24). The laccases are evenly spread out across the genome and no

chromosomes show a larger than average number of laccase genes. We do see clusters of two to five genes in tandem on several chromosomes. The genes in these clusters are not necessarily duplications of the same gene, but we can see at least 9 tandem duplications. The laccase genes also do not cluster with the Cu-miRNA *397, 398,* and *408* genes.

Putative laccase protein sequences from *Arabidopsis* and *Populus* were aligned using ClustalX, showing high conservation of amino acids between the species. After a Bayesian analysis, a radial cladogram of the sequence alignment shows that laccases can be classified into up to seven groups based solely on their sequence similarity (Figure 26). Groups of laccases range from a single member (group 7) up to 11 members (Group 2). Until physiological function can be ascribed to all genes in the gene family, this classification remains subjective (McCaig *et al.*, 2005). However, by including the *Arabidopsis* laccases in the cladogram, we can say something about the members in groups 1, 2, and 4. Groups 1 and 2 contain atLAC17 and atLAC4 respectively. *Arabidopsis* plants containing T-DNA insertions in both of these genes showed a 20-40% decrease in lignin content, while the single mutants showed little to no decrease in lignin (Berthet *et al.*, 2011). AtLAC17 in particular was shown to be responsible for the deposition of g-lignin in the stem fibers (Berthet *et al.*, 2011). On the other hand, the atLAC15 gene in group 4 has been shown to be highly expressed in the seed testa, and knockouts of this gene result in seeds with a brighter coat (Pourcel *et al.*, 2005). As more laccase genes are given precise biochemical functions we will be able to better annotate the phylogenetic tree and see if classification based on sequence indicates conserved biochemical function within the groupings.

Using our phylogenetic classification, we renamed the thirty-eight genes in a more intuitive way. Each gene is labeled with “pop” for *Populus*, “LAC” for laccase, and a number in

which the ones place indicates its grouping (1-7) and the tenths place represents its individual identity (Figure 23). For example, popLAC2.3 is the third sequence in group 2.

A rectangular tree also shows the similarity of the thirty-eight laccase genes. Data from RNA-SEQ were overlaid to show the basal expression of all the genes in each organ and each gene's regulation under copper-deficient conditions (Figure 27). Most of the genes in groups 1, 2, 3, and 7 show high expression (RPKM >5) in the stem and root tissue, or they are ubiquitously highly expressed, while the genes in groups 4, 5 and 6 show low expression (RPKM <5) in every organ.

The mRNAs of all putative laccase genes were input into psRNAtarget to look for potential targets of miRNAs *398* and *408*. The miRNAs *397* and *408* are known to be regulated by the SPL7 transcription factor, which is in turn regulated by copper availability in the cell (Yamasaki *et al.*, 2009). The miRNAs *397* and *408* have also been shown to target laccases in *Arabidopsis* (S. E. Abdel-Ghany and Pilon, 2008; Bonnet *et al*., 2004; Jones-Rhoades and Bartel, 2004; Lu et al., 2005; Yamasaki et al., 2007). miRNA *397* was predicted to target twenty-one of the thirty-eight *Populus* laccases, while miRNA *408* was predicted to target four (Figure 27). Up to 5 base-pair mismatches between a Cu-miRNA and its mRNA target still results in the degradation of the transcript (Palatnik *et al.*, 2003). In this case we set the threshold for 5 with a complete mismatch counting as "1" and a G-U base pair counting as "0.5". Interestingly, the miRNA target prediction shows a group-specific pattern. Groups 1, 2, and 5 contain many potential targets of miRNA *397*, while group 3 is targeted by both miRNA *397* and *408*. It is interesting that these are also the groups that are primarily expressed in stems and roots, because lignin is also distributed this way. The transcripts of genes in groups 4, 6 and 7 are not predicted

miRNA targets and therefore are not predicted to be regulated in response to copper deficiency. Groups 4, 6 and 7 also contain genes ubiquitously expressed at a low level (Figure 27).

For the most part, if a gene is highly expressed and a predicted target of a Cu-miRNA, we see a substantial down-regulation of the gene in copper-deficient conditions. The stems seem to have the highest expression of the laccase genes as well as the greatest down-regulation of the genes. The down-regulations of genes in the stem are almost all statistically significant (Figure 28).

On the other hand, if a gene is low-expressed and a predicted miRNA target, we often do not see down-regulation by the miRNA; in fact, we can often see up-regulation of a few of the low-expressed genes (Figure 27). This up-regulation is not always statistically significant. Plotting only genes with significant differential expression shows that no gene with RPKM higher than 6 in copper-sufficient conditions is significantly up-regulated under copper-deficient conditions (Figure 28). The few genes that are below 6 RPKM and significantly up-regulated are being up-regulated in old and young leaves, but never in the stem. Although these changes in expression are mathematically significant, the question remains, are they biologically significant? Most likely there is no reason for the organism to waste resources attempting to regulate these low expressed genes, but the mechanism for how Cu-miRNAs selectively down-regulate high abundance mRNAs has yet to be discovered.

CONCLUSIONS

In this study we have found twenty-five laccase genes that have not been previously described explicitly as laccases. These twenty-five genes, combined with the thirteen already annotated laccase genes from poplar, make thirty-eight genes in the *Populus* laccase gene family. The laccase gene family in poplar is nearly double that of *Arabidopsis*, even though there have

been two genome-wide duplications since *Arabidopsis* diverged from *Populus* (Tuskan *et al.*, 2006). If genome duplication were the only factor, we would expect to see nearly 68 genes in the *Populus* laccase family. We would also see regions of synteny. Most likely, there is no evolutionary pressure to keep copies of genes with redundant function and many of the laccase gene copies have not been preserved.

The thirty-eight laccase genes are divided into seven distinct phylogenetic groups ranging from a single gene to eleven per group. Originally, laccases were classified by their isoelectric focusing point (pI). The hypothesis was that the pI shows a relationship with the enzyme kinetics and the substrate specificity of the enzymes (Dean *et al.*, 1998). When more laccases were described in a larger number of organisms, it was shown that the pI values vary little between the known laccases and that pI alone was not a good characteristic for describing the variability in the laccase family (McCaig *et al.*, 2005). Using only plant laccases known at the time, the laccase genes were divided into 6 groups based on their amino acid sequence similarity (McCaig *et al.*, 2005). When all thirty-eight *Populus* laccases are included in the phylogenetic tree, a seventh, divergent group containing one *Populus* gene and its *Arabidopsis* homolog is obvious (Figure 26). Until accurate *in vivo* evidence is discovered for plant laccases' function, laccases are clustered based on their phylogenetic relationship, the hypothesis being that genes that have similar sequences will also retain similar enzyme kinetics and substrates.

The genes in each phylogenetic group show similar expression patterns between the organs. The larger groups (1, 2, 3 and 5) are primarily expressed in the stems, the root, or ubiquitously throughout the plant. Interestingly, these four groups also contain the genes that are predicted targets of Cu-miRNAs *397* and *408*. Genes that are low-expressed are generally not potential targets of known miRNA down-regulation. Most likely this is because they have a low

impact on the copper pool and there is no selective pressure to conserve a miRNA target site on these laccase transcripts.

The findings presented here support the copper homeostasis model proposed in *Arabidopsis*. We propose that poplar, under copper-deficient conditions, up-regulates the SPL7 transcription factor, which up-regulates transcription of miRNAs *397* and *408*, which can then down-regulate large numbers of copper-containing proteins, including the laccases. This regulation then allows copper ions to be retained for high-importance proteins such as plastocyanin and cytochrome C-oxidase. Given the proposed role of laccases in lignin formation, this model accounts for the evidence of weakened xylem cell walls under copper-deficient conditions (Marschner, 2011).

This study provides insight into the size and complexity of the laccase family in woody plants. Although these genes have not been given a conclusive function, we may now begin understanding the scope of that task. Determining if laccases are indeed responsible for lignin formation in the cell wall is an essential factor in understanding coppers impact on cell wall formation. With the size and possible redundancy in the laccase gene family, classic forward genetics experiments will most likely have to be elaborate in order to learn about individual laccase gene's function. In the future, miRNA *397* and *408* over-expression poplar lines could be generated, effectively knocking out all highly expressed laccases and solidifying the laccase family's function in lignin formation.

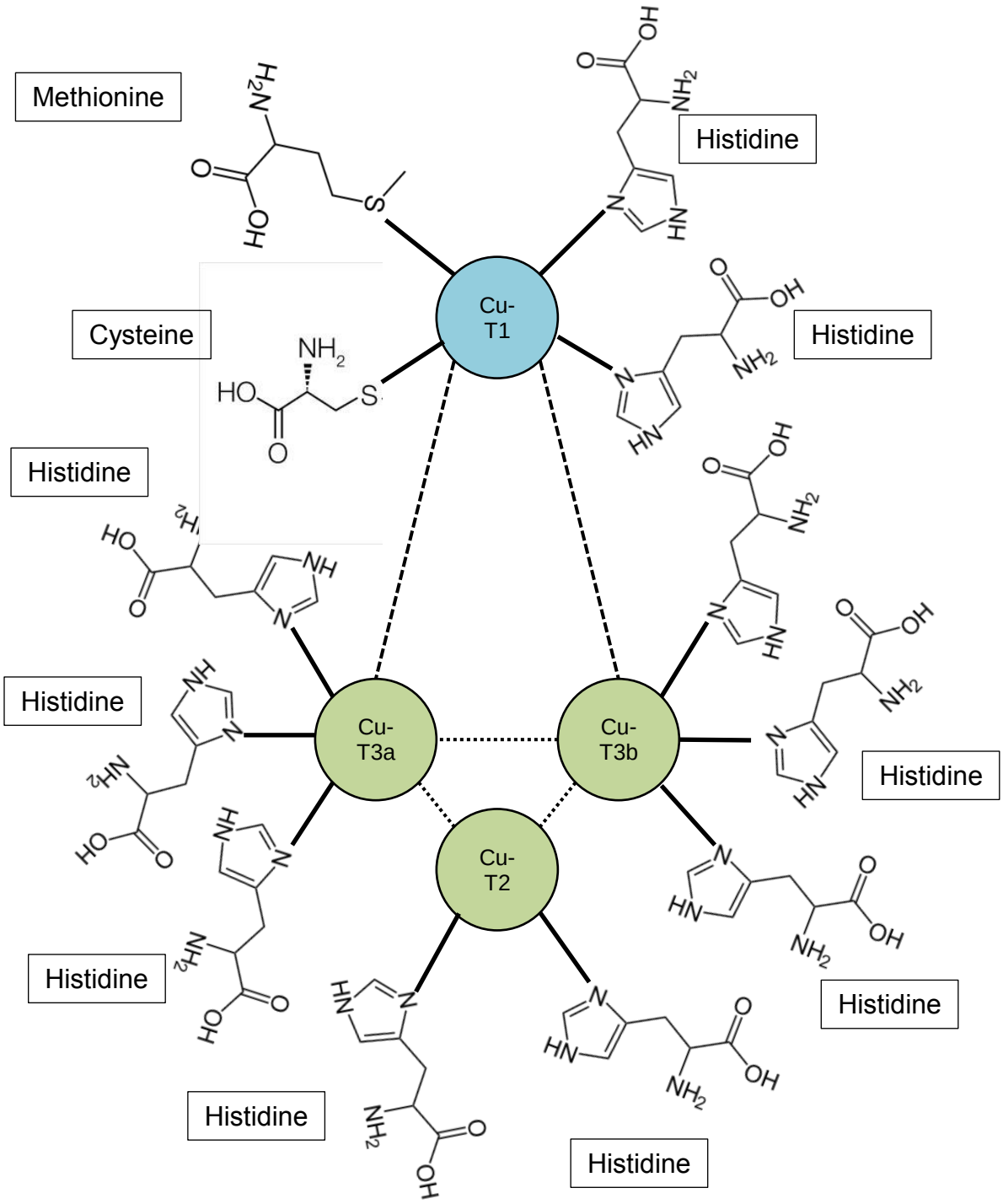

Figure 22: A diagram of the essential amino acids that coordinate the four copper atoms in the active site of laccases. The single T1copper is blue and the trinuclear cluster copper ions are green.

ENSEMBL Accession Number	Gene Name	Alternate Name	Full Length Protein (A.A.)	Signal Peptide length (A.A.)	Predicted Targeting
POPTR_0006s08740	popLAC1.1	LAC110b	580	33	Secreted
POPTR_0009s03940	popLAC1.2		576	32	Secreted
POPTR_0011s12090	popLAC1.3		570	20	Secreted
POPTR_0011s12100	popLAC1.4		581	31	Secreted
POPTR_0001s41160	popLAC1.5		581	31	Secreted
POPTR_0001s41170	popLAC1.6		581	31	Secreted
POPTR_0001s14010	popLAC1.7		581	31	Secreted
POPTR_0001s35740	popLAC1.8		580	30	Secreted
POPTR_0006s08780	popLAC1.9	LAC110a	579	25	Secreted
POPTR_0010s20050	popLAC2.1	LAC3	555	22	Secreted
POPTR_0008s06430	popLAC2.2	LAC2	556	23	Secreted
POPTR_0001s25580	popLAC2.3		554	21	Secreted
POPTR_0009s04720	popLAC2.4		556	23	Secreted
POPTR_0016s11960	popLAC2.5	LAC1b	557	22	Secreted
POPTR_0006s09840	popLAC2.6	LAC1d	550	87	Mitochondria
POPTR_0006s09830	popLAC2.7	LAC1c	560	23	Secreted
POPTR_0016s11950	popLAC2.8	LAC1a	557	22	Secreted
POPTR_0004s16370	popLAC2.9		540	--	--
POPTR_0009s10550	popLAC2.10		561	27	Secreted
POPTR_0007s13050	popLAC2.11		562	28	Secreted
POPTR_0010s19090	popLAC3.1	LAC90d	575	31	Secreted
POPTR_0008s07370	popLAC3.2	LAC90a	574	30	Secreted
POPTR_0010s19080	popLAC3.3	LAC90c	582	34	Secreted
POPTR_0008s07380	popLAC3.4	LAC90b	562	14	Secreted
POPTR_0013s14890	popLAC3.5		576	32	Secreted
POPTR_0019s14530	popLAC3.6		576	29	Secreted
POPTR_0005s22250	popLAC4.1		565	25	Secreted
POPTR_0005s22230	popLAC4.2		565	25	Secreted
POPTR_0005s22240	popLAC4.3		518	--	--
POPTR_0019s11860	popLAC4.4		552	29	Mitochondria
POPTR_0019s11810	popLAC4.5		552	29	Mitochondria
POPTR_0001s21380	popLAC4.6		563	24	Secreted
POPTR_0006s09520	popLAC5.1	LAC110c	562	22	Secreted
POPTR_0016s11520	popLAC5.2		569	22	Secreted
POPTR_0016s11500	popLAC5.3		568	22	Secreted
POPTR_0015s04340	popLAC6.1		579	27	Secreted
POPTR_0012s04620	popLAC6.2		579	27	Secreted
POPTR_0014s09610	popLAC7.1		529	--	--

Figure 23**:** *Populus* laccase genes discovered by amino acid sequence homology with *Arabidopsis* orthologs. Signal peptides and subcellular localization were predicted with TargetP. Alternate names for known *Populus* laccases are from Ranocha *et al.*, 1999.

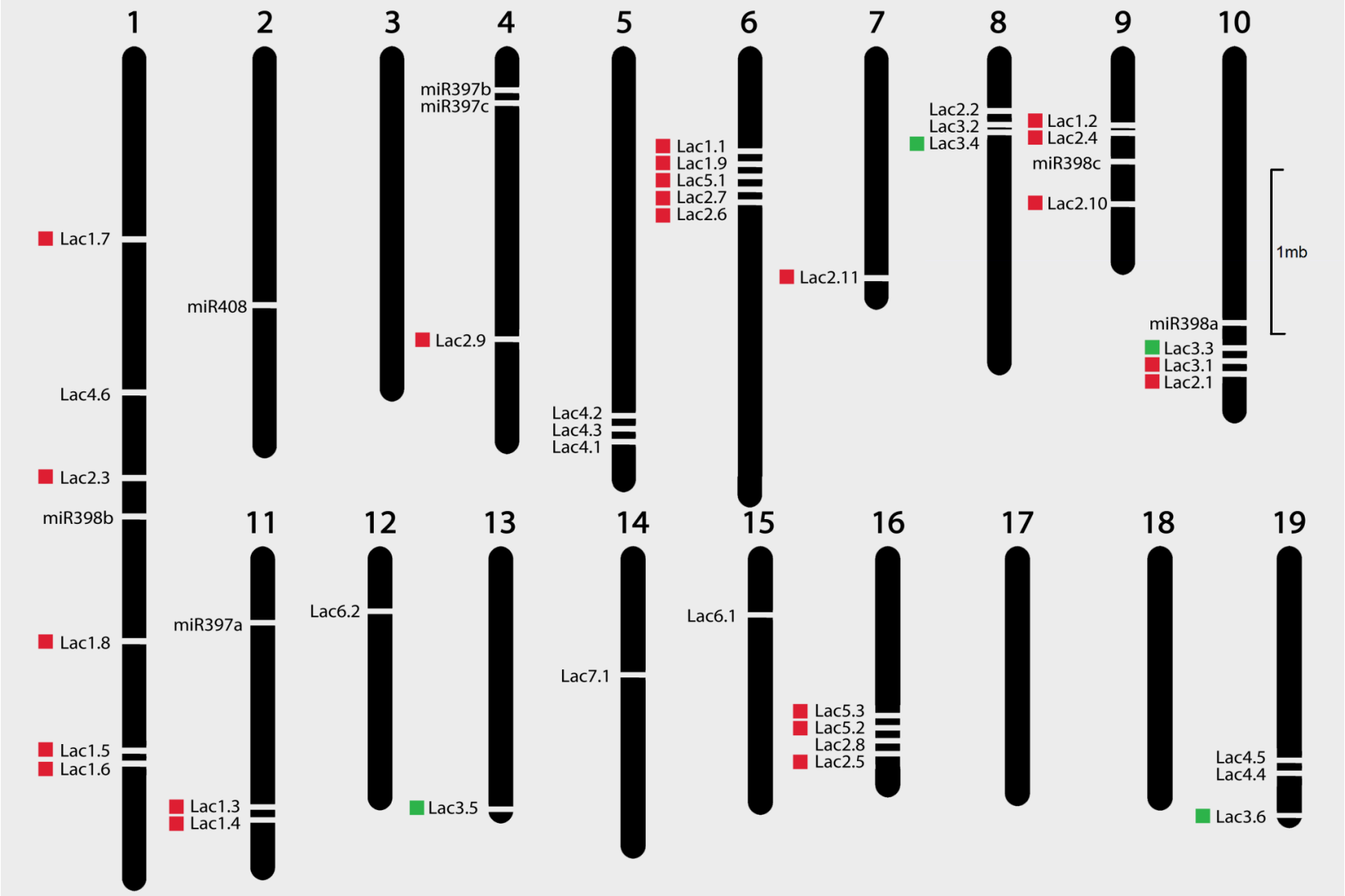

Figure 24: All of the *Populus* laccases and Cu-miRNAs mapped onto a diagram of the chromosomes. Genes with mRNAs targeted by miRNA *397* and miRNA *408* are indicated with a red box and green box respectively.

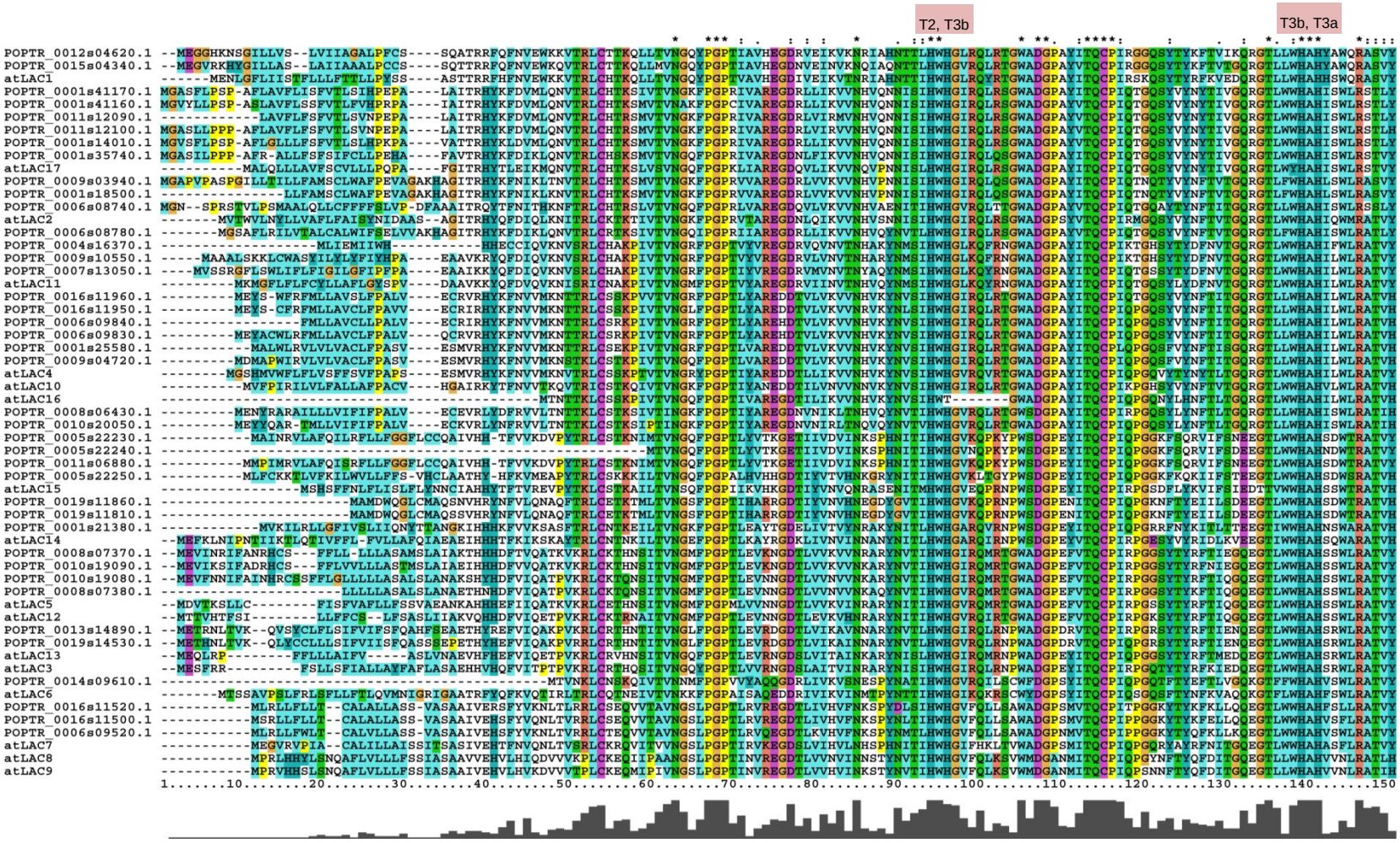

Figure 25: A ClustalW2 alignment of the 17 *Arabidopsis* laccases and 40 peptide sequences found in the *Populus* transcriptome. How conserved a residue is between all of the sequences is indicated by a grey bar at the bottom of the graph. An '*' indicates positions which have a single, fully conserved residue. A ':' indicates conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix. A '.' indicates conservation between groups of weakly similar properties - scoring ≤ 0.5 in the Gonnet PAM 250 matrix. Pink boxes at the top of the alignment indicate regions that bind the copper ions. The copper ion each histidine is thought to bind is in those boxes (Solomon *et al.*, 1996).

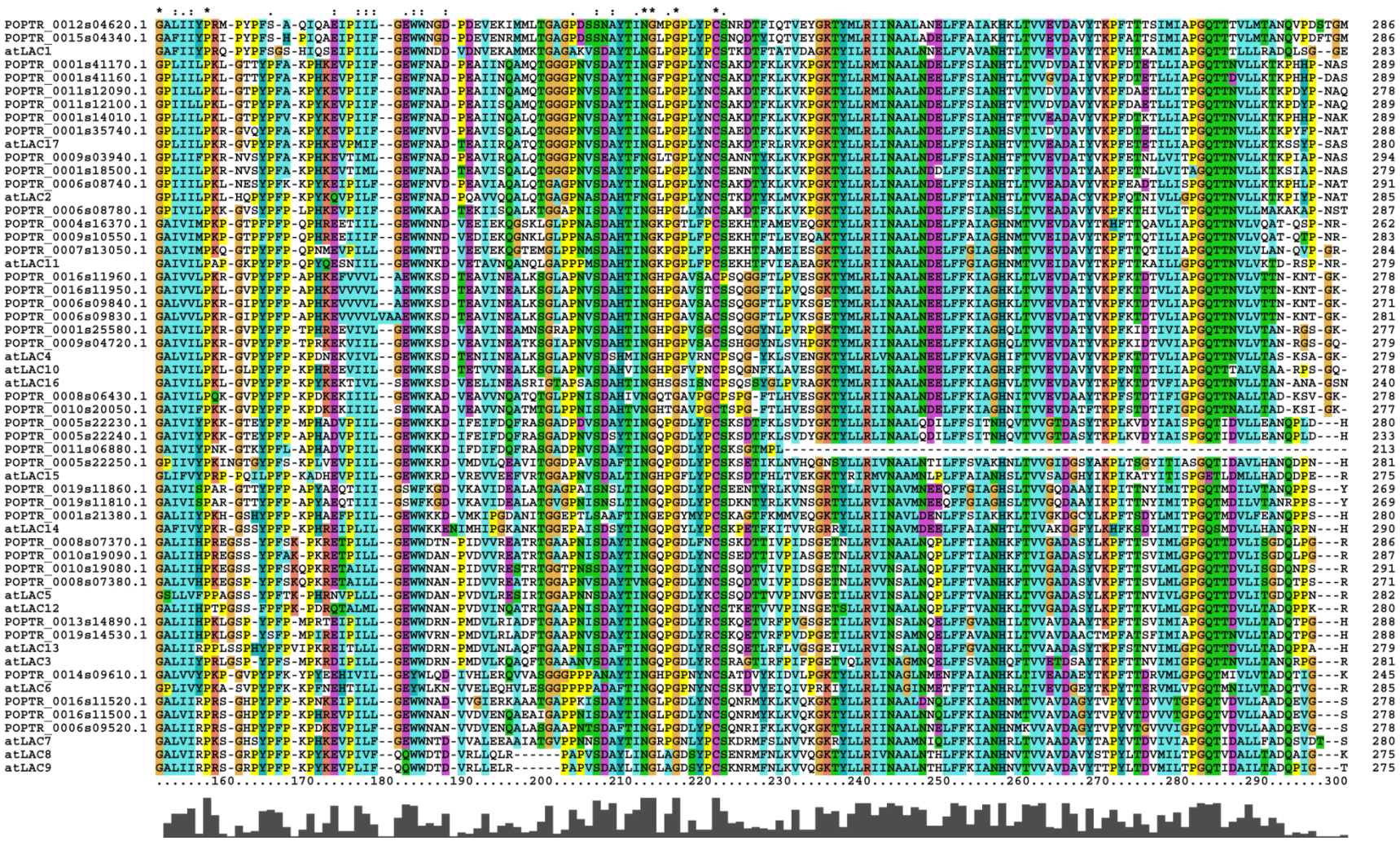

Fig 25: Continued

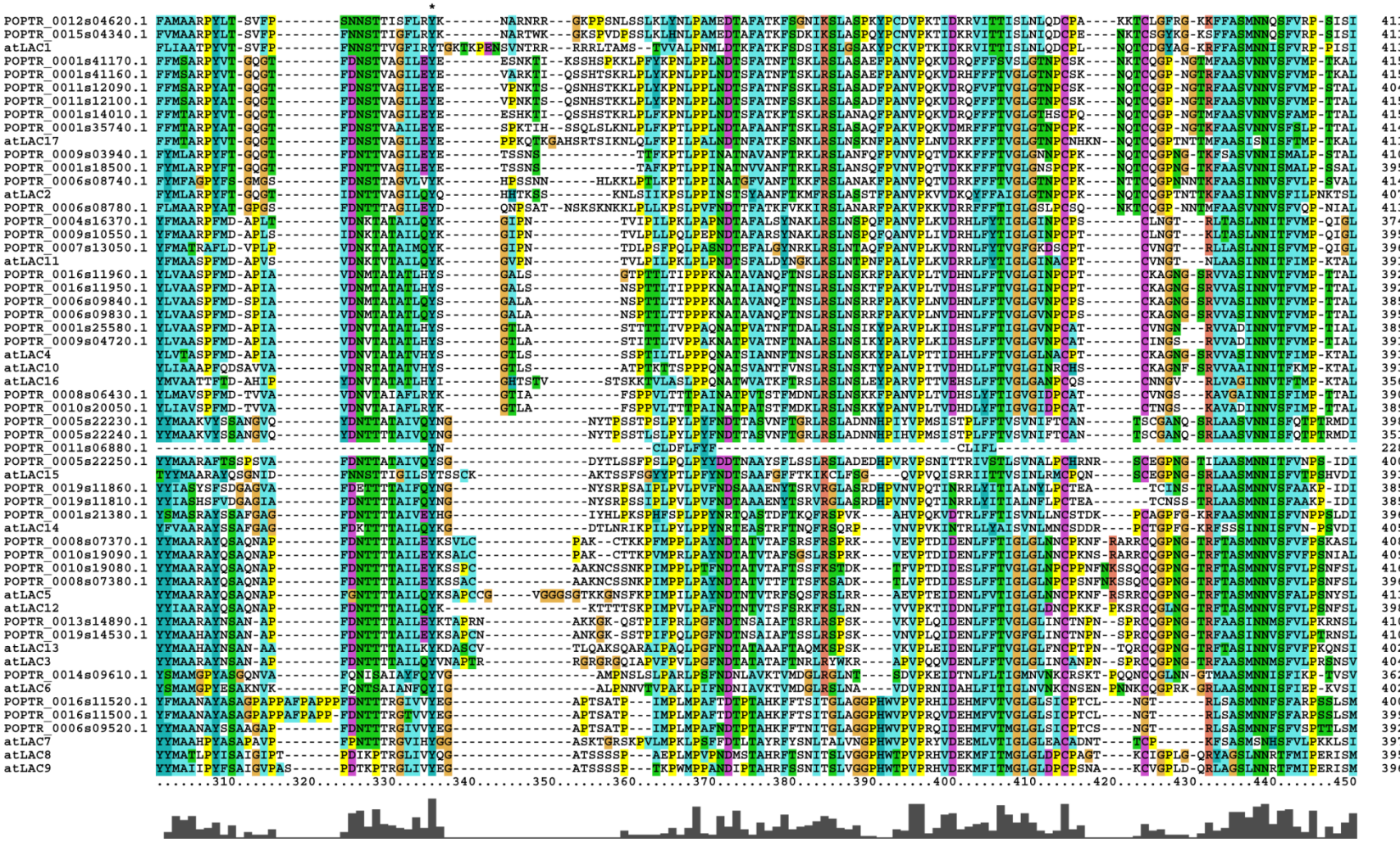

Fig 25: Continued

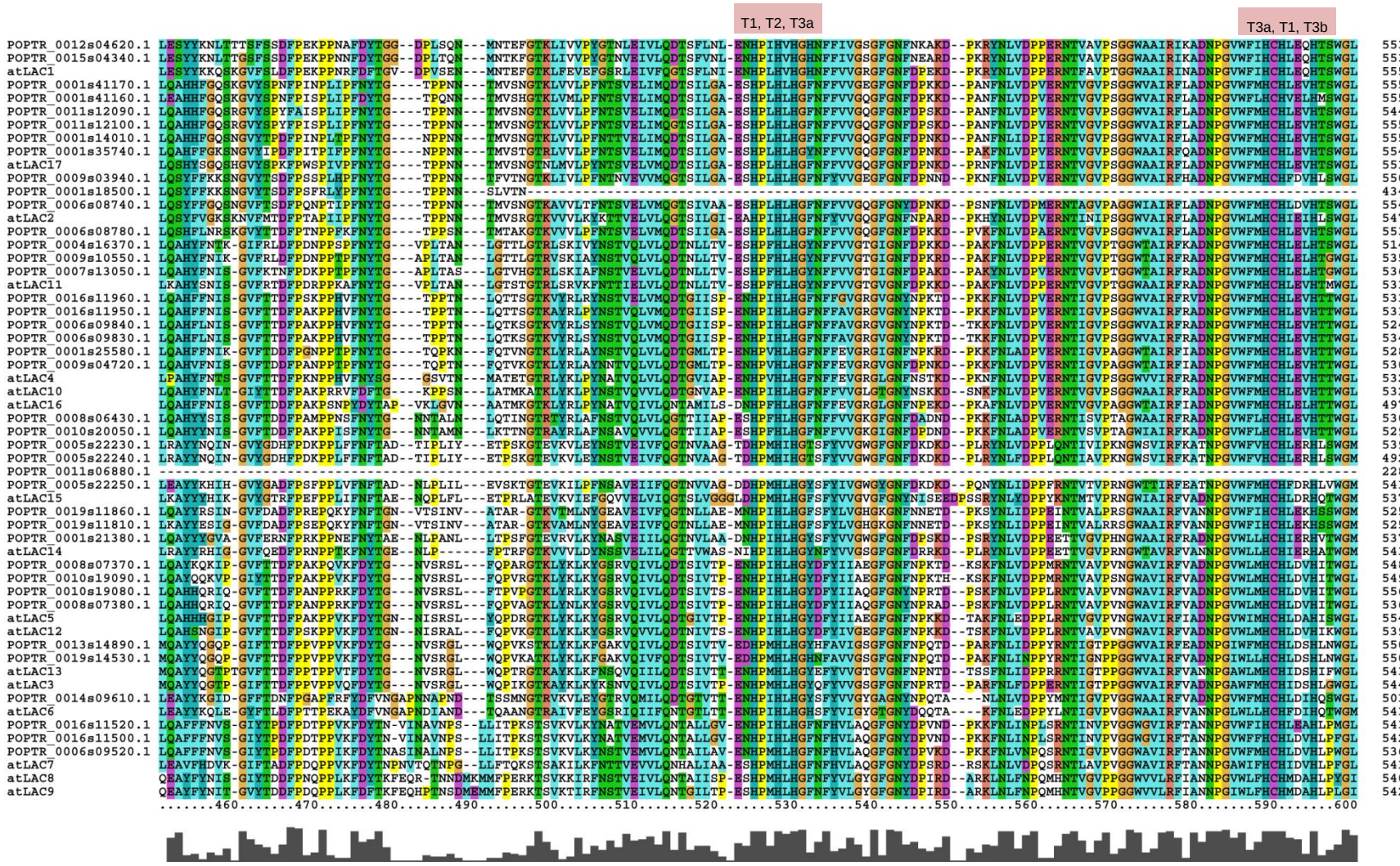

Fig 25: Continued

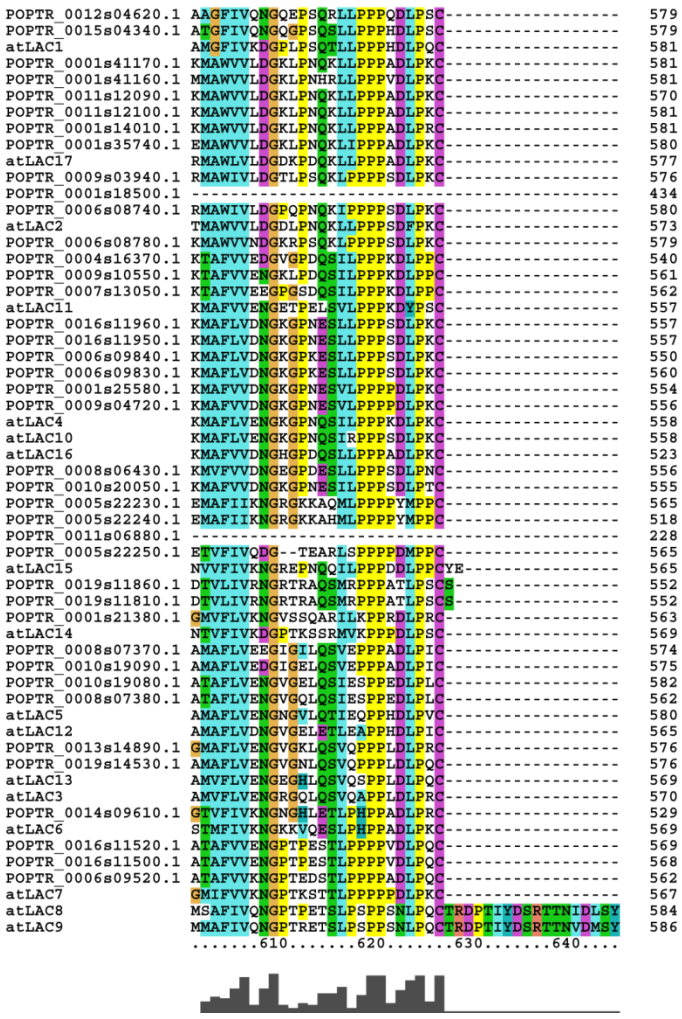

Fig 25: Continued

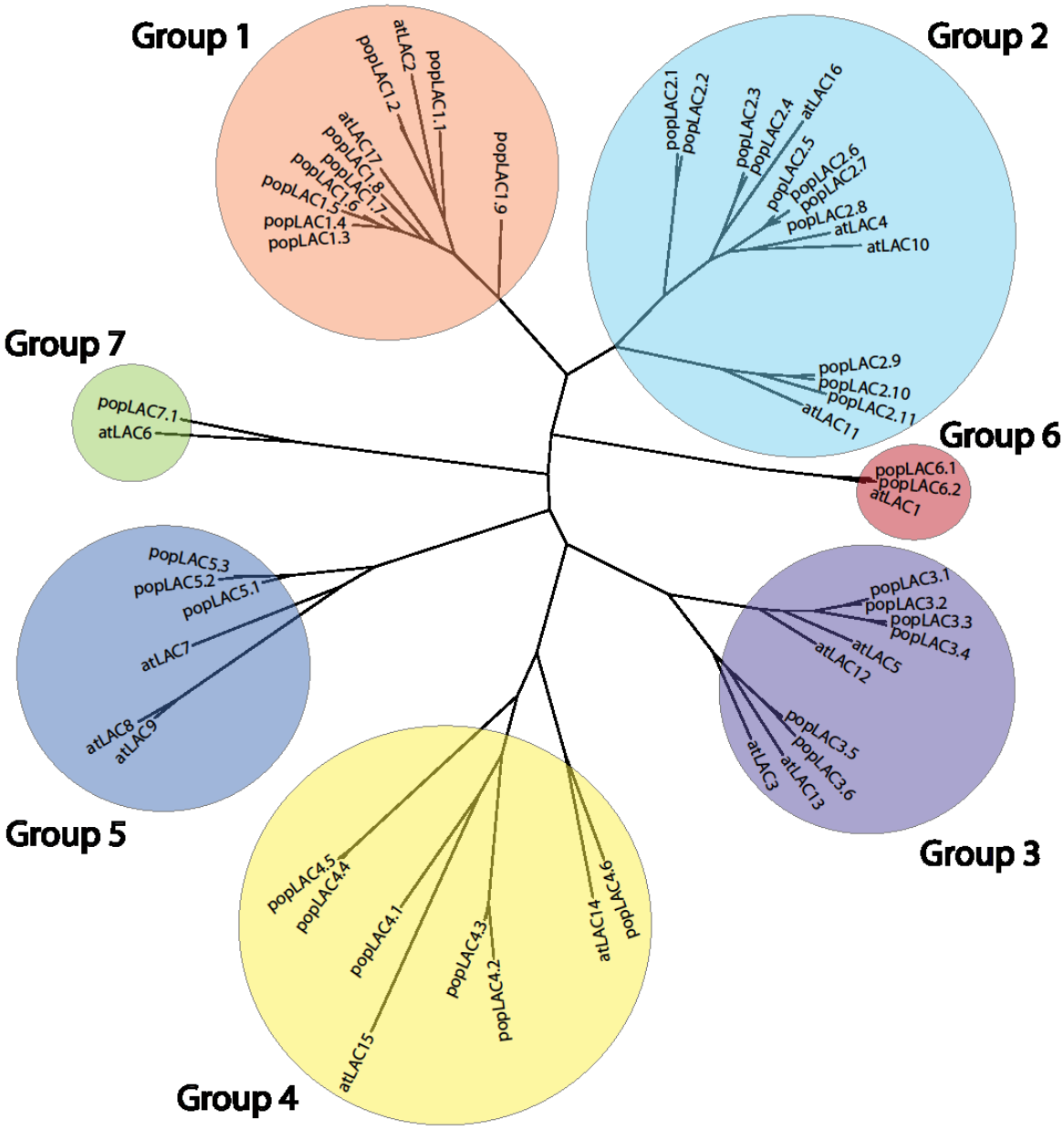

Figure 26: A radial tree showing the parsimony of the popLAC and the atLAC genes.

	Young leaves		Old Leaves		Stem		Root			
Gene Name	50 nM	Log2	50 nM	Log2	50 nM	Log2	50 nM	Log2	Target?	Number of exceptions
popLAC6.1	0.0	0.00	0.0	7.98	0.0	0.00	0.0	0.00		
popLAC6.2	0.0	0.00	0.0	-0.34	0.0	0.00	0.2	0.86		
popLAC3.1	0.8	0.62	0.9	-0.15	21.8	-1.73	11.3	-0.38	X	2
popLAC3.2	0.4	2.59	2.1	-0.24	67.8	-3.09	15.8	-0.65	X	2
popLAC3.3	0.7	1.45	1.8	-0.22	26.0	-2.44	5.6	-1.79	O	4.5
popLAC3.4	19.6	-0.94	25.3	-1.80	45.0	-2.25	27.1	-1.64	O	4.5
popLAC3.5	0.1	2.06	0.1	-1.14	10.4	-2.29	22.5	-2.39	O	4.5
popLAC3.6	2.5	0.17	2.9	-0.38	4.2	-0.40	4.5	-0.22	O	4
popLAC4.1	0.1	-1.51	0.0	0.00	0.6	0.35	0.3	0.58		
popLAC4.2	0.0	-0.58	0.0	-0.86	0.0	2.60	0.8	-0.19		
popLAC4.3	0.0	0.00	0.1	0.37	0.0	0.00	0.0	1.70		
popLAC4.4	0.0	1.46	0.2	-1.38	0.7	-0.13	1.7	0.57		
popLAC4.5	0.5	0.70	0.4	-0.06	5.6	-0.18	9.0	0.41		
popLAC4.6	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00		
popLAC5.1	0.0	1.19	0.0	2.55	0.8	-0.35	13.2	0.16	X	2.5
popLAC5.2	0.0	-1.52	0.0	0.00	0.0	1.49	0.5	-0.88		
popLAC5.3	0.1	-0.85	0.0	2.84	0.3	1.53	27.8	-0.24	X	2.5
popLAC7.1	0.8	-0.66	0.0	1.77	0.6	0.58	7.2	-0.98		
popLAC1.1	0.6	2.73	1.1	1.40	119.5	-1.28	21.0	-0.10	X	2.5
popLAC1.2	2.1	0.16	4.5	0.16	12.0	-1.60	9.5	-0.09	X	3
popLAC1.3	1.9	0.61	2.0	0.93	3.9	-1.43	1.5	-0.32	X	2.5
popLAC1.4	5.3	1.35	34.9	0.27	16.4	-1.27	8.7	-0.32	X	2.5
popLAC1.5	0.0	0.00	0.0	0.00	0.0	0.00	0.0	10.24	X	2
popLAC1.6	5.5	-0.48	2.9	-1.18	22.2	-0.76	11.6	0.05	X	2
popLAC1.7	1.8	-2.21	0.5	-3.24	6.2	-3.98	4.6	-2.82	X	1.5
popLAC1.8	1.4	-0.52	0.2	-1.39	2.6	-1.33	0.5	-2.03		
popLAC1.9	0.4	2.21	0.9	1.17	27.0	-0.28	1.6	0.43	X	3.5
popLAC2.1	0.4	3.26	2.9	1.72	57.8	-0.04	4.7	0.86		
popLAC2.2	0.1	4.70	1.7	1.61	44.7	-0.21	2.8	0.97		
popLAC2.3	0.6	0.31	1.1	0.09	3.0	-0.97	5.3	-0.05	X	2
popLAC2.4	1.0	1.50	3.0	0.39	47.6	-2.15	22.2	-1.46	X	1.5
popLAC2.5	1.0	-0.40	0.2	-1.19	1.7	-1.57	0.6	-0.51	X	3
popLAC2.6	1.2	0.50	2.3	0.03	32.5	-1.81	3.1	-0.18	X	2.5
popLAC2.7	0.9	0.96	3.0	-0.18	19.3	-1.25	3.5	-0.26	X	2.5
popLAC2.8	3.9	0.78	27.1	-0.30	53.1	-2.06	7.3	-1.06	X	1.5
popLAC2.9	0.7	-0.38	0.2	-1.12	1.7	-1.81	1.8	-0.34	X	2
popLAC2.10	5.1	-2.15	0.4	-3.96	2.6	-2.84	3.6	-1.21	X	2
popLAC2.11	8.9	0.03	2.0	-0.44	7.6	-0.64	9.2	0.03	X	3

Figure 27: A table showing the expression (RPKM) of the 38 *Populus* laccase genes in the four organs and their differential expression in copper-deficient conditions (Log2). Genes with high expression in copper-sufficient conditions are grey, while genes with low expression are white. Genes with a decrease in expression in copper-deficient conditions are blue, while genes with an increased expression are red. An 'X' represents a target of Cu-miRNA *397*; a 'O' represents a target of Cu-miRNA *408*. The number of exceptions between the miRNA and the mRNA target is also included.

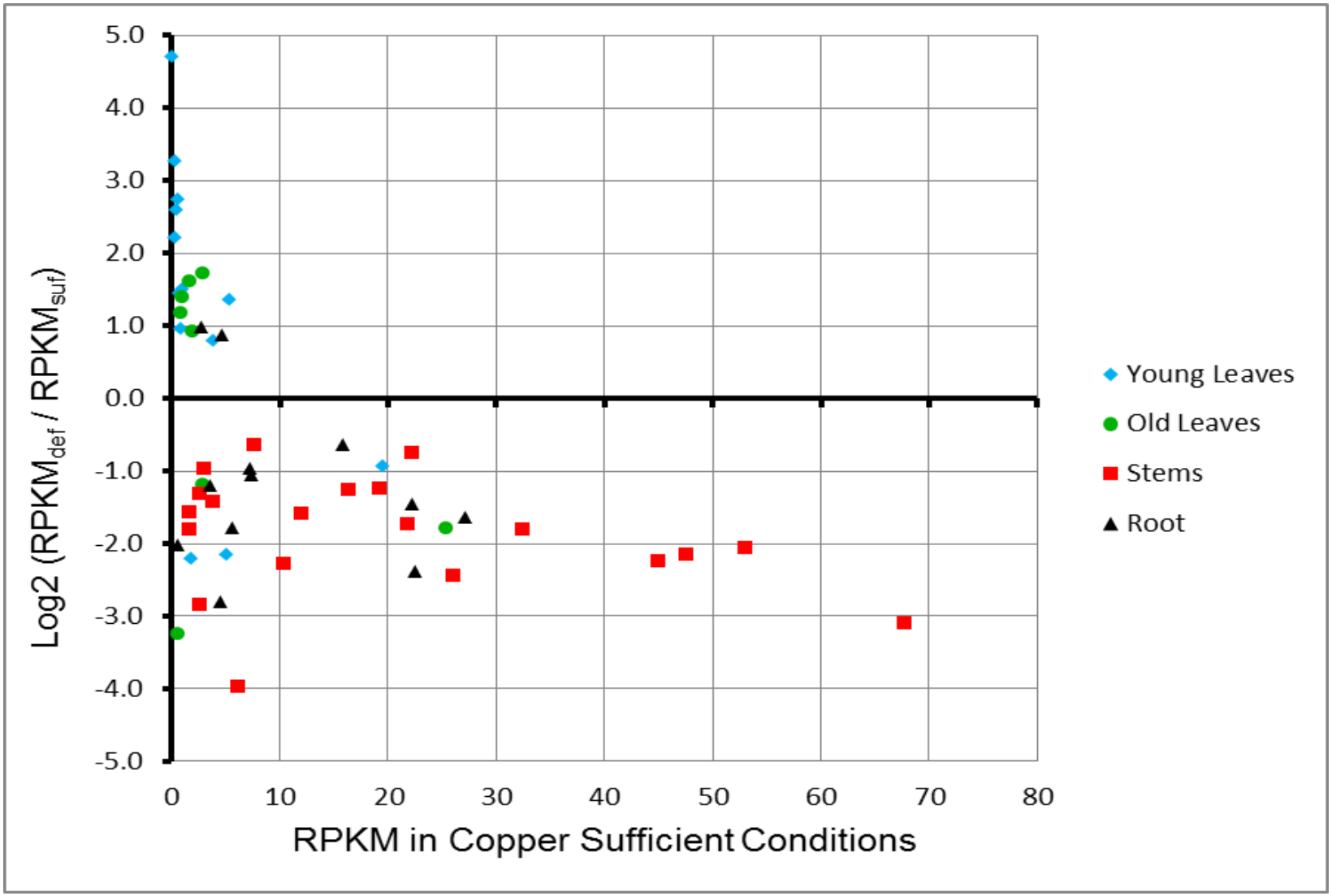

Figure 28: Copper-regulated *Populus* laccase genes with significant p-values ($P < 0.02$)

CHAPTER 4

SUMMARIZING DISCUSSION

In plants, copper homeostasis involves the coordination of expression of a large number of genes. These genes are not limited to genes directly involved with copper binding or transport since many Cu-regulated genes are implicated in stress response, iron homeostasis, and sulfur homeostasis pathways. Until recently our understanding of copper homeostasis has been limited to the model organism *Arabidopsis thaliana*, a small, herbaceous dicot with a relatively simple genome. To expand our knowledge of copper homeostasis we are now looking in more complicated organisms such as *Populus trichocarpa* (Black Poplar). Poplar is a large, perennial dicot with much more elaborate and complicated genomes than *Arabidopsis*. This complexity means that its response mechanisms to copper supply can also be more complex and contain novel mechinisms of homeostasis.

Many of the mechanisms of copper homeostasis have been discovered using molecular and genetics techniques, but recently next-generation sequencing has made transcriptomic experiments an attractive alternative. RNA-SEQ allows us to quantify the relative abundance of all transcripts in a transcriptome and compare these levels between conditions. With this technique we are able to get a complete understanding of a plant's response to copper deficiency and begin focusing new studies in the right direction from the outset. In our study we applied RNA-SEQ to four vegetative organs of poplar under copper-deficient conditions and compared their transcriptomes to plants grown in copper-sufficient conditions.

Given our previous understanding of copper homeostasis, what we found was surprising. First of all, 17.2% of the transcriptome is significantly differentially expressed under copper-deficient conditions, in at least one organ. The transcript responses range from a $\log_2(1)$ to $\log_2(12)$ increase in transcript abundance for some genes in some organs. The makeup of the

transcriptomic response is also variable between the organs. Young leaves, old leaves, and stems, for the most part, up-regulate as many genes as they down regulate in response to copper deficiency. However, the roots down-regulated a significantly larger number of genes compared to the number of genes up-regulated. Many of the down-regulated genes in roots are photosynthetic genes. These findings suggest that models of copper homeostasis need to be tailored to explain organ-specific, if not tissue specific gene regulation.

Laccase genes were among the top ten most down-regulated genes in every organ. Laccases are multi-copper oxidases with the ability to polymerize the formation of complex phenols *in vitro*. This ability has implicated laccases as a key component of lignin formation, but *in vivo* evidence supporting this has been elusive. It has been shown in the past that copper deficiency results in defects in lignin formation, and we have shown that laccases are copper proteins that are down-regulated in response to copper deficiency. These facts taken together suggest a possible mechanistic link between copper supply, the laccases, and lignin formation. Poplar is an excellent organism to study this possibility because it has extensive lignin formation, extensive secondary cell wall production, and a sequenced genome. This research would be of especial interest to the paper and biofuel industries, which look to minimize lignin in their products.

The first step to studying the laccase gene family in poplar is the discovery and annotation of all the family members. In this work, we describe twenty-five new candidate genes, discovered by sequence homology, to add to the thirteen already discovered genes. The thirty-eight genes contain short conserved motifs used for binding four copper ions. The laccase family can be divided into seven groups based on their amino acid sequence similarity; this is

one more group than has been described in the literature, a result of using more genes in the analysis.

In *Arabidopsis*, many laccase genes have been shown by 5' Rapid amplification of cDNA Ends (RACE) to be targets of the Cu-miRNAs *397* and *408.* The thirty-eight poplar laccases were searched for predicted Cu-miRNA target regions; twenty one-candidates for *miR397* and four candidates for *miR408* were found. The potential targets were found in groups 1, 2, 3, and 5, while groups 4, 6 and 7 had no potential targets. When expression data from the RNA-SEQ experiment is overlaid onto a phylogenetic tree of the laccases, we see the high expressed laccases are in groups 1, 2, 3, and 5. We also see that under copper-deficient conditions, highly expressed laccases that are also potential targets for Cu-miRNA mediated down-regulation are indeed significantly down-regulated. This discovery supports the copper economy model and miRNA targeting predictions. It would appear that only genes with high biological impact have enough selective pressure to maintain regulatory mechanisms. In the future, miRNA *397* and *408* over-expression poplar lines could be generated, effectively knocking out all highly expressed laccases and solidifying the laccase family's function in lignin formation.

BIBLIOGRAPHY

Abdel-Ghany, S. E., & Pilon, M. (2008). MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *The Journal of Biological Chemistry*, *283*(23), 15932–15945. doi:10.1074/jbc.M801406200

Abdel-Ghany, S., Müller-Moulé, P., Niyogi, K. K., Pilon, M., & Shikanai, T. (2005). Two P-type ATPases are required for copper delivery in Arabidopsis thaliana chloroplasts. *The Plant Cell …*, *17*(April), 1233–1251. doi:10.1105/tpc.104.030452.In

Andrés-Colás, N., Sancenón, V., Rodríguez-Navarro, S., Mayo, S., Thiele, D. J., Ecker, J. R., Puig, S., et al. (2006). The Arabidopsis heavy metal P-type ATPase, HMA5, interacts with metallochaperones and functions in copper detoxification of roots. *The Plant journal*, *45*, 225–236. doi:10.1111/j.1365-313X.2005.02601.x

Aramayo, R., & Timberlake, E. W. (1990). Sequence and molecular struture of the Aspergilus nidulans yA (laccase I) gene. *Nucleic Acid Research*, *18*(11), 3415–3417.

Ascencio-Ibáñez, J. T., Sozzani, R., Lee, T.-J., Chu, T.-M., Wolfinger, R. D., Cella, R., & Hanley-Bowdoin, L. (2008). Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant physiology*, *148*(1), 436–54. doi:10.1104/pp.108.121038

Bao, W., O'Malley, D. M., Whetten, R., & Sederoff, R. R. (1993). A laccase associated with lignification in loblolly pine xylem. *Science*, *260*(5108), 672–974. doi:10.1126/science.260.5108.672

Bernal, M., Casero, D., Singh, V., Wilson, G. T., Grande, A., Yang, H., Dodani, S. C., et al. (2012). Transcriptome sequencing identifies SPL7-regulated copper acquisition genes FRO4/FRO5 and the copper dependence of iron homeostasis in Arabidopsis. *The Plant cell*, *24*(2), 738–61. doi:10.1105/tpc.111.090431

Bernal, M., Roncel, M., Ortega, J. M., Picorel, R., & Yruela, I. (2004). Copper effect on cytochrome b of photosystem II under photoinhibitory conditions. *Physiologia plantarum*, *120*(4), 686–694. doi:10.1111/j.1399-3054.2004.0286.x

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., Borrega, N., et al. (2011). Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of Arabidopsis thaliana stems. *The Plant Cell*, *23*(3), 1124–1137. doi:10.1105/tpc.110.082792

Bonnet, E., Wuyts, J., Rouzé, P., & Van de Peer, Y. (2004). Detection of 91 potential conserved plant microRNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(31), 11511–6. doi:10.1073/pnas.0404025101

Briat, J.-F., Curie, C., & Gaymard, F. (2007). Iron utilization and metabolism in plants. *Current opinion in plant biology*, *10*(3), 276–82. doi:10.1016/j.pbi.2007.04.003

Burkhead, J. L., Reynolds, K. a G., Abdel-Ghany, S. E., Cohu, C. M., & Pilon, M. (2009). Copper homeostasis. *The New Phytologist*, *182*(4), 799–816. doi:10.1111/j.1469-8137.2009.02846.x

Butt, V. S. (1980). Direct oxidases and related enzymes. *Plant Physiology and Biochemisty*, *2*, 81–123.

Cardon, G., Höhmann, S., Klein, J., Nettesheim, K., Saedler, H., & Huijser, P. (1999). Molecular characterisation of the Arabidopsis SBP-box genes. *Gene*, *237*(1), 91–104.

Carr, H. S., & Winge, D. R. (2003). Assembly of cytochrome c oxidase within the mitochondrion. *Accounts of chemical research*, *36*(5), 309–16. doi:10.1021/ar0200807

Charron, J.-B. F., Ouellet, F., Houde, M., & Sarhan, F. (2008). The plant Apolipoprotein D ortholog protects Arabidopsis against oxidative stress. *BMC plant biology*, *8*, 86. doi:10.1186/1471-2229-8-86

Claus, H. (2003). Laccases and their occurrence in prokaryotes. *Archives of Microbiology*, *179*(3), 145–150. doi:10.1007/s00203-002-0510-7

Claus, H. (2004). Laccases: structure, reactions, distribution. *Micron*, *35*(1-2), 93–96. doi:10.1016/j.micron.2003.10.029

Cohu, C. M., & Pilon, M. (2010). Cell Biology of Copper. In R. Hell & R.-R. Mendel (Eds.), *Cell biology of metals and nutrients* (Vol. 17, pp. 55–74). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-10613-2

Davin, L. B., Bedgar, D. L., Katayama, T., & Lewis, N. G. (1992). On the stereoselective synthesis of (+)-pinoresinol in Forsythia suspensa from its achiral precursor, coniferyl alcohol. *Phytochemistry*, *31*(11), 3869–74.

Dean, J., LaFayette, P., Tristram, A., Hoopes, J., Eriksson, K.-E. L., & Merkle, S. (1998). *Laccases associated with lignifying vascular tissues*. (N. Lewis & S. Sarkanen, Eds.) (pp. 96–108). ACS Symposium Series.

Driouich, A., Lainé, A., Vian, B., & Faye, L. (1992). Characterization and localization of laccase forms in stem and cell cultures of sycamore. *The Plant Journal*, *2*(1), 13–24.

Ducros, V., Brzozowski, A., Wilson, K., Brown, S., Ostergaard, P., Schneider, P., Yaver, D., et al. (1998). Crystal structure of the type-2 Cu depleted laccase from Coprinus cinereus at 2.2 Å resolution. *Nature Structural Biology*, *5*(4), 310–316. doi:0.1038/nsb0498-310

Dykema, P. E., Sipes, P. R., Marie, A., Biermann, B. J., Crowell, D. N., & Randall, S. K. (1999). A new class of proteins capable of binding transition metals. *Plant molecular biology*, *41*(1), 139–50.

Eggert, C., Temp, U., & Eriksson, K.-E. L. (1997). Laccase is essential for lignin degradation by the white-rot fungus Pycnoporus cinnabarinus. *FEBS Letters*, *407*(1), 89–92. doi:10.1016/S0014-5793(97)00301-3

Epstein, E., & Bloom, A. (2005). *Mineral Nutrion in Plant: Principles and Perspectives*. Sunderland, MA: 2nd ed. Sinauer Associates.

Food and Agriculture Orginization. (2007). *State of the World's Forests* (pp. 2–27).

Freudenberg, K. (1959). Biosynthesis and constitution of lignin. *Nature*, *183*, 1152–1155. doi:10.1038/1831152a0

Garcia-Molina, A., Andrés-Colás, N., Perea-García, A., Del Valle-Tascón, S., Peñarrubia, L., & Puig, S. (2011). The intracellular Arabidopsis COPT5 transport protein is required for photosynthetic electron transport under severe copper deficiency. *The Plant journal : for cell and molecular biology*, *65*(6), 848–60. doi:10.1111/j.1365-313X.2010.04472.x

Halliwell, B., & Gutteridge, J. (1984). Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochemical Journal*, *219*, 1–14.

Irshad, M., Canut, H., Borderies, G., Pont-Lezica, R., & Jamet, E. (2008). A new picture of cell wall protein dynamics in elongating cells of Arabidopsis thaliana: confirmed actors and newcomers. *BMC plant biology*, *8*, 94. doi:10.1186/1471-2229-8-94

Jones-Rhoades, M. W., & Bartel, D. P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular cell*, *14*(6), 787–99. doi:10.1016/j.molcel.2004.05.027

Jung, H., Gayomba, S. R., Rutzke, M. a, Craft, E., Kochian, L. V, & Vatamaniuk, O. K. (2012). COPT6 is a plasma membrane transporter that functions in copper homeostasis in Arabidopsis and is a novel target of SQUAMOSA promoter-binding protein-like 7. *The Journal of biological chemistry*, *287*(40), 33252–33267. doi:10.1074/jbc.M112.397810

Kampfenkel, K., Kushnir, S., Babiychuk, E., Inzé, D., & Van Montagu, M. (1995). Molecular characterization of a putative Arabidopsis thaliana copper transporter and its yeast homologue. *The Journal of biological chemistry*, *270*(47), 28479–86.

Kieselbach, T., Hagman, Å., Andersson, B., & Schro, W. P. (1998). The Thylakoid Lumen of Chloroplasts. *The Journal of Biological Chemistry*, *273*(12), 6710–6716.

Kropat, J., Tottey, S., Birkenbihl, R. P., Depège, N., Huijser, P., & Merchant, S. (2005). A regulator of nutritional copper signaling in Chlamydomonas is an SBP domain protein that

recognizes the GTAC core of copper response element. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18730–5. doi:10.1073/pnas.0507693102

Kumar, S. V. S., Phale, P. S., Durani, S., & Wangikar, P. P. (2003). Combined sequence and structure analysis of the fungal laccase family. *Biotechnology and Bioengineering*, *83*(4), 386–94. doi:10.1002/bit.10681

Linder, M. C., & Goode, C. A. (1991). *Biochemistry of Copper*. New York, New York, USA: Plenum Press.

Lippard, S., & Berg, J. (1994). *Principles of Bioinorganic Chemistry*. New York, New York, USA: University Science Books.

Liu, L., Dean, J. F. D., Friedman, W. E., & Eriksson, K.-E. L. (1994). A laccase-like phenoloxidase is correlated with lignin biosynthesis in Zinnia elegans stem tissues. *The Plant Journal*, *6*(2), 213–224. doi:10.1046/j.1365-313X.1994.6020213.x

Lu, S., Sun, Y., Shi, R., Clark, C., Li, L., & Chiang, V. L. (2005). Novel and mechanical stress–responsive microRNAs in Populus trichocarpa that are absent from Arabidopsis. *The Plant Cell*, *17*(August), 2186–2203. doi:10.1105/tpc.105.033456.1

Malmstrom, B. G., Andreasson, L. E., & Reinhammar, B. (1975). *The Enzymes* (p. 507).

Marschner, H. (2011). *Mineral nutrition of higher plants* (pp. 333–347). Academic Press.

Mayer, A. M. (2006). Polyphenol oxidases in plants and fungi: going places? A review. *Phytochemistry*, *67*(21), 2318–31. doi:10.1016/j.phytochem.2006.08.006

Mayer, A. M., & Staples, R. C. (2002). Laccase: new functions for an old enzyme. *Phytochemistry*, *60*(6), 551–65.

McCaig, B. C., Meagher, R. B., & Dean, J. F. D. (2005). Gene structure and molecular analysis of the laccase-like multicopper oxidase (LMCO) gene family in Arabidopsis thaliana. *Planta*, *221*(5), 619–36. doi:10.1007/s00425-004-1472-6

McDougall, G. J., & Morrison, L. M. (1996). Extraction and partial purification of cell-wall-associated coniferyl alcohol oxidase from developing xylem of sitka spruce. *Holzforschung*, *50*(6), 549–553. doi:10.1515/hfsg.1996.50.6.549

O'Malley, D. M., Whetten, R., Bao, W., Chen, C.-L., & Sederoff, R. R. (1993). The role of laccase in lignification. *The Plant Journal*, *4*(5), 751–757. doi:10.1046/j.1365-313X.1993.04050751.x

Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., & Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature*, *425*(6955), 257–63. doi:10.1038/nature01958

Patsikka, E., Kairavuo, M., Frantisek, S., Aro, E., & Tyystjarvi, E. (2002). Excess copper predisposes photosystem II to photoinhibition in vivo by outcompeting iron and causing decrease in leaf chlorophyll 1. *Plant Physiology*, *129*(July), 1359–1367. doi:10.1104/pp.004788.1

Pich, A., & Scholz, G. (1996). Translocation of copper and other micronutrients in tomato plants (Lycopersicon esculentum Mill.): nicotianamine-stimulated copper transport in the xylem. *Journal of Experimental Botany*, *47*(294), 41–47.

Pilon, M. (2011). Moving copper in plants. *New Phytologist*, *192*, 305–307.

Pourcel, L., Routaboul, J.-M., Herhoas, L., Caboche, M., Lepiniec, L., & Debeaujon, I. (2005). Transparent testa 10 encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in Arabidopsis seed coat. *The Plant Cell …*, *17*(November), 2966–2980. doi:10.1105/tpc.105.035154.1

Puig, S., Mira, H., Dorcey, E., Sancenón, V., Andrés-Colás, N., Garcia-Molina, A., Burkhead, J. L., et al. (2007). Higher plants possess two different types of ATX1-like copper chaperones. *Biochemical and biophysical research communications*, *354*(2), 385–90. doi:10.1016/j.bbrc.2006.12.215

Ranocha, P., McDougall, G., Hawkins, S., Sterjiades, R., Borderies, G., Stewart, D., Cabanes-Macheteau, M., et al. (1999). Biochemical characterization, molecular cloning and expression of laccases - a divergent gene family - in poplar. *European journal of biochemistry / FEBS*, *259*(1-2), 485–95.

Ravet, K., Danford, F. L., Dihle, A., Pittarello, M., & Pilon, M. (2011). Spatiotemporal analysis of copper homeostasis in Populus trichocarpa reveals an integrated molecular remodeling for a preferential allocation of copper to plastocyanin in the chloroplasts of developing leaves. *Plant physiology*, *157*(3), 1300–12. doi:10.1104/pp.111.183350

Reinhammar, B., & Malmstroem, B. (1981). "Blue"' copper-containing oxidases. *Spiro TG (ed) Copper proteins*, *3*, 109–149.

Richardson, A., Duncan, J., & McDougall, G. J. (2000). Oxidase activity in lignifying xylem of a taxonomically diverse range of trees: identification of a conifer laccase. *Tree physiology*, *20*(15), 1039–47.

Ruiter, J. H. (1969). Suspected copper deficiency in radiata pine. *Plant and Soil*, *31*(1), 197–200. doi:10.1080/00382167.1969.9629214

Salas, B. S. D., Bennett, J. E., Kwon-Chung, K. J., & Perfect, J. R. (1996). Effect of the laccase gene, CNLAC1, on virulence of Cryptococcus neoforman. *Journal of Experimental Medicine*, *184*(August), 377–386.

Sancenón, V., Puig, S., Mira, H., Thiele, D. J., & Peñarrubia, L. (2003). Identification of a copper transporter family in Arabidopsis thaliana. *Plant molecular biology*, *51*(4), 577–87.

Sawchuk, M. G., Donner, T. J., Head, P., & Scarpella, E. (2008). Unique and overlapping expression patterns among members of photosynthesis-associated nuclear gene families in Arabidopsis. *Plant physiology*, *148*(4), 1908–24. doi:10.1104/pp.108.126946

Schubert, M., Petersson, U. a, Haas, B. J., Funk, C., Schröder, W. P., & Kieselbach, T. (2002). Proteome map of the chloroplast lumen of Arabidopsis thaliana. *The Journal of biological chemistry*, *277*(10), 8354–65. doi:10.1074/jbc.M108575200

Silva, D. De, Askwith, C., Eide, D., & Kaplan, J. (1995). The FET3 gene product required for high affinity iron transport in yeast is a cell surface ferroxidase. *Journal of Biological Chemistry*, *270*(January), 1098–1101.

Solomon, E. I., Sundaram, U. M., & Machonkin, T. E. (1996). Multicopper oxidases and oxygenases. *Chemical reviews*, *96*(7), 2563–2606.

Sterjiades, R., Dean, J. F., & Eriksson, K. E. (1992). Laccase from Sycamore Maple (Acer pseudoplatanus) Polymerizes Monolignols. *Plant physiology*, *99*(3), 1162–8.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., et al. (2004). Mapman: a User-Driven Tool To Display Genomics Data Sets Onto Diagrams of Metabolic Pathways and Other Biological Processes. *The Plant Journal*, *37*(6), 914–939. doi:10.1111/j.1365-313X.2004.02016.x

Turlapati, P., Kim, K.-W., Davin, L., & Lewis, N. (2011). The laccase multigene family in Arabidopsis thaliana: towards addressing the mystery of their gene function(s). *Planta*, *233*(3), 439–70. doi:10.1007/s00425-010-1298-3

Tuskan, G. a, Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., et al. (2006). The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science (New York, N.Y.)*, *313*(5793), 1596–604. doi:10.1126/science.1128691

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews*, *10*(1), 57–63. doi:10.1038/nrg2484

Weigel, M., Varotto, C., Pesaresi, P., Finazzi, G., Rappaport, F., Salamini, F., & Leister, D. (2003). Plastocyanin is indispensable for photosynthetic electron flow in Arabidopsis thaliana. *The Journal of biological chemistry*, *278*(33), 31286–9. doi:10.1074/jbc.M302876200

Welch, R. M., Norvell, W. A., Schaefer, S. C., Shaft, J. E., & Kochian, L. V. (1993). Induction of iron (III) and copper (II) reduction in pea (Pisum satirum L.) roots by Fe and Cu status : Does the root-cell plasmalemma Fe (III)-chelate reductase perform a general role in regulating cation uptake? *Planta*, *190*, 555–561.

Wintz, H., Fox, T., Wu, Y.-Y., Feng, V., Chen, W., Chang, H.-S., Zhu, T., et al. (2003). Expression profiles of Arabidopsis thaliana in mineral deficiencies reveal novel transporters involved in metal homeostasis. *The Journal of biological chemistry*, *278*(48), 47644–53. doi:10.1074/jbc.M309338200

Yamasaki, H., Abdel-Ghany, S. E., Cohu, C. M., Kobayashi, Y., Shikanai, T., & Pilon, M. (2007). Regulation of copper homeostasis by micro-RNA in Arabidopsis. *The Journal of biological chemistry*, *282*(22), 16369–78. doi:10.1074/jbc.M700138200

Yamasaki, H., Hayashi, M., Fukazawa, M., Kobayashi, Y., & Shikanai, T. (2009). SQUAMOSA Promoter Binding Protein-Like7 Is a Central Regulator for Copper Homeostasis in Arabidopsis. *The Plant cell*, *21*(1), 347–61. doi:10.1105/tpc.108.060137

Yoshida, H. (1883). Chemistry af Lacquer. *Journal of the Chemical Society*, *43*, 472–486.

Yruela, I., Gatzen, G., Picorel, R., & Holzwarth, a R. (1996). Cu(II)-inhibitory effect on photosystem II from higher plants. A picosecond time-resolved fluorescence study. *Biochemistry*, *35*(29), 9469–74. doi:10.1021/bi951667e