TRANSPORT OF DITERPENES BETWEEN THE PLASTID AND ENDOPLASMIC RETICULUM

By

Cassandra Johnny

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

CELL AND MOLECULAR BIOLOGY –Doctor of Philosophy

2018

ABSTRACT

TRANSPORT OF DITERPENES BETWEEN THE PLASTID AND ENDOPLASMIC RETICULUM

By

Cassandra Johnny

Plant cells are highly compartmentalized with membranes that delineate organelles and provide physical barriers against movement of proteins and compounds. Many metabolic pathways span subcellular compartments, which necessitates transport of pathway intermediates between compartments. Diterpenes number more than 12,000 structures and whose syntheses are prime examples of membrane-spanning plant metabolic pathways. Plastid-localized diterpene synthases cyclize the universal precursor geranylgeranyl diphosphate (GGDP) into various diterpene olefins that are then trafficked into the endoplasmic reticulum (ER) for further oxidation reactions by ER-resident cytochrome P450s. Fifty-three plastid envelope transporters have been identified and characterized and all but two move polar metabolites and most notably, none have been identified for non-polar diterpene olefins. A previous study showed that plastid-localized tocopherol and carotenoid precursors were accessible in a bidirectional fashion from the ER lumen, suggesting that a novel interface exists between plastids and the ER that could allow access of non-polar compounds between the two organelles independent of transporters.

In this dissertation, the accessibility of diterpene olefins between the plastid and ER was studied using two diterpene pathways, the gibberellin (GA) and diterpene resin acid (DRA) pathways. In Chapter 2, the accessibility of *ent*-kaurene synthesized in the ER lumen by outer envelope membrane-localized kaurene oxidase is tested by retargeting both plastid-localized copalyl diphosphate synthase (CPS) and kaurene synthase (KS) into the ER in Arabidopsis mutant backgrounds lacking both activities. Access to ER-synthesized *ent*-kaurene was

evidenced by complementation of the dwarf mutant phenotypes. Surprisingly, the polar intermediate, *ent*-copalyl diphosphate (*ent*-CDP) was also shown to be accessible in a unidirectional fashion from the ER back into the plastid. This was observed by complementation of the dwarf phenotype of the CPS mutant, *ga1-6*, by ER-targeted CPS. Localization studies of ER-targeted CPS-YFP protein in transgenic lines showed that ER:CPS-YFP was correctly targeted to the ER lumen. Metabolite analysis of various GA intermediates showed levels in the ER:CPS-YFP lines similar to WT.

In Chapter 3, the accessibility of conifer-specific diterpene olefins, abietadiene and isopimaradiene by ER-localized CYP720B4 was investigated. These diterpene olefins were produced in stable Arabidopsis transgenic lines by overexpressing plastid-localized abietadiene and isopimaradine synthases. Expression of these conifer-specific bifunctional diterpene synthases however, only resulted in extremely low levels of the corresponding diterpene olefins. The low levels of the bifunctional diterpene synthase proteins produced, coupled to the apparent inability of these proteins to form productive complexes with the existing GGDP-producing machinery in Arabidopsis are the most likely explanations for the low levels of diterpene olefins produced.

As a whole, the work in this dissertation provided more insight into the transport of *ent*-CDP and *ent*-kaurene in the GA pathway and that the transport processes for the two compounds are most likely mediated by two separate mechanisms.

ACKNOWLEDGEMENTS

As the excitement of summer gives way to a new school year, my heart flutters with joy, and at times, disbelief at the completion of this dissertation. I have spent the past six years learning and growing in my scientific career, and am extremely grateful for having the opportunity to work in the DellaPenna laboratory.

First, I would like to thank my supervisor, Dr. Dean DellaPenna. Without his support and advice, the pursuit of this advanced degree would not have been possible. Being a graduate student in his laboratory enabled me to learn not only various lab skills but also how to give good presentations and answer questions effectively. I am also extremely grateful to the rest of my committee members, Dr. Dan Jones, Dr. Christoph Benning and Dr. Jianping Hu for their time, support and expertise. Thank you for the inspiration, thought-provoking questions, and helpful suggestions that have led to the completion of this dissertation.

I would like to thank Dr. Payam Mehrshahi for his mentorship and showing me "the ropes" in the early part of my graduate career. I also had the wonderful opportunity to work with Violet Whitney, who assisted in various experiments that were instrumental to this work. To the past and present members of the DellaPenna lab, thank you for all the helpful scientific and career discussions.

Nobody has been more important to me during the pursuit of this dissertation than my family. It was ingrained in me from a young age that only education can bring someone to a better future, and my family has never ceased to encourage and motivate me to achieve this doctoral degree. Thank you for your love and support. To the many friends that I have made here in East Lansing, thank you for your friendship and support.

Lastly, this work would not have been possible without funding from the National Science Foundation and Michigan State University.

TABLE OF CONTENTS

LIST OF TABLES

LISTS OF FIGURES

KEY TO ABBREVIATIONS

GA	Gibberellin
DRA	Diterpene resin acid
IDP	Isopentyl diphosphate
DMADP	Dimethylallyl diphosphate
MEP	2-*C*-methyl-D-erythritol 4-phosphate
MVA	Mevalonate
ER	Endoplasmic reticulum
GDP	Geranyl diphosphate
FDP	Farnesyl diphosphate
GGDP	Geranylgeranyl diphosphate
IEM	Inner envelope membrane
OEM	Outer envelope membrane
CPS	Copalyl diphosphate synthase
KS	Kaurene synthase
KO	Kaurene oxidase
Ent-CDP	*Ent*-copalyl diphosphate
FA	Fatty acid
GGDPS	Geranylgeranyl diphosphate synthase
DiTPS	Diterpene synthase

CHAPTER 1: Literature review

The role of plastids in plant metabolism

Plastids are the defining organelles of photosynthetic eukaryotes. They are the site of synthesis for many metabolites such as chlorophyll, carotenoids and gibberellins (GAs) that are necessary to support plant growth and development and in addition, plastids synthesize specialized metabolites such as diterpene resin acids (DRAs), that are important for plant protection against biotic and abiotic stresses. In addition to housing the entirety of some of these metabolic pathways (e.g., carotenoids, tocopherols and fatty acids), plastids also house portions of pathways that span multiple subcellular compartments, such as those for monoterpenes, diterpenes, and membrane lipids. This spatial separation of pathways requires coordinated exchange of pathway intermediates across the double membrane plastid envelope. Some of these compounds are polar or charged and therefore cannot diffuse through the hydrophobic core of membrane bilayers (Flügge, 2000). The plastid envelope contains 53 characterized transporter proteins that mediate the transport of such polar metabolites (Appendix A). In contrast, only two transporters have been identified for transports of fatty acids and membrane lipids (Appendix A).

Terpenoid biosynthesis

Terpenoids encompass the most structurally and functionally diverse class of plant metabolites (Tholl and Lee, 2011). All terpenoids are derived from two five-carbon precursors, isopentenyl diphosphate (IDP) and its isomer, dimethylallyl diphosphate (DMADP). In plants, these precursors are both produced by two parallel and compartmentally separated biosynthetic pathways, the plastid-localized 2-*C*-methyl-D-erythritol 4-phosphate (MEP) and the cytosol/peroxisome-localized mevalonate (MVA) pathways (Tholl and Lee, 2011). The formation of prenyl diphosphate intermediates of increasing chain lengths is catalyzed by short-

chain prenyltransferases, producing C_{10}-, C_{15}-, C_{20}- and C_{25}-prenyl diphosphates, which are substrates for a large family of terpene synthases (Tholl and Lee, 2011; Nagel *et al.*, 2015; C., Wang *et al.*, 2016). Generally, mono- (C_{10}), di- (C_{20}) and sesterterpenes (C_{25}) are synthesized in plastids while sesqui- (C_{15}) and triterpenes (C_{30}) are synthesized in the cytosol. In Arabidopsis, plastid-localized monoterpene and diterpene synthases produce an array of structurally diverse C_{10}- and C_{20}- hydrocarbon skeletal structures. These hydrophobic pathway intermediates are typically further decorated by oxygenation reactions of endoplasmic reticulum (ER)-localized cytochrome P450s (Tholl and Lee, 2011; Zi *et al.*, 2014).

Interaction between the MEP and MVA pathways

Despite both subcellular compartments containing pools of IDP and DMADP precursors, there is considerable evidence supporting a low level of metabolic interaction between the MEP and MVA pathway (Flügge and Gao, 2005; Bick and Lange, 2003). However, such low-level exchanges between the two pathways cannot fully compensate the loss of precursors necessary for terpenoid production within a single compartment in a MEP or MVA pathway mutant (Nagata *et al.*, 2002; Kasahara *et al.*, 2002; Tholl and Lee, 2011). IDP is widely accepted to be the prenyl diphosphate intermediate that is exchanged between these two compartments (Bick and Lange, 2003; Flügge and Gao, 2005), but the true mode and identity of any IDP transporter remains unknown. Bick and Lange (2003) proposed that transport of IDP is likely mediated by a Ca^{2+}-gated IDP/proton symporter mechanism while Flügge and Gao (2005) proposed that IDP is transported by a uniport system, with the requirement that phosphorylated compounds are present on the opposite side of the membrane. In addition to IDP, geranyl diphosphate (GDP), farnesyl diphosphate (FDP) and geranylgeranyl diphosphate (GGDP) have also been shown to

have low level exchanges between the two compartments (Bick and Lange, 2003; Gutensohn *et al.*, 2013).

These exchanges were shown by **(1)** expressing prenyl diphosphate synthases and terpene synthases into two different subcellular compartments and, **(2)** utilizing MEP- or MVA-specific inhibitors or mutants and feeding radiolabeled precursors and tracking incorporation of labeled precursors into final products. A comprehensive review covering both the biochemical and genetic tools used to assess the contribution of MEP or MVA-specific precursors into final terpenoid products has been published, and interested readers are referred to the following publication (Lipko and Swiezewska, 2016). Gutensohn *et al.* (2013) demonstrated that the levels of novel monoterpene production in the tomato fruit cytosol increased by 3.5-fold when plastid-localized snapdragon geranyl diphosphate synthase (GDPS) and cytosol-localized α-zingiberene synthase (which possesses both mono- and sesquiterpene synthase activity) were coexpressed. This is in comparison to the levels of novel monoterpenes when only α-zingiberene synthase was expressed in the cytosol. The authors confirmed that the activity of α-zingiberene synthase was mostly enriched in the cytosol, and reasoned that the production of novel monoterpenes was due to channeling of plastidic GDP to the cytosol (Gutensohn *et al.*, 2013). Dong *et al.* (2015) also demonstrated the trafficking of GDP across different subcellular compartment by coexpressing different combinations of GDP synthase and geraniol synthase in the plastid, cytosol, and mitochondria. By quantifying the resulting production of geraniol or geraniol-derived products, Dong *et al.* (2015) showed that there was a difference in the levels of GDP exchanged between these different subcellular compartments, the highest being 100% of the GDP produced in the mitochondria being accessible to plastid-localized geraniol synthase. Kasahara *et al.* (2002) used both a null mutant of the MEP pathway and a chemical inhibitor to block the MVA pathway to

study the flux of MEP- or MVA-derived precursor into the final terpenoid products in Arabidopsis. They treated the seedlings with either radiolabeled 1-deoxy-D-xylulose (a MEP precursor) or mevalonolactone (a MVA precursor) to track which precursor can contribute to the formation of *ent*-kaurene (synthesized in the plastids) and campesterol (synthesized in the cytosol). By assessing the incorporation percentage of labeled precursors into metabolites of the GA and sterol pathways, Kasahara *et al*. (2002) determined that 99% and 53% of the precursors derived from the MEP and MVA pathways, respectively, contributed to GA formation. On the other hand, 27% and 98% of precursors derived from the MEP and MVA pathway were incorporated to sterol metabolites. The difference in these percentages of incorporation showed the MEP and MVA pathways contribute precursors at different levels. May *et al.* (2013) also investigated the contribution of the MEP or MVA pathways in the formation of sesquiterpenes in wine grapes. Similar to Kasahara *et al.* (2002), they also found that both labeled-MEP and MVA precursors can be incorporated into various sesquiterpenes synthesized in the cytosol, although MEP-derived IDP was incorporated at less than 10% while MVA-derived IDP was incorporated at a range between 16-50%. Wölver-Rieck *et al.* (2014) also investigated the incorporation of labeled MEP and MVA-derived precursors into mono-, sesqui- and diterpenes of *Stevia rebaudiana* plant. They concluded that mono- and diterpene biosynthesis largely rely on MEP-derived precursors, but sesquiterpene biosynthesis is supported by both MEP and MVA pathway as 50% incorporation rates were seen for both pathways. The common underlying observation from all these studies is that the formation of terpenoids in plants is not exclusively either through the precursors derived from the MEP or MVA pathway. The formation of these "mixed-origin" terpenoids may be enhanced during scenarios such as genetic mutation in one pathway or at different developmental stages of the plant.

Diterpene pathways: General and specialized metabolism

Diterpenes are a subclass of terpenoids that are made from the 20-carbon intermediate GGDP. An estimated 12,000 diterpene structures have been identified so far in plants (Zi *et al.*, 2014). This structural diversity of diterpenes is generated by the action of two classes of enzymes: diterpene synthases in the plastid that first catalyze the formation of non-polar diterpene olefins which are then trafficked to the ER for further oxidation reactions by cytochrome P450s (**Figure 1.1**). In general metabolism, GAs are the only diterpenes that are required for plant growth and development. The majority of the other structurally diverse diterpenes are involved in specialized metabolism, and are important for plant protection against pests and pathogen in an ecological niche (Zerbe and Bohlmann, 2015; Schmelz *et al.*, 2014).

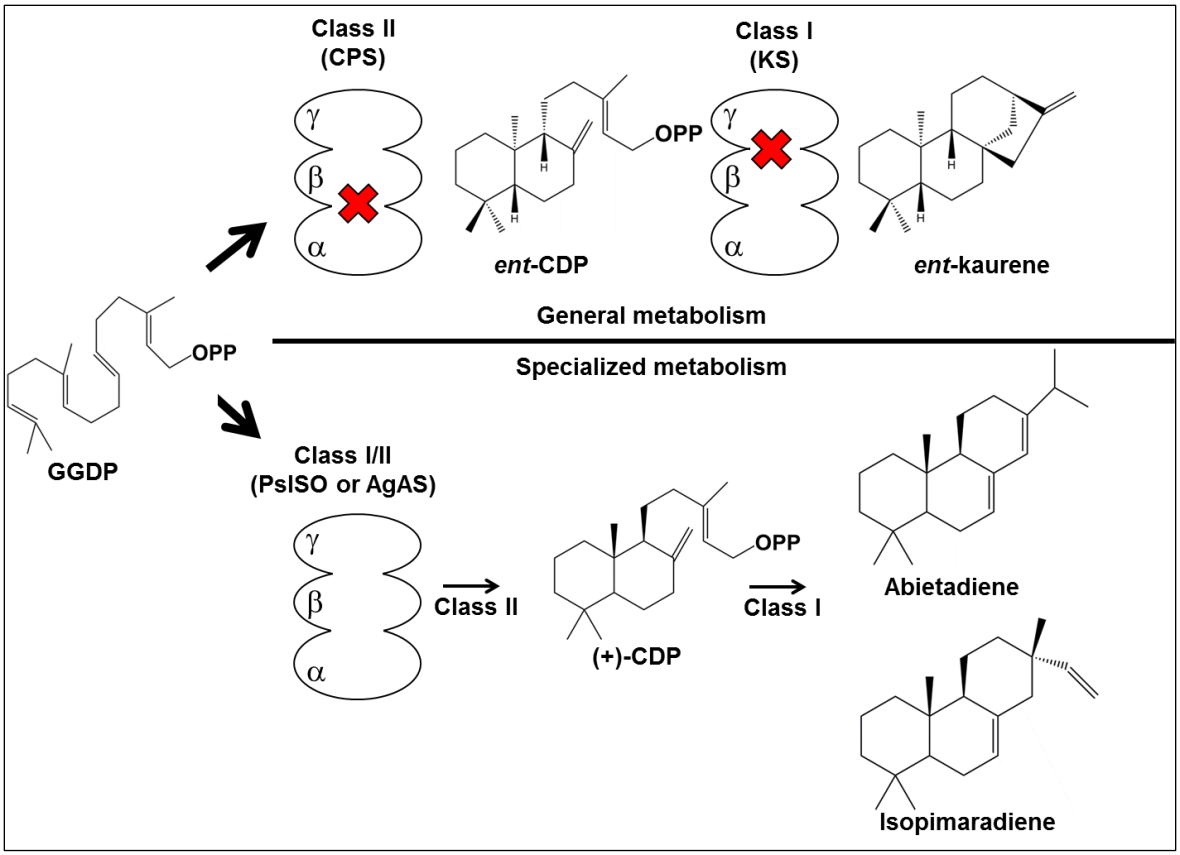


Figure 1.1: Diterpene synthases and their enzymatic reactions. Plastid-localized diterpene synthases have a three-domain protein structure: the α, β and γ domains. CPS, a class II

Figure 1.1: (cont'd) diterpene synthase converts geranylgeranyl diphosphate (GGDP) to *ent*-copalyl diphosphate (*ent*-CDP). KS, a class I diterpene synthase, then catalyzes the formation of *ent*-kaurene from *ent*-CDP. **X** represents the non-functional class I active site in CPS and class II active site in KS.In specialized metabolism, abietadiene synthase (AgAS) from grand fir and isopimaradiene synthase (PsISO) from Sitka spruce are involved in the synthesis of diterpene resin acids (DRAs). Both AgAS and PsISO are examples of bifunctional diterpene synthases that contain both class I and II activities on a single protein. The major diterpenes synthesized by AgAS and PsISO are abietadiene and isopimaradiene, respectively.

The gibberellin (GA) pathway:

In GA biosynthesis, formation of the non-polar compound *ent*-kaurene is catalyzed by two separate, plastid-localized diterpene synthases: copalyl diphosphate synthase (CPS) and kaurene synthase (KS). The *ent*-kaurene produced by the combined action of these two enzymes is then trafficked to the outer envelope membrane (OEM) where the cytochrome P450, kaurene oxidase (KO) is localized and catalyzes formation of *ent*-kaurenoic acid. *Ent*-kaurenoic acid is then transported into the ER where additional oxidation reactions occur. Bioactive GAs, GA_1, and GA_4 are then synthesized in the cytosol (**Figure 1.2**). Intensive studies of this pathway have fully elucidated the genes involved in biosynthesis, signaling, perception and turnover pathways of GAs (Sun, 2008; Hedden, 2016; Hedden and Thomas, 2012). Intracellular transporters mediating movement of *ent*-kaurene from the inner to OEM or *ent*-kaurenoic acid from the plastid into the ER have yet to be discovered.

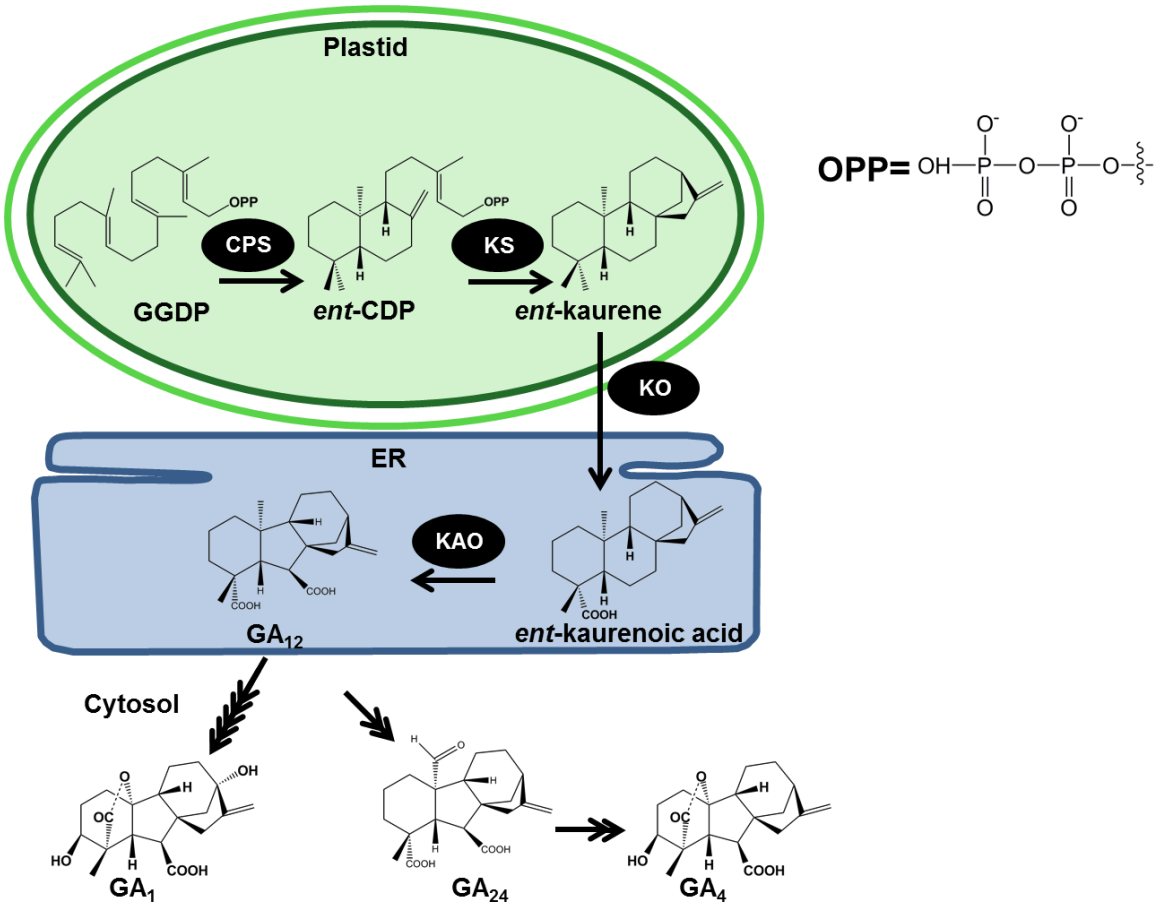


Figure 1.2: Gibberellin (GA) biosynthesis in Arabidopsis. Plastid-localized diterpene synthases, copalyl diphosphate synthase (CPS) and kaurene synthase (KS) catalyze the cyclization reaction of geranylgeranyl diphosphate (GGDP) to *ent*-copalyl diphosphate (*ent*-CDP) and then to *ent*-kaurene, respectively. *Ent*-kaurene (a non-polar compound) is then converted to *ent*-kaurenoic acid by a cytochrome P450 on the outer envelope membrane, kaurene oxidase (KO). *ent*-kaurenoic acid is converted to GA_{12} by an ER-localized cytochrome P450. GA_{12} lies at the branch point in the pathway and either undergoes hydroxylation at C-13 position or does not. Two families of soluble 2-oxoglutarate dependent dioxygenases catalyze the formation of GA_{24} (an intermediate of the non-13-hydroxyated GA branch) which eventually form the bioactive GA, GA_4. GA_1 is another bioactive GA from the 13-hydroxylated branch of the GA pathway.

The diterpene resin acid (DRA) pathway:

Diterpene resin acids (DRAs) and a mixture of mono- and sesquiterpenes are major compounds produced in specialized resin ducts in conifers (order Coniferales) for protection against pests and pathogens (Keeling and Bohlmann, 2006; Langenheim, 2003). In the synthesis of DRAs, a single, plastid-localized bifunctional diterpene synthase catalyzes the formation of various

diterpene olefins with tricyclic parent skeletons from GGDP (Keeling and Bohlmann, 2006) (Figure 1.1). These diterpene olefins are then translocated into the ER and converted into the corresponding DRAs by the action of cytochrome P450 monooxygenases. Members of the conifer-specific cytochrome P450 CYP720B family catalyze the three-step oxidation reactions that convert the various diterpene olefins into their corresponding DRAs. Members of the CYP720B family are clustered into four distinct clades (I to IV) (Hamberger *et al.,* 2011; Geisler *et al.*, 2016). Members of clade I catalyze the formation of DRAs from an initial unstable diterpene alcohol while members of clade III catalyze the formation of DRAs from diterpene olefins (Geisler *et al.,* 2016; Hamberger *et al.*, 2011). For example, the clade III enzymes, PtCYP720B1 from lolloby pine (*Pinus taeda*) and PsCYP720B4 from Sitka spruce (*Picea sitchensis*), have been characterized and shown to catalyze the three-step oxidation reactions of various diterpene olefins into their corresponding DRAs (Ro *et al.*, 2005; Hamberger *et al.*, 2011). The diterpenes in the DRA pathway are composed of members of three structural groups: abietanes, pimaranes and dehydroabietanes (Hamberger *et al.*, 2011; Keeling and Bohlmann, 2006; Keeling *et al.*, 2011).

Plastid envelope transporters

As described earlier, the transport of polar compounds across the OEM is largely mediated by outer envelope proteins (OEPs), which are β-barrel proteins that form anion- or cation-selective pores with broad substrate specificity (Flügge, 2000; Breuers *et al.*, 2011; Pottosin and Dobrovinskaya, 2015; Fischer *et al.*, 2016). On the other hand, transport across the inner envelope membrane (IEM) is largely mediated by α-helical proteins with high substrate specificities (Fischer *et al.*, 2016; Linka and Weber, 2010). Bioinformatic and proteomic studies

have identified 97 putative transporters of ion, metals, and metabolites localized to the envelope membranes (Froehlich *et al.*, 2003; Zybailov *et al.*, 2008; Baginsky and Gruissem, 2009; Ferro *et al.*, 2010, Appendix A). Of these proteins, 30 are categorized as putative transporters, 53 have been functionally characterized, while 16 are annotated as highly homologous to the 53 characterized transporters and likely have functions similar to their characterized paralogs (Mehrshahi *et al.*, 2013). Fifty-one of the 53 functionally characterized transporters transport polar compounds, or metals and ions across the plastid envelope membranes while only two are transporters for fatty acids/ lipids (Mehrshahi *et al.*, 2013; Fischer *et al.*, 2016). In contrast, the transport process for other non-polar compounds such as mono- or diterpene olefins from plastids is still unknown.

Fatty acid/lipid transporters

In plants, plastids are the site of *de novo* fatty acid synthesis with most research in this field being focused on lipid synthesis in photosynthetic plastids, chloroplasts (Hurlock *et al.*, 2014). Plastid-synthesized fatty acids can either be incorporated into glycerolipids within the chloroplasts (the so-called "prokaryotic pathway) or exported to the ER (the so-called eukaryotic pathway). Due to the specific preference of acyltransferases in the two organelles for certain fatty acids, the origin of glycerolipids can be determined from their fatty acid content. Lipids derived from the prokaryotic pathway have 16-carbon fatty acids esterified to the *sn-2* position, while lipids derived from the eukaryotic pathway invariably have 18-carbon fatty acids esterified to the *sn-2* position. Several transporter proteins have been identified in lipid trafficking between the chloroplasts and ER in Arabidopsis (Zhou *et al.*, 2016). The characterization of fatty acid export 1 (FAX1), a novel protein localized on the inner envelope membrane suggests that this

protein is involved in the export of fatty acids to the ER (Li *et al.*, 2015). The ability of the FAX1 protein to transport fatty acids was shown in a yeast mutant lacking a membrane-spanning fatty acid transport protein. In addition to FAX1, three additional members of this protein family (FAX2-FAX4) are also predicted to be localized in the chloroplasts and are thought to be redundant in fatty acid export. For further metabolism in the ER, fatty acids need to be converted to their corresponding acyl-CoAs. One member of the long-chain CoA synthetase (LACS) protein that is localized in chloroplast envelope, LACS9 is presumed to catalyze this reaction (Schnurr *et al.*, 2002). However, pulse-chase labeling experiments showed that LACS9 protein, in concert with the ER-localized LACS4 protein, participate in ER-to-chloroplast lipid trafficking instead (Jessen *et al.*, 2015). Uptake of fatty acids into the ER can also be mediated by ATP-binding cassette (ABC) transporter protein ABCA9 localized in the ER (Kim *et al.*, 2013). An alternative transport mechanism of fatty acids export from the chloroplast is transport by acyl-CoA binding proteins (ACBPs). ACBP4 and ACBP5 are localized in the cytosol and have been shown to have a high binding affinity towards C18-CoAs (Xiao *et al.*, 2008). Another transporter complex, the trigalactosyldiacylglycerol (TGD) transporter complex has also been shown to mediate transport of lipid from the ER back into the chloroplasts (Xu *et al.*, 2003; Awai *et al.*, 2006; Lu *et al.*, 2007; Xu *et al.*, 2008). TGD1, TGD2, and TGD3 form a multi-subunit ABC transporter in the inner envelope membrane, while TGD4 forms a homodimer in the outer envelope membrane (Roston *et al.*, 2012; Wang *et al.*, 2013). Another TGD protein, TGD5, has been shown to interact with the TGD1, 2, and 3 ABC transporter complex proteins and TGD4, linking lipid transport from the outer envelope membrane to the inner envelope membrane (Fan *et al.*, 2015). Although both TGD2 and 4 have been shown to bind phosphatidic acid (PA), there

is no direct evidence demonstrating that this is the transported lipid species (Awai *et al.*, 2006; Wang *et al.*, 2013).

Characterized terpenoid transporters at the plasma membrane

Although how terpenoid pathway intermediates are exchanged within the cell is largely unknown, several plasma membrane-localized terpenoid transporters have been identified and characterized. These transporters or carrier proteins include members of the nitrate transporter /peptide transporter family (NPF), SWEET transporters, ATP-binding cassette (ABC) subfamily G proteins and lipid transfer proteins (LTPs). In general, these transporters have been shown to either export or import highly oxidized sesquiterpenes or diterpenes, in contrast to the diterpene hydrocarbons (*ent*-kaurene, abietadiene or isopimaradiene) that are exchanged between the plastids and ER. Furthermore, these plasma membrane transporters are able to transport multiple classes of compounds.

The NPF transporters

The NPF plasma membrane-localized transporters transport not only GA intermediates that are more oxidized, but also a range of other compounds including nitrate, di/tripeptides, glucosinolates and plant hormones (Kanno *et al.*, 2012; Nour-Eldin *et al.*, 2012; Saito *et al.*, 2015; Tal *et al.*, 2016). The NPF transporter family is composed of 53 members, which share sequence homology with transporter proteins that are ubiquitous in bacteria, animals, and fungi (Léran *et al.*, 2014). Three members of this NPF family (AtNPF2.10, AtNPF3.1, and AtNPF4.1) are plasma membrane-localized and have been shown to transport GA_3, either *in vitro* using yeast or *Xenopus* oocytes or *in vivo* using Arabidopsis seedlings (Kanno *et al.*, 2012; Saito *et al.*,

2015; Tal *et al.*, 2016). Interestingly, mutants of these transporters rarely have dramatic phenotypes such as the dwarf stature or seed germination problems that are hallmarks of mutants in the GA biosynthetic pathway, consistent with functional redundancy in transport of GAs at the plasma membrane. The loss-of-function mutant for NPF2.10 (previously named glucosinolate transporter 1), the *gtr1* mutant, was identified during a screening for novel regulators for jasmonic acid (JA) signaling (Saito *et al.*, 2015). The *gtr1* mutant has small siliques and reduced fertility, and these phenotypes are attributed to its shorter stamen filaments and failure in anther dehiscence (Saito *et al.*, 2015). Two alleles of the *npf3* mutant were characterized in a screening for putative GA transporter using fluorescently-labeled GAs (Tal *et al.*, 2016). These mutants are similar to WT in terms of seed germination, root and shoot elongation and flowering time (Tal *et al.*, 2016). A mutant for NPF4.1 (previously named ABA-importing transporter 1), the *ait3* mutant, was identified during a screening for putative abscisic acid (ABA) transporters (Kanno *et al.*, 2012). Since Kanno *et al.* (2012) were interested in transport of ABA, characterization was based on phenotypes typically observed for ABA deficiencies. ABA deficiencies can be observed by monitoring stomatal aperture as ABA is transported into the guard cells (Kanno *et al.*, 2012). This is done by measuring rosette leaves surface temperature to indirectly monitor stomatal aperture, in which the *ait3* mutant showed no temperature difference compared to WT (Kanno *et al.*, 2012). No further characterization has been reported for the *ait3* mutant.

Studies of the NPF2.10, NPF3.1 and NPF4.1 demonstrated transport of GA_3 at the plasma membrane, however, GA_3 is not the only substrate for these transporters (Kanno *et al.*, 2012; Saito *et al.*, 2015; Tal *et al.*, 2016). AtNPF2.10 was shown to be able to transport JA-Ile (bioactive form of JA), GA_3 and 4-methylthiobutyl glucosinolate (4MTB; a glucosinolate with high affinity for NPF2.10 (Nour-Eldin *et al.*, 2012)) in *Xenopus* oocytes (Saito *et al.*, 2015). In

addition, transport activities of JA-Ile and GA_3 were diminished in the presence of 4MTB, indicating preferential transport of 4MTB by NPF2.10. Characterization of NPF2.10 in *Xenopus* oocyte showed an apparent K_m of 20 μM towards 4MTB (Nour-Eldin *et al.*, 2012) and 301 μM towards GA_3 (Saito *et al.*, 2015).

The NPF3.1 proteins were also shown to transport various GAs (GA_1, GA_3, GA_4 (bioactive GAs), GA_{20} (a GA intermediate) and GA_8 (a GA catabolite)) (Tal *et al.*, 2016). Unlike the NPF2.1 transporter protein, the specificity of GA_4 transport by NPF3.1 was not diminished in the presence of other compounds that are typically transported by members of NPF family proteins such as NO_3^-, dipeptide or amino acid. Despite demonstrating the ability of NPF3.1 to transport GA_4, Tal *et al.* (2016) estimated the apparent affinity for GA_4 to be 500 μM. They also showed decreased accumulation of fluorescently-labeled GAs by the *npr3* Arabidopsis mutants compared to WT. Using the *Xenopus* oocyte system, Tal *et al.* (2016) also demonstrated that the NPF4.1 transporter protein was able to transport various GAs (GA_3, GA_4, GA_8 and GA_{20}). In addition, NPF4.1 was also shown to transport both ABA and GA_3 in yeast cells expressing the NPF4.1 protein (Kanno *et al.*, 2012). These three transporters, NPF2.1, 3.1, and 4.1, transport various GA molecules, which suggests that these transport activities may be shared redundantly between two or more proteins. Their different tissue expression patterns (roots versus flower) may also explain the biochemically redundant nature of these transporters characterized for GA transport.

The SWEET transporters

Another transporter family at the plasma membrane that has been linked to GA transport is the SWEET protein family. SWEET proteins were originally identified as transporters of glucose,

fructose and sucrose (Chen *et al.*, 2010). Using a yeast-two-hybrid system expressing the GA perception complex GID/DELLA, Kanno *et al.* (2016) identified both AtSWEET13 and 14 as potential GA transporters. As with NPF loss of function mutants, AtSWEET13 and 14 loss-of-function single mutants appear phenotypically similar to WT (Kanno *et al.*, 2016). Since both genes are expressed in the anthers and vascular tissues in the leaves and roots of young seedlings, the authors hypothesized that they may be functionally redundant, and generated the double mutant *sweet13 sweet14*. The double mutant showed reduced fertility and smaller seed set compared to WT due to a delay in anther dehiscence in the double mutant. This phenotype could be reverted with application of GA_3. Interestingly, quantification of various GAs in various tissues of the double mutant largely showed no significant difference compared to WT, except for lower levels of the deactivated GAs, GA_8 and GA_{29} 10 days after seed fertilization.

The GA transport function of SWEET transporters was further tested in yeast cells and *Xenopus* oocytes. In yeast cells expressing either AtSWEET13 or 14, significant GA_3 transport was shown compared to the control while in the *Xenopus* oocyte system, accumulation of 11 different GAs was higher than the control (Kanno *et al.*, 2016). The authors estimated that both transporters have an apparent affinity for GA_3 in the range of several hundred micromolar (Kanno *et al.*, 2016). Kanno *et al.* (2016) also acknowledged that even though these SWEET proteins have been characterized as low affinity sugar transporters (*K*ms in the mM range), their affinity (*K*m) for GAs is still lower than that of the sugars. It is still debatable if these proteins do transport GAs *in vivo* as the levels of endogenous GAs are very low (often in < 20 ng g^{-1} fresh weight (Talon *et al.*, 1990; Fleet *et al.*, 2003)).

In conclusion, all GA transporters that have been characterized thus far in plants are localized to the plasma membrane. In addition, these GA transporters are most likely influx

transporters (Binenbaum *et al.*, 2018) responsible for transport of downstream GA metabolites that are more oxidized than *ent*-kaurene or *ent*-kaurenoic acid. In addition to potentially transporting GAs, these plasma membrane-localized transporters transport other metabolites such as glucosinolates, other phytohormones and sugars.

An ABCG family transporter involved in terpenoid transport

Sclareol is an oxidized bicyclic diterpene alcohol, produced by *Nicotiana glutinosa* (Guo and Wagner, 1995). Several transporters in the ATP-binding cassette (ABC) G subfamily have been implicated in the transport of sclareol, including *Nicotiana plumbaginafolia* pleiotropic drug resistance1 (PDR1) and *Nicotiana tabacum* PDR1(Stukkens *et al.*, 2001; Stukkens *et al.*, 2005; Crouzet *et al.*, 2013), both of which are localized to the plasma membrane. The expression of both NpPDR1 and NtPDR1 can be induced by elicitors such as methyl jasmonate or yeast extract, which suggests that these proteins may play a role in plant defense, presumably by exporting sclareol onto cell surfaces. By using a tobacco cell system with constitutive expression of transporter proteins, NpPDR1 transported sclareol and its close analog sclareolide (Stukkens *et al.*, 2001). Silencing of NpPDR1 also made cells susceptible to sclareol (Stukkens *et al.*, 2005). On the other hand, NtPDR1, which shares 84% amino acid identity to NpPDR1, transported several different diterpenes such as sclareol, manool, abietic acid, and dehydroabetic acid, which have different degrees of saturation, numbers of rings and degrees of oxidation (Crouzet *et al.*, 2013). It was also recently shown that the ATPase activity of NtPDR1 is stimulated by the diterpenes sclareol, cembrene and the sesquiterpene capsidiol (Pierman *et al.*, 2017). Another close ortholog of NtPDR1, a plasma membrane-localized ABCG1 protein in

Nicotiana benthamiana, has also been shown to be involved in the export of capsidiol during *Phytophthora infestans* infection (Shibata *et al.*, 2016).

Lipid transfer proteins

Lipid transfer proteins (LTPs) are 6.5 to 10.5 kDa proteins that are ubiquitous in eukaryotes (Liu *et al.*, 2015; Kader, 1996). The tertiary protein structure of LTPs is characterized by eight cysteine residues that form four disulfide bridges and stabilize a hydrophobic pocket (Liu *et al.*, 2015; Kader, 1996). Some LTPs are extracellular, consistent with the presence of a signal peptide at the N-terminus but no H/KDEL ER retention signal (Kader, 1996). Using a transient expression system in *Nicotiana benthamiana*, Wang *et al*. (2016) investigated the role of LTPs that are coexpressed with the artemisinin biosynthetic pathway, an oxygenated sesquiterpene produced in *Artemisia annua.* Coexpression of AaLTP3 with artemisinin biosynthetic pathway genes resulted in higher levels of the end products, artemisinin, and arteannuin B compared to empty vector controls (Wang *et al.*, 2016). AaLTP3 was specific in boosting the production of artemisinin and artennuin B as coexpression of AaLTP3 with costunolide (a sesquiterpene lactone) biosynthetic pathway genes did not result in increased levels of costunolide (Wang *et al.*, 2016). In addition, a different LTP in *Nicotiana tabacum*, NtLTP1, is involved in lipid secretion from long glandular trichomes (Choi *et al.*, 2012). These lipid secretions contained increased levels of diterpenes and alkane hydrocarbons in NtLTP1-overexpressing transgenic tobacco lines (Choi *et al.*, 2012).

Volatile organic compound transporters

In comparison to diterpenes, mono- and sesquiterpenes have lower molecular weights (100-200 Da) (Widhalm *et al.*, 2015), and unlike diterpenes are volatile and considered volatile organic compounds (VOCs), which includes oxidized fatty acid derivatives, benzoids, and some phenylpropanoids. Since VOCs are readily detected in plant headspace analysis, it was presumed that these compounds are transported out of the cells by passive diffusion owing to their hydrophobicity and high vapor pressure at ambient temperature (Widhalm *et al.*, 2015; Borghi *et al.*, 2017). However, this notion has recently been challenged based on mathematical models of VOC emission according to Fick's first law (Widhalm *et al.*, 2015). Passive diffusion of a molecule across a membrane bilayer can be represented by Fick's first law, which is dependent on parameters such as concentration difference along the diffusion path and resistance to diffusion through a given barrier (Widhalm *et al.*, 2015). Passive diffusion of VOCs would result in cytotoxic levels of VOC accumulation in membranes, and cannot fully explain the high emission rates reported for these compounds (Widhalm *et al.*, 2015). Several biological mechanisms have been proposed for emission of volatile terpenoids across cell membranes: **1)** direct transport of pathway intermediates across the plastid:ER contact or hemifused sites (Mehrshahi *et al.*, 2013; Mehrshahi *et al.*, 2014), **2)** partitioning of hydrophobic terpenoids into vesicles from the ER, followed by vesicle transport, fusion, and compound release at the plasma membrane (Skubatz *et al.*, 1995), **3)** binding of hydrophobic terpenoids by carrier or lipid transfer proteins (LTPs) across intracellular spaces and the plasma membrane (Widhalm *et al.*, 2015; Ting *et al.*, 2015) and **4)** transport of terpenoids across plasma membrane by ABC transporters (Widhalm *et al.*, 2015; Adebesin *et al.*, 2017). Most recently, an ABCG transporter in petunia flower (PhABCG1) was shown to be a plasma membrane-localized VOC transporter

that could transport phenylpropanoid and benzoic compounds, but not volatile terpenoids (Adebesin *et al.*, 2017).

The knowledge gap in terpenoid transport

The sheer diversity of compounds exchanged between the plastid and other subcellular compartments is astounding. The transport of polar compounds across plastids is well understood with 51 envelope transporters of polar compounds now characterized, while in contrast, only two transporters of non-polar compounds across the plastid envelope have been identified. All terpenoid transporters identified thus far are localized to and function at the plasma membrane, and the only non-polar transporters localized to plastid envelopes or the ER are involved in FA/lipid transport. While polar compounds would require the presence of transporter to cross a membrane bilayer, non-polar compounds are able to partition into the membrane bilayer.

Synthesis of some non-polar compounds would also require transport of intermediates into other subcellular compartments. For example, the synthesis of monoterpenes, diterpenes, and membrane lipids occurs in both the plastid and ER. While transporters have been identified for fatty acids/membrane lipids (FAX1 and the TGD transporter complex), no plastidic transporters for mono- and diterpenes have yet been identified (Li *et al.*, 2015; Hurlock *et al.*, 2014). Monoterpenes have a high vapor pressure at ambient temperature and are thus volatile, but diffusion-based movement of compounds cannot fully explain their rate of emission into the environment and non-polar pathway intermediates could potentially accumulate at a toxic level based on Fick's first law of diffusion (Widhalm *et al.*, 2015). This same diffusion limitation also applies to the transport of diterpenes.

The contact sites between plastids and the ER are hypothesized to play a role in the transport of non-polar compounds (Mehrshahi *et al.,* 2014). These contact sites are known as plastid-associated membranes (PLAMs) and have been visualized in transmission electron microscopy (Crotty and Ledbetter, 1973; McLean *et al.*, 1988; Turner and Lange, 2015) and freeze-fracture scanning electron microscopy (McLean *et al.*, 1988; Whatley *et al.*, 1991) in various plant species. In addition, in specialized plant structures that actively produce terpenoids, such as the peltate glandular trichomes of peppermint, resin ducts of pine trees or secretory cavities of grapefruit exocarp, these PLAM sites are also often observed in greater abundance (Turner *et al.*, 2000; Benayoun and Fahn, 1979; Turner and Lange, 2015). In a previous study in the DellaPenna laboratory (Mehrshahi *et al.*, 2013), an alternative model for movement of non-polar compounds was proposed to occur via membrane hemifusion sites at contact sites between the plastid and ER. Hemifusion is an intermediate step in membrane fusion where the inner leaflets of two membranes remain in contact and form a stable (hemifused) bilayer while the outer membrane leaflets display lipid mixing (Chernomordik and Kozlov, 2008). Due to the non-polar nature, mono- and diterpene olefins are able to partition into the membrane bilayers, and could presumably be accessible from the plastid or ER at hemifusion sites between the two organelles and thus "move" or be accessible between two organelles independent of transporters.

Transorganellar complementation as an in vivo functional approach to probe for substrate accessibility and transport.

The transorganellar complementation approach is a functional assay to probe the *in vivo* accessibility of non-polar substrates between organelles through complementation of mutated pathway activities in one organelle by enzyme targeted to a different (companion) organelle. In

general, an enzyme is retargeted to a different organelle by removing any targeting peptide and transformed into a mutant background lacking that activity in the original organelle. In the context of this dissertation, the accessibility of a pool of non-polar diterpene olefin between the plastid and ER is tested. In Chapter 2 of this dissertation, I describe transplanting the two steps synthesis of the diterpene olefin in the GA pathway, *ent*-kaurene, from the plastid to the ER and the subsequent accessibility of *ent*-kaurene by OEM-localized KO. In Chapter 3 of this dissertation, I describe experiments designed to transplant a non-endogenous plastid:ER spanning DRA pathway from conifers to Arabidopsis in order to assess accessibility of these compounds across the plastid:ER interface.

Aims of the study

This study aims to further define the operating principles of the interface between the plastid and ER by functionally determining the types of non-polar compounds that can be accessed across the plastid:ER interface *in vivo*, the directionality of this complementation and the classes of enzymes and pathways capable of transorganellar complementation. Two diterpene pathways, the GA and DRA pathways are chosen as model diterpene synthesis pathways that naturally span the plastid and ER. The transorganellar complementation approach is used to probe whether diterpene pathway intermediates can be accessible at the interface between the plastid and ER.

APPENDIX

Table 1: List of plastid envelope transporters as published in Mehshahi *et al.* (2013). This list has been updated on July 2018 to reflect most current functional identification of these proteins listed. In summary, there are 97 plastid envelope transporters and out of these 97, there are 24 known ion/metal transporters, 27 known polar metabolite transporters, 2 known non-polar transporters, 16 homologous to known metal/ion/polar metabolite transporters, 8 unknown ABC transporters, 2 unknown MATE transporters and 18 unknown transporters.

No	Protein Annotation	Proteomic Studies				Aramemnon	TMHMM	localization (PubMed ID)	Published Function	Predicted localization	Protein annotation	General Classification
		(1)	(2)	(3)	(4)							
1	AT2G29650		x	x	x	9	10	18086223; 14564522	Y	C	ANTR1 - Pi transporter; also named PHT4;1	Known metal/ion transporter
2	AT2G38060					12	10	18086223	Y	C	ANTR3 - Pi transporter; also named PHT4;2	Known metal/ion transporter
3	AT4G00370		x	x	x	10	10	18086223; 14564522	Y	C	ANTR1 - Pi transporter; also named PHT4;4	Known metal/ion transporter
4	AT1G01790	x	x		x	13	11		Y	-	K+ efflux antiporter, putative (KEA1)	Known metal/ion transporter
5	AT1G68570		x			11	10		Y	-	Nitrite transporter (CsNitrl1) - proton-dependent oligopeptide transport (POT) protein	Known metal/ion transporter
6	AT3G26570	x	x	x	x	12	13	12172020	Y	C	PHt2;1 Phosphate translocator - Na+/Pi transporter	Known metal/ion transporter
7	AT4G00630	x	x		x	12	10		Y	-	K+ efflux antiporter 2	Known metal/ion transporter
8	AT4G04770		x	x	x		0	11156608	Y	C	ATNAP1 or AtABC1, LAF6 - SufB-like (ABC/ATPase) - interacts with NAP7	Known metal/ion transporter
9	AT4G37270		x	x	x	6	5		Y	-	Cu-ATPase, HMA1, also involved Zn detoxification	Known metal/ion transporter
10	AT3G51895					13	12	23095126	Y	-	SULTR3;1, Sulfate transporter 3;1	Known metal/ion transporter
11	AT5G13550			x		12	12		Y	C	SULTR4;1, Sulfate transporter 4;1	Known metal/ion transporter
12	AT5G20380					10	7	18086223	Y	-	ANTR6 also named APHT4;5	Known metal/ion transporter

Table 1: (cont'd)

13	AT5G22830	x		x	x	2	2		Y	C	magnesium transporter protein (GMN10)	Known metal/ion transporter
14	AT5G42130		x		x	3	3		Y	C	MitoFerrinLike1 (MFL1) iron metabolism	Known metal/ion transporter
15	AT5G49740					10	10		Y	-	ferric-chelate reductase/ oxidoreductase 7 (FRO7)	Known metal/ion transporter
16	AT5G59040					3	2		Y	-	copper transporter 3	Known metal/ion transporter
17	AT4G33520		x	x	x	7	0	15772282	Y	C	ATPASE COPPER UPTAKE (HMA6)	Known metal/ion transporter
18	AT2G15290	x	x		x	4	3	19224954	Y	C	PERMEASE FE IMPORT (PIC1)	Known metal/ion transporter
19	AT5G14100		x		x	0	0		Y	C	ABC FE HOMEOSTASIS (NAP14)	Known metal/ion transporter
20	AT5G10490		x		x	2	5		Y	C	MECHANOSENSITIVE PORE (MSL2)	Known metal/ion transporter
21	AT1G58200		x		x	4	0	16401419	Y	C	MECHANOSENSITIVE PORE (MSL3)	Known metal/ion transporter
22	AT5G26820					12	11	19675150	Y	C	FE HOMEOSTASIS (MAR1)	Known ion/metal transporter
23	AT5G62720	x	x	x		5	4	24904028	Y	C	nitrite transporter; can transport nitrite transporter in cyanobacterial cell	known metal/ion transporter
24	AT4G13590			x	x	6	5	29734002	Y	C	Chloroplast manganese transporter1	known metal/ion transporter
25	AT1G65410			x	x		0		Y	C	TGD3- small ATPase involved in lipid import (ATNAP11)	Known non-polar metabolite transporter
26	AT3G57280	x		x	x	4	3	25646734	Y	C	Fatty acid export 1	Known non-polar metabolite transporter
27	AT5G01500			x	x	3	0	17261580	Y	C	ATP/ADP carrier (TAAC) - (likely envelope)	Known polar metabolite transporter

Table 1: (cont'd)

28	AT5G16150	x	x	x	x	12	10	14704427	Y	C	IEP62 (putative sugar transporter)	Known polar metabolite transporter
29	AT1G61800					8	7		Y	C	GPT2 (glucose-6-phosphate transporter 1) hexoseP/Pi	Known polar metabolite transporter
30	AT1G80300	x	x	x	x	11	9		Y	C	ATP/ADP translocator (AATP1 ort AtNTT1)	Known polar metabolite transporter
31	AT2G26900			x	x	9	9		Y	C	BASS2 - pyruvate importer	Known polar metabolite transporter
32	AT2G28900	x	x	x	x		0	17098851	Y	-	Amino acid transporter (OEP16:HP15)	Known polar metabolite transporter
33	AT2G32040			x		11	9	16162503	Y	C	folate transporter	Known polar metabolite transporter
34	AT3G01550					8	7		Y	C	phosphoenolpyruvate (pep)/phosphate translocator 2	Known polar metabolite transporter
35	AT3G08580	x		x	x	3	3		Y	M	ADP/ATP carrier 1 (AAC1)	Known polar metabolite transporter
36	AT4G24460				x	10	10		Y	C	CRT (chloroquine-resistance transporter)-like transporter 2	Known polar metabolite transporter
37	AT4G32400						0	21330298	Y	-	BT1- nucleotide uniport carrier - export AMP,ADP,ATP (Brittle-1 like) - dual mitos & plastid	Known polar metabolite transporter
38	AT4G39460		x	x	x	5	5	17261580	Y	C	S-Adenosylmethionine transporter,S-homocysteine (SAMT1)	Known polar metabolite transporter
39	AT5G03555					12	12	22474184	Y	C	permease, cytosine/purines, uracil, thiamine, allantoin family protein	Known polar metabolite transporter

Table 1: (cont'd)

40	AT5G12860	x	x	x	x	14	13		Y	C	2-oxoglutarate/malate translocator (DIT1; EP45; OMT1)	Known polar metabolite transporter
41	AT5G17520			x	x	9	9	14704427	Y	-	maltose exporter Mex1 (formerly root cap 1 -RCP1)	Known polar metabolite transporter
42	AT5G17630				x	8	7		Y	C	xylulose-5-phosphate/phosphate translocator (XPT)	Known polar metabolite transporter
43	AT5G33320		x	x	x	6	6		Y	C	PPT - IEP33 = Phosphate/phosphoenolpyruvate translocator	Known polar metabolite transporter
44	AT5G46110	x	x	x	x	8	9		Y	M	TPT - IEP30 = Phosphate/triose-phosphate translocator	Known polar metabolite transporter
45	AT5G54800					8	8		Y	C	GPT1 (glucose-6-phosphate transporter 1) hexoseP/Pi	Known polar metabolite transporter
46	AT5G64280			x		11	10		Y	C	2-oxoglutarate/malate translocator 2 (DiT2;IEP45 family; DCT2) - very low expression in leaf	Known polar metabolite transporter
47	AT5G64290	x	x	x		10	9		Y	C	2-oxoglutarate/malate translocator 1 (DiT1;IEP45 family;DCT1)	Known polar metabolite transporter
48	AT5G66380					2	2	16055441	Y	M	folate transporter 1 (FOLT1)	Known polar metabolite transporter
49	AT2G47490					3	4	19745225	Y	C	NAD/NAMN transporter (ATNDT1)	Known polar metabolite transporter
50	AT4G12030					9	7	19542295	Y	C	bile acid transporter 5 (BAT 5) - transports short- and long-chain 2-keto acids	Known polar metabolite transporter
51	AT4G22840					8	8	28351992	Y	C	Glycolate transporter	Known polar metabolite transporter

Table 1: (cont'd)

52	AT1G32080	x	x	x	x	12	12		Y	C	Plastidial glycolate glycerate transporter 1	Known polar metabolite transporter
53	AT1G78560				x	9	9	29382740	Y	C	Plastidial pantoate transporter	Known polar metabolite transporter
54	AT2G16800					6	5		N	C	high-affinity nickel-transport family protein	Homologous to known ion/metal transporter
55	AT3G21580				x	6	5		N	C	cobalt ion transmembrane transporters	Homologous to known ion/metal transporter
56	AT4G33460	x		x	x		0		N	C	ABC transporter family protein	Homologous to known ion/metal transporter
57	AT4G35080					5	5		N	C	high-affinity nickel-transport family protein	Homologous to known ion/metal transporter
58	AT2G21340			x	x	9	8	26055508	N	C	MATE efflux family protein - Proton dependent (putative phenolic acid transporter)	Homologous to known polar metabolite transporter
59	AT1G15500		x	x	x	11	11		N	-	TLC ATP/ADP transporter (AtNTT2)	Homologous to known polar metabolite transporter
60	AT3G51870			x	x	2	2		N	M	Mitochondrial substrate carrier family protein	Homologous to known polar metabolite transporter
61	AT2G40420					11	11		N	-	Transmembrane amino acid transporter family protein	Homologous to known polar metabolite transporter

Table 1: (cont'd)

62	AT3G10290			x		10	10		N	C	Nucleotide-sugar transporter family protein	Homologous to known polar metabolite transporter
63	AT3G46980				x	12	12		N	C	phosphate transporter, ANTR4 also named APHT4;3	Homologous to known polar metabolite transporter
64	AT1G04570					10	8		N	C	putative folate/biopterin transporter homologs	Homologous to known polar metabolite transporter
65	AT2G35800				x	2	0	22062157	N	C	SAM transporter-like (SAMTL) (likely envelope)	Homologous to known polar metabolite transporter
66	AT3G62880					2	0		N	-	OEP16-4	Homologous to known polar metabolite transporter
67	AT4G16160			x			0	17098851	Y	-	OEP16-2	Homologous to known polar metabolite transporter
68	AT3G25410					9	9		N	C	Sodium Bile acid symporter family	Homologous to known polar metabolite transporter
69	AT5G59250	x	x	x		11	11		N	C	Major facilitator superfamily protein-Carbohydrate transport	Homologous to known polar metabolite transporter
70	AT5G09930						0		N	C	ABC transporter family protein	unknown ABC transporter
71	AT1G54350		x			4	4		N	C	ABC transporter family protein	unknown ABC transporter

Table 1: (cont'd)

72	AT1G70610			x	x	4	3		N	C	transporter associated with antigen processing protein 1	unknown ABC transporter
73	AT2G01320				x	4	4		N	-	ABC-2 type transporter family protein WBC7	unknown ABC transporter
74	AT4G25450	x	x		x	5	5		N	C	non-intrinsic ABC protein 8	unknown ABC transporter
75	AT5G03910	x		x	x	6	2		N	C	ABC2 homolog 12	unknown ABC transporter
76	AT5G19410					5	5		N	C	ABC-2 type transporter family protein	unknown ABC transporter
77	AT5G52860					6	7		N	C	ABC-2 type transporter family protein	unknown ABC transporter
78	AT2G38330					12	10		N	C	MATE efflux family protein - Proton dependent	unknown MATE transporter
79	AT4G38380		x			13	9	17098813	N	C	MATE efflux family protein - Proton dependent	unknown MATE transporter
80	AT1G48460	x		x	x	5	5		N	-	tRNA-pcocessing ribonuclease (source Araport11)	unknown transporter
81	AT1G55930				x	5	5		N	C	CBS domain-containing protein / transporter associated domain-containing protein	unknown transporter
82	AT2G02590				x	6	6		N	C	small multi-drug export protein (source:Araport11)	unknown transporter
83	AT2G42770	x	x		x	2	2		N	-	Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein	unknown transporter
84	AT3G04440					9	8		N	C	Plasma-membrane choline transporter family protein	unknown transporter
85	AT3G13070				x	5	4		N	C	CBS domain-containing protein / transporter associated domain-containing protein	unknown transporter
86	AT3G52420						1	11595795; 12853455; 21515817	N	-	outer envelope membrane protein 7 (OEP7)	unknown transporter
87	AT3G60590			x	x	5	5		N	-	cytochrome P450 family protein (source:Araport11)	unknown transporter

Table 1: (cont'd)

88	AT3G63160	x	x	x	x	1	1		N	-	OEP6	unknown transporter
89	AT4G28210				x	7	3		N	C	embryo defective 1923, chloroplast organization, embryo development ending in seed dormancy, rRNA processing, tRNA metabolic process	unknown transporter
90	AT5G02940	x	x		x	3	2		N	-	ion channel POLLUX-like protein, protein of unknown function (DUF1012) (sourceAraport11)	unknown transporter
91	AT5G19750			x	x	3	2	17666025	N	C	Peroxisomal membrane 22 kDa-like (Mpv17/PMP22) family protein	unknown transporter
92	AT5G42960	x	x		x		0		N	C	OEP24-II - transporter of charged solutes	unknown transporter
93	AT5G43745	x		x	x	2	2		N	-	ion channel POLLUX-like protein, protein of unknown function (DUF1012)	unknown transporter
94	AT5G52540			x	x	9	10		N	C	keratin-associated protein, protein of unknown function (DUF819)	unknown transporter
95	AT3G11560	x		x	x	2	0		N	C	LETM1-like protein (PG in mutant?)	unknown transporter
96	AT3G52230	x	x	x	x		0		N	-	OM24 orthologue	unknown transporter
97	AT2G39190					2	0		N	C	ABC1 kinase 4 (ABC1K4) (likely plastid)	unknown transporter

REFERENCES

REFERENCES

Adebesin, F., Widhalm, J.R., Boachon, B., et al. (2017) Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter. *Science (80-.).*, **356**, 1386–1388.

Awai, K., Xu, C., Tamot, B. and Benning, C. (2006) A phosphatidic acid-binding protein of the chloroplast inner envelope membrane involved in lipid trafficking. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 10817–22.

Baginsky, S. and Gruissem, W. (2009) The Chloroplast Kinase Network: New Insights from Large-Scale Phosphoproteome Profiling. *Mol. Plant*, **2**, 1141–1153.

Benayoun, J. and Fahn, A. (1979) Intracellular transport and elimination of resin from epithelial duct-cells of Pinus halepensis. *Ann. Bot.*, **43**, 179–181.

Bick, J.A. and Lange, B.M. (2003) Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch. Biochem. Biophys.*, **415**, 146–154.

Binenbaum, J., Weinstain, R. and Shani, E. (2018) Gibberellin Localization and Transport in Plants. *Trends Plant Sci.*, **23**, 410–421.

Borghi, M., Fernie, A.R., Schiestl, F.P. and Bouwmeester, H.J. (2017) The Sexual Advantage of Looking, Smelling, and Tasting Good: The Metabolic Network that Produces Signals for Pollinators. *Trends Plant Sci.*, **22**, 338–350.

Breuers, F.K.H., Bräutigam, A. and Weber, A.P.M. (2011) The Plastid Outer Envelope - A Highly Dynamic Interface between Plastid and Cytoplasm. *Front. Plant Sci.*, **2**, 97.

Chen, L.-Q., Hou, B.-H., Lalonde, S., et al. (2010) Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature*, **468**, 527.

Chernomordik, L. V and Kozlov, M.M. (2008) Mechanics of membrane fusion. *Nat. Struct. Mol. Biol.*, **15**, 675–683.

Choi, Y.E., Lim, S., Kim, H.J., Han, J.Y., Lee, M.H., Yang, Y., Kim, J.A. and Kim, Y.S. (2012) Tobacco NtLTP1, a glandular-specific lipid transfer protein, is required for lipid secretion from glandular trichomes. *Plant J.*, **70**, 480–491.

Crotty, W.J. and Ledbetter, M.C. (1973) Membrane continuities involving chloroplasts and other organelles in plant cells. *Science (80-.).*, **182**, 839–41.

Crouzet, J., Roland, J., Peeters, E., Trombik, T., Ducos, E., Nader, J. and Boutry, M. (2013) NtPDR1, a plasma membrane ABC transporter from Nicotiana tabacum, is involved in diterpene transport. *Plant Mol. Biol.*, **82**, 181–192.

Fan, J., Zhai, Z., Yan, C. and Xu, C. (2015) Arabidopsis TRIGALACTOSYLDIACYLGLYCEROL5 Interacts with TGD1, TGD2, and TGD4 to Facilitate Lipid Transfer from the Endoplasmic Reticulum to Plastids. *Plant Cell*, **27**, 2941–55.

Ferro, M., Brugière, S., Salvi, D., et al. (2010) AT_CHLORO, a Comprehensive Chloroplast Proteome Database with Subplastidial Localization and Curated Information on Envelope Proteins. *Mol. Cell. Proteomics*, **9**, 1063–1084.

Fischer, K., Weber, A.P.M.. and Kunz, H.-H. (2016) The Transporters of Plastids-New Insghts into an Old Field. In H. Kirchhoff, ed. *Chloroplasts: Current Research and Future Trends*. Norfolk: Caister Academic Press, pp. 209–240.

Fleet, C.M., Yamaguchi, S., Hanada, A., Kawaide, H., David, C.J., Kamiya, Y. and Sun, T. (2003) Overexpression of AtCPS and AtKS in Arabidopsis Confers Increased ent -Kaurene Production But No Increase in Bioactive Gibberellins 1. *Plant Physiol.*, **132**, 830–839.

Flügge, U. and Gao, W. (2005) Transport of Isoprenoid Intermediates Across Chloroplast Envelope Membranes. *Plant Biol.*, **7**, 91–97.

Flügge, U.I. (2000) Transport in and out of plastids: does the outer envelope membrane control the flow? *Trends Plant Sci.*, **5**, 135–7.

Froehlich, J.E., Wilkerson, C.G., Ray, W.K., McAndrew, R.S., Osteryoung, K.W., Gage, D.A. and Phinney, B.S. (2003) Proteomic study of the Arabidopsis thaliana chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *J. Proteome Res.*, **2**, 413–25.

Geisler, K., Jensen, N.B., Yuen, M.M.S., Madilao, L. and Bohlmann, J. (2016) Modularity of Conifer Diterpene Resin Acid Biosynthesis: P450 Enzymes of Different CYP720B Clades Use Alternative Substrates and Converge on the Same Products. *Plant Physiol.*, **171**, 152–164.

Guo, Z. and Wagner, G. (1995) Biosynthesis of labdenediol and sclareol in cell-free extracts from trichomes of Nicotiana glutinosa. *Planta*, **197**, 1–10.

Gutensohn, M., Orlova, I., Nguyen, T.T.H., Davidovich-Rikanati, R., Ferruzzi, M.G., Sitrit, Y., Lewinsohn, E., Pichersky, E. and Dudareva, N. (2013) Cytosolic monoterpene biosynthesis is supported by plastid-generated geranyl diphosphate substrate in transgenic tomato fruits. *Plant J.*, **75**, 351–363.

Hamberger, B., Ohnishi, T., Hamberger, B., Séguin, A. and Bohlmann, J. (2011) Evolution of Diterpene Metabolism: Sitka Spruce CYP720B4 Catalyzes Multiple Oxidations in Resin Acid Biosynthesis of Conifer Defense against Insects. *Plant Physiol.*, **157**, 1677–1695.

Hedden, P. (2016) Gibberellin biosynthesis in higher plants. In *Annual Plant Reviews, Volume 49*. Chichester, UK: John Wiley & Sons, Ltd, pp. 37–72.

Hedden, P. and Thomas, S.G. (2012) Gibberellin biosynthesis and its regulation. *Biochem. J.*, **444**, 11–25.

Hurlock, A.K., Roston, R.L., Wang, K. and Benning, C. (2014) Lipid trafficking in plant cells. *Traffic*, **15**, 915–32.

Jessen, D., Roth, C., Wiermer, M. and Fulda, M. (2015) Two Activities of Long-Chain Acyl-Coenzyme A Synthetase Are Involved in Lipid Trafficking between the Endoplasmic Reticulum and the Plastid in Arabidopsis. *Plant Physiol.*, **167**, 351–366.

Kader, J.-C. (1996) Lipid-transfer proteins in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **47**, 627–654.

Kanno, Y., Hanada, A., Chiba, Y., Ichikawa, T., Nakazawa, M., Matsui, M., Koshiba, T., Kamiya, Y. and Seo, M. (2012) Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. *Proc. Natl. Acad. Sci.*, **109**, 9653–9658.

Kanno, Y., Oikawa, T., Chiba, Y., et al. (2016) AtSWEET13 and AtSWEET14 regulate gibberellin-mediated physiological processes. *Nat. Commun.*, **7**, 13245.

Kasahara, H., Hanada, A., Kuzuyama, T., Takagi, M., Kamiya, Y. and Yamaguchi, S. (2002) Contribution of the mevalonate and methylerythritol phosphate pathways to the biosynthesis of gibberellins in Arabidopsis. *J. Biol. Chem.*, **277**, 45188–94.

Keeling, C.I. and Bohlmann, J. (2006) Diterpene resin acids in conifers. *Phytochemistry*, **67**, 2415–23.

Keeling, C.I., Weisshaar, S., Ralph, S.G., Jancsik, S., Hamberger, B., Dullat, H.K. and Bohlmann, J. (2011) Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (Picea spp.). *BMC Plant Biol.*, **11**, 43.

Kim, S., Yamaoka, Y., Ono, H., et al. (2013) AtABCA9 transporter supplies fatty acids for lipid synthesis to the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 773–8.

LaBrant, E., Barnes, A.C. and Roston, R.L. (2018) Lipid transport required to make lipids of photosynthetic membranes. *Photosynth. Res.*, **0**, 0.

Langenheim, J.H. (2003) *Plant Resins: Chemistry, Evolution, Ecology, and Ethnobotany*, Portland, OR: Timber Press.

Léran, S., Varala, K., Boyer, J.-C., et al. (2014) A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. *Trends Plant Sci.*, **19**, 5–9.

Li, N., Gügel, I.L., Giavalisco, P., Zeisler, V., Schreiber, L., Soll, J. and Philippar, K. (2015) FAX1, a novel membrane protein mediating plastid fatty acid export. *PLoS Biol.*, **13**, e1002053.

Linka, N. and Weber, A.P.M. (2010) Intracellular metabolite transporters in plants. *Mol. Plant*, **3**, 21–53.

Lipko, A. and Swiezewska, E. (2016) Isoprenoid generating systems in plants - A handy toolbox how to assess contribution of the mevalonate and methylerythritol phosphate

pathways to the biosynthetic process. *Prog. Lipid Res.*, **63**, 70–92.

Liu, F., Zhang, X., Lu, C., Zeng, X., Li, Y., Fu, D. and Wu, G. (2015) Non-specific lipid transfer proteins in plants: Presenting new advances and an integrated functional analysis. *J. Exp. Bot.*, **66**, 5663–5681.

Lu, B., Xu, C., Awai, K., Jones, A.D. and Benning, C. (2007) A small ATPase protein of Arabidopsis, TGD3, involved in chloroplast lipid import. *J. Biol. Chem.*, **282**, 35945–53.

McLean, B., Whatley, J.M. and Juniper, B.E. (1988) Continuity of chloroplast and endoplasmic Reticulum membranes in Chara and Equisetum. *New Phytol.*, **109**, 59–65.

Mehrshahi, P., Johnny, C. and DellaPenna, D. (2014) Redefining the metabolic continuity of chloroplasts and ER. *Trends Plant Sci.*, **19**, 501–7.

Mehrshahi, P., Stefano, G., Andaloro, J.M., Brandizzi, F., Froehlich, J.E. and DellaPenna, D. (2013) Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 12126–31.

Nagata, N., Suzuki, M., Yoshida, S. and Muranaka, T. (2002) Mevalonic acid partially restores chloroplast and etioplast development in Arabidopsis lacking the non-mevalonate pathway. *Planta*, **216**, 345–350.

Nagel, R., Bernholz, C., Vranová, E., et al. (2015) Arabidopsis thaliana isoprenyl diphosphate synthases produce the C 25 intermediate geranylfarnesyl diphosphate. *Plant J.*, **84**, 847–859.

Nour-Eldin, H.H., Andersen, T.G., Burow, M., et al. (2012) NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature*, **488**, 531–534.

Pierman, B., Toussaint, F., Bertin, A., Lévy, D., Smargiasso, N., Pauw, E. De and Boutry, M. (2017) Activity of the purified plant ABC transporter NtPDR1 is stimulated by diterpenes and sesquiterpenes involved in constitutive and induced defenses. *J. Biol. Chem.*, **292**, 19491–19502.

Pottosin, I. and Dobrovinskaya, O. (2015) Ion Channels in Native Chloroplast Membranes: Challenges and Potential for Direct Patch-Clamp Studies. *Front. Physiol.*, **6**, 396.

Ro, D.-K., Arimura, G.-I., Lau, S.Y.W., Piers, E. and Bohlmann, J. (2005) Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 8060–5.

Roston, R.L., Gao, J., Murcha, M.W., Whelan, J. and Benning, C. (2012) TGD1, -2, and -3 proteins involved in lipid trafficking form ATP-binding cassette (ABC) transporter with multiple substrate-binding proteins. *J. Biol. Chem.*, **287**, 21406–15.

Saito, H., Oikawa, T., Hamamoto, S., et al. (2015) The jasmonate-responsive GTR1 transporter is required for gibberellin-mediated stamen development in Arabidopsis. *Nat. Commun.*, **6**, 6095.

Schmelz, E.A., Huffaker, A., Sims, J.W., Christensen, S.A., Lu, X., Okada, K. and Peters, R.J. (2014) Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J.*, **79**, 659–678.

Schnurr, J.A., Shockey, J.M., Boer, G.-J. de and Browse, J.A. (2002) Fatty acid export from the chloroplast. Molecular characterization of a major plastidial acyl-coenzyme A synthetase from Arabidopsis. *Plant Physiol.*, **129**, 1700–9.

Shibata, Y., Ojika, M., Sugiyama, A., Yazaki, K., Jones, D.A., Kawakita, K. and Takemoto, D. (2016) The Full-Size ABCG Transporters Nb-ABCG1 and Nb-ABCG2 Function in Pre- and Postinvasion Defense against Phytophthora infestans in Nicotiana benthamiana. *Plant Cell*, **28**, 1163–1181.

Skubatz, H., Kunkel, D.D., Patt, J.M., Howald, W.N., Hartman, T.G. and Meeuse, B.J. (1995) Pathway of terpene excretion by the appendix of Sauromatum guttatum. *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 10084–8.

Stukkens, Y., Bultreys, A., Trombik, T., Vanham, D. and Boutry, M. (2005) NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana plumbaginifolia*, plays a major role in plant pathogen defense. *Plant Physiol.*, **139**, 341–352.

Stukkens, Y., Degand, H., Purnelle, B., Marchand-brynaert, J. and Boutry, M. (2001) A Plant Plasma Membrane ATP Binding Cassette – Type Transporter Is Involved in Antifungal Terpenoid Secretion. , **13**, 1095–1107.

Sun, T. (2008) Gibberellin metabolism, perception and signaling pathways in Arabidopsis. *Arab. B.*, **6**, e0103.

Tal, I., Zhang, Y., Jørgensen, M.E., et al. (2016) The Arabidopsis NPF3 protein is a GA transporter. *Nat. Commun.*, **7**, 11486.

Talon, M., Koornneef, M. and Zeevaart, J. a (1990) Endogenous gibberellins in Arabidopsis thaliana and possible steps blocked in the biosynthetic pathways of the semidwarf ga4 and ga5 mutants. *Proc. Natl. Acad. Sci. U. S. A.*, **87**, 7983–7.

Tholl, D. and Lee, S. (2011) Terpene Specialized Metabolism in Arabidopsis thaliana. *Arab. B.*, **9**, e0143.

Ting, H.M., Delatte, T.L., Kolkman, P., Misas-Villamil, J.C., Hoorn, R.A.L. Van Der, Bouwmeester, H.J. and Krol, A.R. Van Der (2015) SNARE-RNAi results in higher terpene emission from ectopically expressed caryophyllene synthase in nicotiana benthamiana. *Mol. Plant*, **8**, 454–466.

Turner, G.W., Gershenzon, J. and Croteau, R.B. (2000) Development of peltate glandular trichomes of peppermint. *Plant Physiol.*, **124**, 665–80.

Turner, G.W. and Lange, B.M. (2015) Ultrastructure of grapefruit secretory cavities and immunocytochemical localization of (+)-limonene synthase. *Int. J. Plant Sci.*, **176**, 643–661.

Wang, B., Kashkooli, A.B., Sallets, A., et al. (2016) Transient production of artemisinin in Nicotiana benthamiana is boosted by a specific lipid transfer protein from A. annua. *Metab. Eng.*, **38**, 159–169.

Wang, C., Chen, Q., Fan, D., Li, J., Wang, G. and Zhang, P. (2016) Structural Analyses of Short-Chain Prenyltransferases Identify an Evolutionarily Conserved GFPPS Clade in Brassicaceae Plants. *Mol. Plant*, **9**, 195–204.

Wang, Z., Anderson, N.S. and Benning, C. (2013) The Phosphatidic Acid Binding Site of the Arabidopsis Trigalactosyldiacylglycerol 4 (TGD4) Protein Required for Lipid Import into Chloroplasts. *J. Biol. Chem.*, **288**, 4763–4771.

Whatley, J.M., McLean, B. and Juniper, B.E. (1991) Continuity of chloroplast and endoplasmicreticulum membranes in Phaseolus vulgaris. *New Phytol.*, **117**, 209–217.

Widhalm, J.R., Jaini, R., Morgan, J.A. and Dudareva, N. (2015) Rethinking how volatiles are released from plant cells. *Trends Plant Sci.*, **20**, 545–550.

Xiao, S., Li, H.Y., Zhang, J.P., Chan, S.W. and Chye, M.L. (2008) Arabidopsis acyl-CoA-binding proteins ACBP4 and ACBP5 are subcellularly localized to the cytosol and ACBP4 depletion affects membrane lipid composition. *Plant Mol. Biol.*, **68**, 571–583.

Xu, C., Fan, J., Cornish, A.J. and Benning, C. (2008) Lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis requires the extraplastidic TGD4 protein. *Plant Cell*, **20**, 2190–204.

Xu, C., Fan, J., Riekhof, W., Froehlich, J.E. and Benning, C. (2003) A permease-like protein involved in ER to thylakoid lipid transfer in Arabidopsis. *EMBO J.*, **22**, 2370–9.

Zerbe, P. and Bohlmann, J. (2015) Plant diterpene synthases: Exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol.*, **33**, 419–428.

Zhou, Y., Dorp, K. vom, Dormann, P. and Holzl, P. (2016) Chloroplast Lipids. In *Chloroplasts: Current Research and Future Trends*. pp. 1–24.

Zi, J., Mafu, S. and Peters, R.J. (2014) To Gibberellins and Beyond! Surveying the Evolution of (Di)Terpenoid Metabolism. *Annu. Rev. Plant Biol.*, **65**, 259–286.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q. and Wijk, K.J. van (2008) Sorting Signals, N-Terminal Modifications and Abundance of the Chloroplast Proteome. *PLoS One*, **3**, e1994.

CHAPTER 2: Probing the biochemical continuity of the plastid and ER using the gibberellin biosynthetic pathway

Abstract:

Gibberellin (GA) biosynthesis is initiated by diterpene synthases localized to the plastid with intermediates being oxidized by endoplasmic reticulum (ER)-localized P450s and finally bioactive GAs are produced by activities resident in the cytosol. Despite intensive study, plastid- or ER-localized transporter mediating movement of non-polar GA intermediates, *ent*-kaurene and *ent*-kaurenoic acid, have not been identified, a theme common in the synthesis of many non-polar plant metabolites. Previous work has shown that non-polar substrates in the plastid-localized tocopherol and carotenoid pathways were accessible by pathway enzymes retargeted to the ER lumen, suggesting a novel bidirectional mechanism exists for access of non-polar carotenoid and tocopherol biosynthetic compounds between the two organelles. In this work, we probed the *in vivo* accessibility of *ent*-kaurene produced in the ER by kaurene oxidase (KO) localized to the outer envelope membrane (OEM). This was achieved by transplanting the first two steps of the pathway, copalyl diphosphate synthase (CPS) and kaurene synthase (KS), to the ER lumen in mutant background(s) deficient for one or both activities and assessing for functional complementation of GA synthesis. Single targeting of CPS, but not KS, to the ER complemented the dwarf phenotype of its corresponding mutant. Quantification of the GA pathway intermediates showed that ER-targeted CPS transgenic lines have levels similar to wild-type and suggests that the product of CPS, *ent*-CDP, is transported in a unidirectional fashion from the ER into the plastid. Retargeting of both CPS and KS to the ER in the double mutant background also complemented the dwarf phenotype, indicating that OEM-localized KO can access its substrate, *ent*-kaurene, when it is synthesized in the ER lumen instead of the plastid stroma.

Introduction:

Plastids are the defining organelle of photosynthetic eukaryotes. In plants, plastids represent a group of developmentally-related organelles that include undifferentiated proplastids, photosynthetically active chloroplasts, starch-storing amyloplasts and terpenoid-producing leucoplasts (Jarvis and López-Juez, 2013; Lopez-Juez and Pyke, 2005; Pyke, 2007), the differentiation of which depends on the cell types and developmental stage of the tissue. Plastids are also metabolic hubs involved in carbon, nitrogen, and sulfur assimilation and the site for synthesis of an enormous variety of compounds that include fatty acids, terpenoids, amino acids, and numerous hormones (Jarvis and López-Juez, 2013). The double-membrane envelope delineates plastids from the cytosol. While the outer envelope membrane (OEM) is permeable to small molecules (<10 kD) (Breuers *et al.*, 2011), the inner envelope membrane (IEM) is a tightly controlled permeability barrier (Facchinelli and Weber, 2011).

Many plant pathways span multiple compartments in the cell, including plastids, and exchange of numerous substrates and pathway intermediates occurs between plastids and other subcellular compartments. Transport across the OEM is largely mediated by β-barrel proteins termed outer envelope proteins (OEP) that have broad substrate specificities (Flügge, 2000; Breuers *et al.*, 2011; Pottosin and Dobrovinskaya, 2015; Fischer *et al.*, 2016). Transport across the IEM, in contrast, is mediated by α-helical proteins with high substrate specificities (Weber and Linka, 2011; Linka and Weber, 2010; Rolland *et al.*, 2012). Various biochemical, genetic and proteomic studies have elucidated the activities of fifty-one polar metabolite transporters in the transport of organic acids, nucleotide sugars, metals, ions and other polar metabolites, but only two transporters of non-polar metabolites have been identified. These are the trigalactosyldiacylglycerol (TGD) transporter complex and fatty acid export 1 (FAX1) (Froehlich

et al., 2003; Zybailov *et al.*, 2008; Baginsky and Gruissem, 2009; Ferro *et al.*, 2010; Mehrshahi *et al.*, 2013; Fischer *et al.*, 2016), both of which are involved in transport of membrane lipids/fatty acids. This disparity in the numbers of characterized polar compound transporters versus non-polar compound transporters raises fundamental questions of how plastid-synthesized non-polar pathway intermediates (monoterpene and diterpene olefins) are transported from plastids to the endoplasmic reticulum (ER) during synthesis.

Gibberellins (GAs) are one class of diterpenes (20-carbon terpenoids) that as phytohormones play essential roles in processes such as seed germination, cell elongation, and flowering (Sun, 2008). GA synthesis is initiated in the stroma of plastids where two diterpene synthases, copalyl diphosphate synthase (CPS, encoded by the *GA1* locus) and kaurene synthase (KS, encoded by the *GA2* locus) cyclize the general precursor geranylgeranyl diphosphate (GGDP) into the diterpene olefin, *ent*-kaurene (Figure 2.1). Kaurene oxidase (KO) in the OEM catalyzes the three-step oxidation reaction of *ent*-kaurene to *ent*-kaurenoic acid, which is the substrate for additional oxidations by ER-localized kaurenoic acid oxidases that convert it to GA_{12}. Subsequent steps of the GA pathway occur in the cytosol, where GA_{12} serves as a branch point intermediate, undergoing hydroxylation at the C-13 position or not. GA_{24}, an intermediate in the non C-13 hydroxylated GA branch, is ultimately converted to the bioactive GA_4, while GA_1 is the C-13 hydroxylated form of bioactive GA. The GA biosynthetic pathway has been intensely studied since the 1980s, resulting in the identification of all biosynthetic steps and also a large number of components involved in GA perception, transcriptional regulation, and turnover (Hedden, 2016). However, despite these efforts, no insight has been gained into how *ent*-kaurene or *ent*-kaurenoic acid are transported between organelles, a common theme for most other non-polar metabolite pathways in plants.

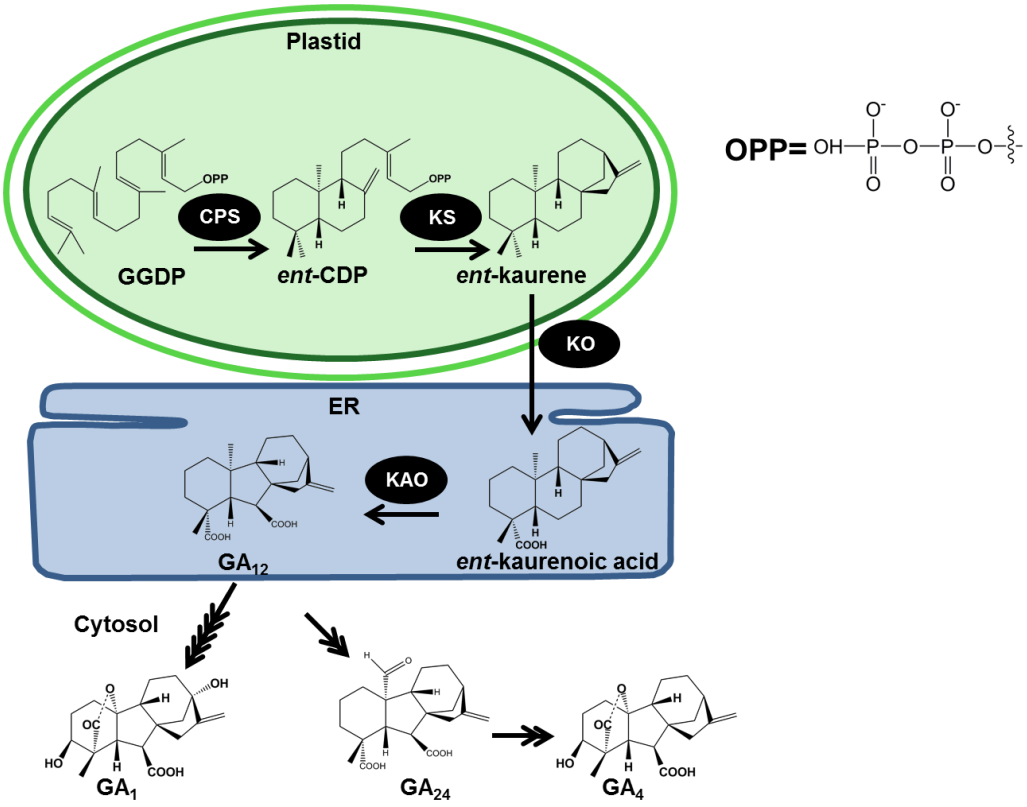


Figure 2.1: Gibberellins (GAs) biosynthesis in Arabidopsis. The plastid-localized diterpene synthases, copalyl diphosphate synthase (CPS) and kaurene synthase (KS) catalyze the cyclization of geranylgeranyl diphosphate (GGDP) first to *ent*- copalyl diphosphate (*ent*-CDP), and subsequently to *ent*-kaurene. *ent*-Kaurene (a non-polar compound) is then converted to *ent*-kaurenoic acid by a cytochrome P450 on the outer chloroplast envelope membrane, kaurene oxidase (KO). *ent*-Kaurenoic acid is then converted to GA_{12} by endoplasmic reticulum (ER)-localized cytochrome P450s, kaurenoic acid oxidases. GA_{12} lies at the branch point in the pathway, either undergoing hydroxylation at the C-13 position or not. Two families of soluble 2-oxoglutarate dependent dioxygenases catalyze the formation of GA_{24} (an intermediate of the non-13-hydroxylated GAs) and eventually form the bioactive GAs, $GA_{1,}$ and $GA_{4.}$ Figure adapted from Hedden and Thomas, 2012. Single arrow indicates one reaction step while multiple arrows indicate multiple steps. **OPP** represents the diphosphate group.

In a previous study, we developed a functional complementation approach, termed transorganellar complementation, to determine the *in vivo* accessibility of compounds in one compartment by enzymes in another organelle (Mehrshahi *et al.*, 2013). Briefly, a mutant line defective for an activity is tested for complementation by expressing the functional enzyme targeted to the native compartment (e.g., the plastid, as a positive control) or companion

organelle (e.g., the ER). Using this approach we demonstrated that multiple activities in the tocopherol and carotenoid biosynthetic pathways, whose substrates and enzymes are restricted to the plastid, could access their biosynthetic intermediates from the ER lumen in a bidirectional fashion (Mehrshahi *et al.*, 2013; Mehrshahi *et al.*, 2014). This suggested the existence of a novel interface between the plastid and ER that provides the companion organelle with access to non-polar compounds in the native organelle, even when the pathways and compounds are not present in the companion organelle (Mehrshahi *et al.*, 2013; Mehrshahi *et al.*, 2014). This prior study focused on pathways and substrates solely localized to the plastid (Mehrshahi *et al.*, 2013); the present study utilizes the GA pathway, to assess the bidirectional access by enzymes in a pathway that natively spans the plastid and ER.

Results:

Phenotypes of the GA biosynthetic mutants

The *ga1-6* and *ga2-1* mutants used in this study are EMS mutants defective in CPS and KS activities, respectively, in the *Landsberg erecta* background (Koornneef and van der Veen, 1980). The CPS in *ga1-6* contains a Ser151 to Phe151 mutation (Keeling, Madilao, *et al.*, 2011) while the KS in *ga2-1* has a premature stop codon that eliminates 14 kDa of the enzyme's carboxy terminus (Yamaguchi *et al.*, 1998). The *ga1-6* mutant is considered leaky, as it can germinate without exogenous GA treatment but still retains a dwarf mutant phenotype without continuous GA application during the vegetative growth (Koornneef and van der Veen, 1980; Sun *et al.*, 1992). The *ga2-1* mutant, on the other hand, is considered a severe null mutant, as without exogenous GA application it remains a non-germinating, male-sterile, and extreme dwarf (Zeevaart and Talon, 1992; Yamaguchi *et al.*, 1998). Bioactive GAs have been reported to

be drastically reduced in both mutants (Zeevaart and Talon, 1992) and in the absence of exogenous GA application, both mutants develop into dwarf, dark green, and sterile plants. However, with continuous GA application, both mutants develop into plants phenotypically indistinguishable from wild-type (Figure 2.2, A and B). Thus, complementation experiments for CPS and KS activities can be assessed at two levels, physiologically at the whole plant phenotype and analytically for the levels of GA pathway intermediates accumulated.

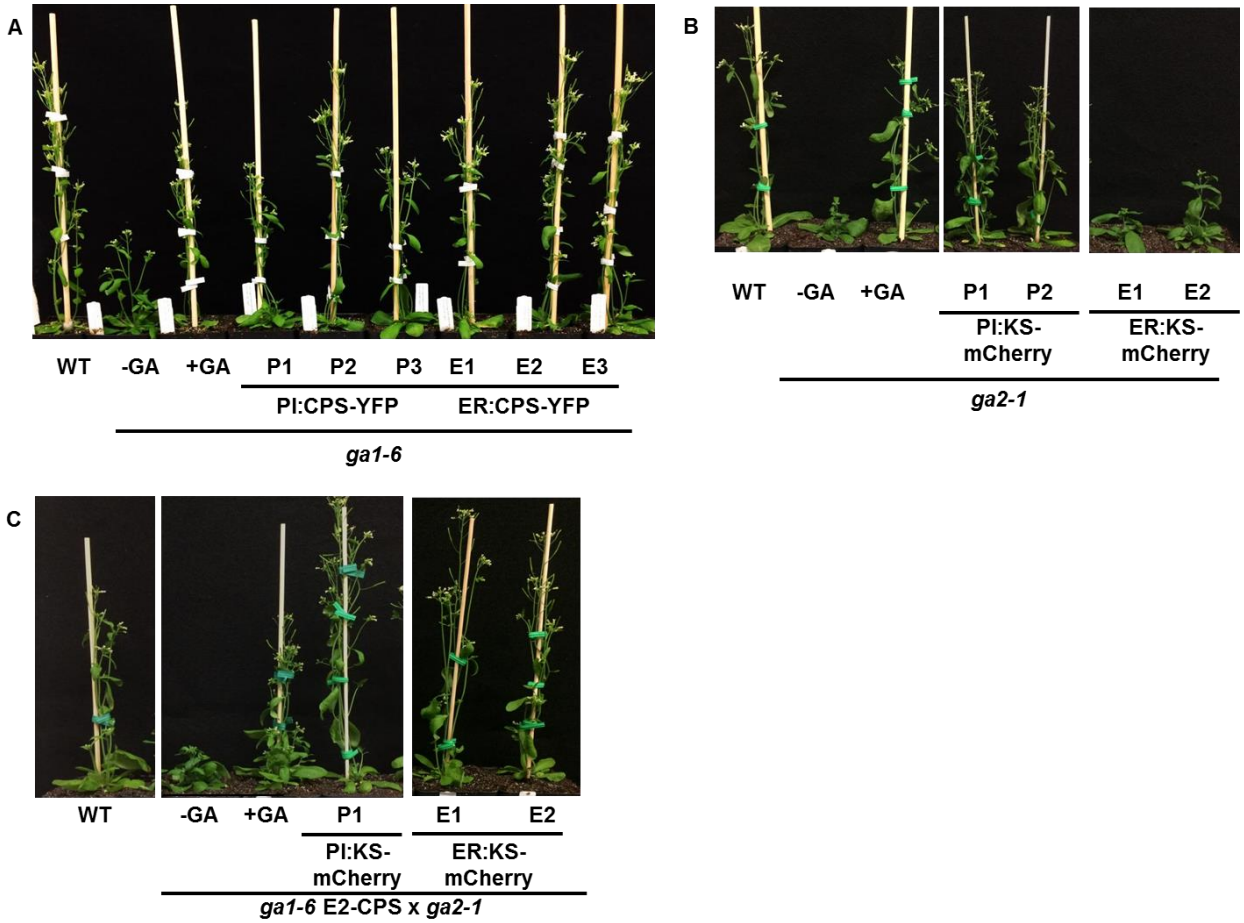


Figure 2.2: Growth and development of transgenic lines.
A) Growth of five-week-old plants for plastid- (P1, P2, and P3) or ER- (E1, E2, and E3) targeted CPS-YFP in the *ga1-6* mutant background on soil. B) Growth of five-week-old plants for plastid- (P1 and P2) or ER- (E1 and E2) targeted KS-mCherry in the *ga2-1* mutant background on soil. C) Growth of five-week-old plastid- (P1-KS) and ER- (E1 and E2) targeted KS-mCherry in the *ga1-6* E2-CPS x *ga2-1* background. The *ga1-6* E2-CPS x *ga2-1* background was generated by crossing E2-CPS in (A) to a *ga2-1* mutant. All alleles and loci were confirmed to be homozygous before transformation of the KS-mCherry construct. +GA plants were sprayed once a week with 100 μM GA_3. Wild-type (WT) *Landsberg* plants are shown for comparison.

Expression of CPS and KS in the plastids and ER and whole plant phenotypes of mutants and complemented mutants

CPS and KS are localized in the stroma (Sun and Kamiya, 1994; Helliwell *et al.*, 2001) and to test their abilities to access their respective substrate from a different subcellular compartment and complement their respective mutations, we engineered both proteins for expression and re-targeting to the plastid and ER. Their predicted N-terminal chloroplast transit peptides (cTPs) and cleavage sites were estimated computationally using ChloroP (http://www.cbs.dtu.dk/services/ChloroP/) and also by comparing protein sequence alignments of orthologs from a range of plant species, as transit peptides show little evolutionary conservation. Based on these data, we removed 62 and 33 N-terminal amino acids from CPS and KS, respectively. For ER retargeting, retention and detection (by confocal microscopy and immunoblots using antibodies to YFP or mCherry), we added a 21 amino acid ER signal peptide (Matsuoka *et al.*, 1995) to the N-termini of the truncated proteins and a fluorescent protein (YFP for CPS and mCherry-FLAG for KS) followed by an HDEL ER retention signal to their C-termini. These constructs are designated ER:CPS-YFP and ER:KS-mCherry. As positive controls for complementation, CPS and KS constructs with their native cTPs were generated with the same fluorescent reporter (designated plastid:CPS-YFP and plastid:KS-mCherry) to assess for complementation in their native plastid compartment. All constructs were introduced into their respective mutant backgrounds and at least four to ten independent transgenic lines were assessed for target protein expression, localization and single insertion segregation patterns. Two to three lines for each construct were selected and used for further analyses (Figure 2.3).

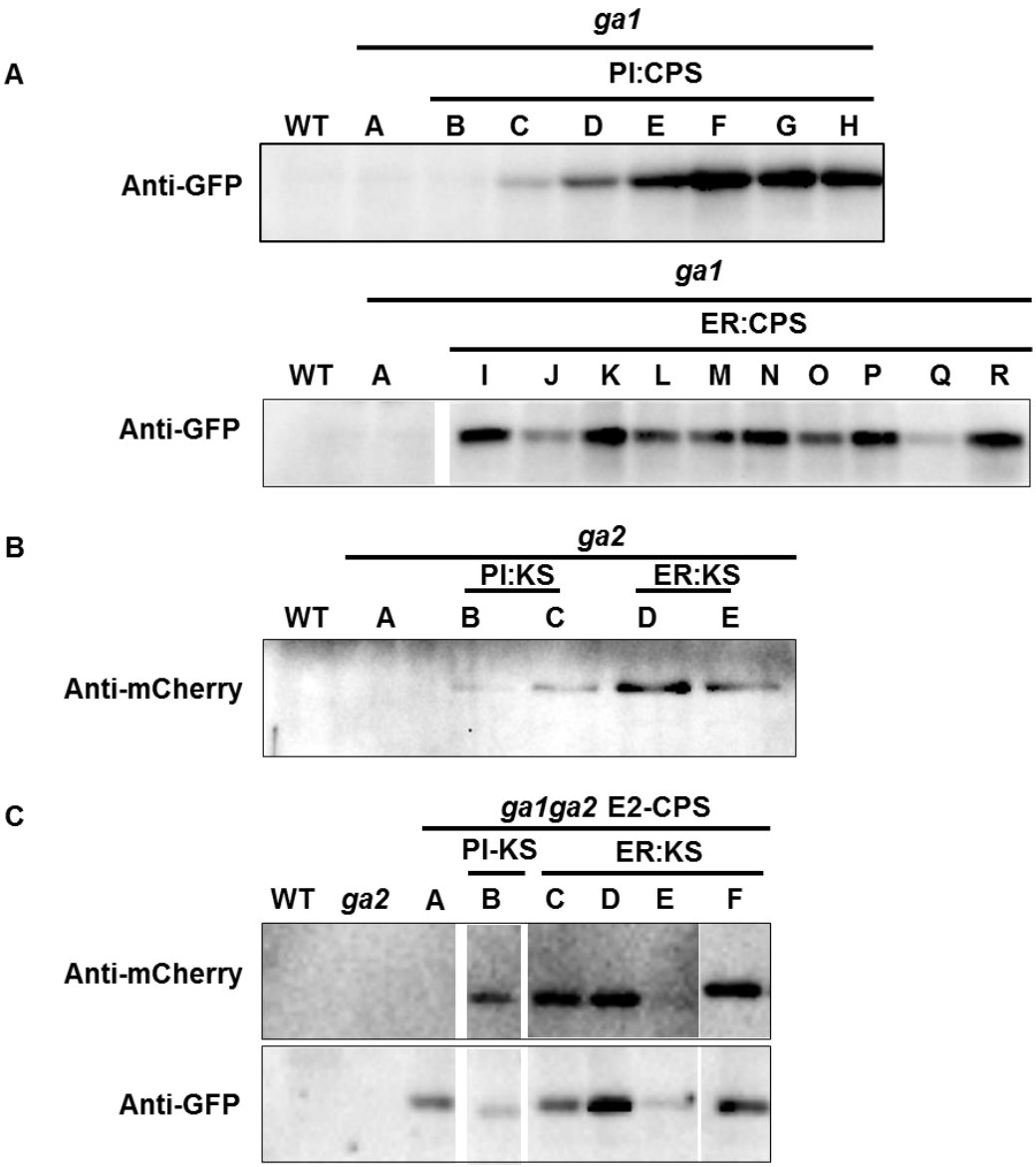


Figure 2.3: Protein expression in transgenic lines. A) Selection of CPS transgenic lines in the *ga1-6* mutant background. Eight μg of total protein was loaded onto an 8%SDS-PAGE gel. CPS-YFP expression in the independent transgenic lines was assessed by immunoblotting with anti-GFP antibody. A: the *ga1-6* mutant; B: P1-CPS; D: P2-CPS; H: P3-CPS; I: E2-CPS; L: E3-CPS and Q: E1-CPS were selected for further characterization.

B) Selection of the KS transgenic lines in the *ga2-1* mutant background. Ten μg of total protein was loaded onto an 8% SDS-PAGE gel. KS-mCherry expression in independent transgenic lines was assessed by immunoblotting with anti-mCherry antibody. A: the *ga2-1* mutant; B:P1-KS; C:P2-KS; D:E1-KS; E:E2-KS.

C) Selection of KS transgenic lines in the *ga1-6* E2-CPS x *ga2-1* double mutant background. Ten μg of total protein was loaded onto an 8% SDS-PAGE gel. KS-mCherry expression in the independent transgenic lines was assessed by immunoblotting with anti-mCherry antibody, while CPS-YFP expression was assessed by immunobotting with anti-GFP antibody. A: E2-CPS in the *ga1-6 ga2-1*double mutant; B:P1-KS; C:E1-KS; D:E2-KS

As anticipated, plastid:CPS-YFP and plastid:KS-mCherry both phenotypically complemented their respective *ga1-6* and *ga2-1* mutant, yielding transgenic plants that were visually indistinguishable from wild-type (WT) (Figure 2.2, A and B). Confocal imaging of stable transgenic lines expressing plastid:CPS-YFP and transiently expressed of plastid:KS-mCherry in tobacco showed that both tagged-proteins are plastid-localized as evidenced by colocalization of their respective fluorescent signals with chlorophyll fluorescence (Figure 2.4, A and B). Three plastid:CPS-YFP lines (P1-CPS, P2-CPS, and P3-CPS, respectively) expressing low, medium and high levels of fusion protein detected with an anti-GFP antibody were selected for further analyses (Figure 2.3 A). Similarly, two plastid:KS-mCherry (P1-KS, and P2-KS) expressing low and medium levels of fusion protein detectable by anti-mCherry antibody were also selected (Figure 2.3 B).

CPS and KS act sequentially to produce *ent*-kaurene in the plastid, and as the overall goal of this work was to generate a pool of *ent*-kaurene in the ER to test for its accessibility by KO localized in the OEM, both CPS and KS needed to be retargeted into the ER lumen in a *ga1-6 ga2-1* double mutant background, which lacks both activities in the plastid. As negative controls for the experiment, ER:CPS-YFP and ER:KS-mCherry were independently targeted to the ER lumen by transformation of their corresponding single mutant backgrounds, *ga1-6* and *ga2-1*, respectively. Imaging by confocal microscopy showed that fluorescent signals for both stable transgenic line expressing ER:CPS-YFP and transient expression of ER:KS-mCherry in tobacco localized to reticulate structures characteristic of the ER, with no detectable signal colocalizing to plastids or the cytosol (Figure 2.4, C and D). As further confirmation of ER localization, a homozygous ER:CPS-YFP line was stably transformed with an ER marker [a cyan fluorescent protein (CFP) with the sporamin signal peptide and an HDEL retention signal; (Brandizzi *et al.*,

2004)]. Imaging of ER:CPS-YFP and ER:CFP double transgenic showed both signals colocalized to the same reticulate structures (Figure 2.4E). As with the plastid-targeted lines, three ER:CPS-YFP lines (E1-CPS, E2-CPS, and E3-CPS) and two ER:KS-mCherry lines (E1-KS, and E2-KS) with differing levels of immunologically detectable proteins were selected for further analysis.

We had anticipated that neither the *ga1-6* nor *ga2-1* mutation could be complemented by expression of their corresponding WT proteins in the ER lumen as this would require movement of the polar CPS product and KS substrate, *ent*-copalyl diphosphate (*ent*-CDP) (Figure 2.1) across the ER and plastid envelope membranes. In the case of ER:CPS-YFP expressed in the *ga1-6* background, *ent*-CDP would be produced in the ER, while the endogenous KS is plastid-localized. Similarly, for ER:KS-mCherry expressed in the *ga2-1* background, *ent*-CDP would be produced in the stroma by the endogenous CPS and therefore would not be accessible by ER-localized KS. Assessment of the whole plant phenotypes of transformants indicated that this is true for ER:KS-mCherry in the *ga2-1* background, transgenic lines were indistinguishable from the *ga2-1* background in growth and seed set in the absence of GA_3 treatment (Figure 2.2B). However, plants expressing ER:CPS-YFP in the *ga1-6* background, complemented the mutation and were visually indistinguishable from WT, suggesting endogenous KS in the plastid stroma was able to access *ent*-CDP produced in the ER lumen.

As the goal of this study is to study the *in vivo* movement of *ent*-kaurene produced in the ER and assess its accessibility by kaurene oxidase localized on the OEM, we generated a *ga1 ga2* double mutant background by crossing *ga2-1* with *ga1-6* E2-CPS. Homozygosity at all loci was determined by genotyping and segregation of the resistance marker for the ER:CPS-YFP transgene. Introducing the *ga2-1* mutation in the E2-CPS background suppressed the WT

phenotype of E2-CPS to that of a GA dwarf mutant, indicating that the complementation observed in E2-CPS was indeed due to plastid-localized KS activity (Figure 2.2C). We then separately transformed the plastid: and ER:KS-mCherry constructs into the *ga1-6* E2-CPS x *ga2-1* background. As anticipated, introducing plastid:KS-mCherry into the *ga1-6* E2-CPS x *ga2-1* background complemented the dwarf mutant phenotypes (Figure 2.2C). Likewise, plants expressing both ER-targeted CPS and KS in the *ga1ga2* background also complemented the dwarf mutant phenotype, indicating that KO is able to access *ent*-kaurene synthesized in the ER (Figure 2.2C).

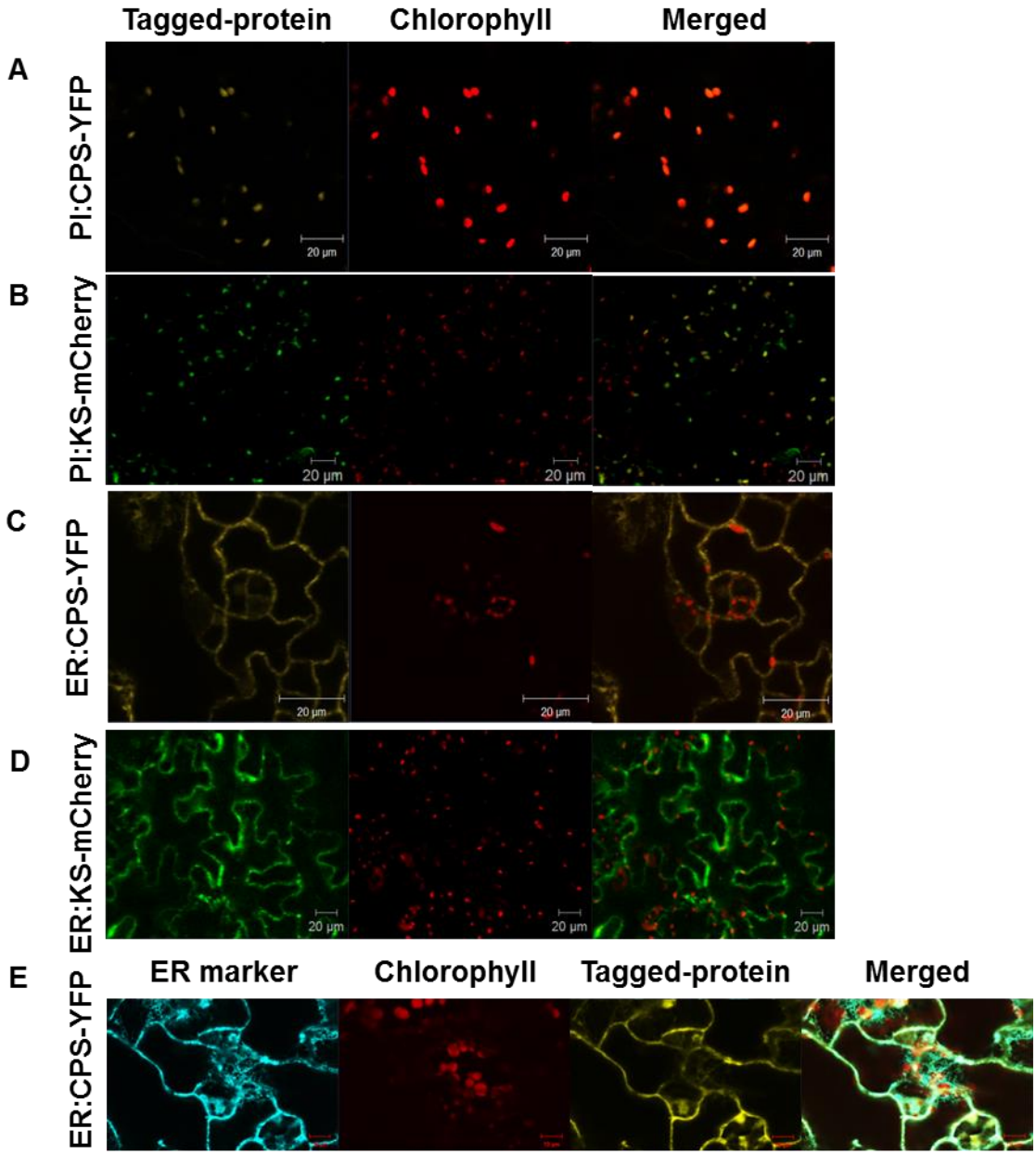


Figure 2.4: Subcellular localization of fluorescent-tagged CPS and KS proteins.
A) In plastid:CPS-YFP, YFP signals (*yellow*) colocalized with chlorophyll autofluorescence (*red*).
B) In plastid:KS-mCherry, mCherry signals (*green*) colocalized with chlorophyll autofluorescence (*red*).
C) In ER:CPS-YFP, YFP (*yellow*) signals did not overlap with chlorophyll autofluorescence (*red*).
D) In ER:KS-mCherry, mCherry (*green*) did not overlap with chlorophyll autofluorescence (*red*).
E) ER:CPS-YFP (*yellow*) colocalizes with ER marker signal (*blue*), but not with chlorophyll autofluorescence (*red*).

Subcellular fractionation of Plastid:CPS-YFP and ER:CPS-YFP

To independently corroborate the subcellular localization of retargeted proteins determined by confocal microscopy, we isolated intact chloroplasts and crude microsome fractions from three-week-old plate-grown seedlings of P2-CPS and E2-CPS (Figure 2.5A). Localization of the retargeted fusion proteins and known organellar markers were assessed by immunoblots with antibody to GFP (which cross-reacts with YFP) and known organelle compartment markers including the translocon at the outer membrane of chloroplast 75 (Toc75) for the OEM, the oxygen evolving complex 33 (OE33) for the thylakoid lumen and the luminal binding protein (BiP) as a marker for the ER lumen (Tranel *et al.*, 1995; Hashimoto *et al.*, 1997; Oliver *et al.*, 1995). Immunoblot analysis of P2-CPS showed CPS-YFP signal in intact chloroplasts and no detectable signal in the crude microsome fraction (Figure 2.5A). In contrast, E2-CPS showed CPS-YFP signal in crude microsomes and none detectable in the intact chloroplast fraction. The vast majority of BiP signal was in the crude microsome fraction of both lines and only very faint signals present in the chloroplast fraction, indicating slight contamination of ER in the chloroplast fraction (Figure 2.5A). The signal for OE33 was only present in intact chloroplasts in both lines, however, the signal for Toc75 was present in both the chloroplast and crude microsome fractions indicating contamination of broken chloroplast envelope membranes in the crude microsome fraction, but little to no thylakoid contamination (Figure 2.5A), as has been reported previously (Lord, 1987; Wang *et al.*, 2012).

To assess whether the YFP signal in E2-CPS microsomes was in the ER lumen or external to the ER membrane, protease protection assays of the crude microsome fraction were performed (Figure 2.5B). In the presence of thermolysin, only the OEM marker, Toc34 whose N-terminus is exposed to the cytosol (Sveshnikova *et al.*, 2000) was digested by the protease. In the

presence of thermolysin and the detergent Triton X-100, both the BiP and E2-CPS were also degraded, consistent with both BiP and ER:CPS-YFP in the microsome fraction being localized in the ER lumen.

Because high level of Toc75 in the crude microsomes indicated their contamination with OEMs, we further separated the crude microsome fraction on a continuous sucrose (20-50 %, w/w) gradient in the presence or absence of $MgCl_2$ (Figure 2.5C) and assessed the distribution of organelle compartment markers and CPS-YFP. Mg^{2+} allows ribosomes to remain attached to the ER, thereby inducing a shift towards higher density fractions in the sucrose gradient (Lord, 1987) while removal of Mg^{2+} shifts the ER fractions towards lower densities, resulting a characteristic shift of ER-associated, but not other membrane-associated proteins (Schaller, 2017). We isolated crude microsome fraction from the E3-CPS line and separated this fraction on the continuous sucrose gradient. Immunoblot analysis of continuous sucrose gradient fractions in the absence of $MgCl_2$ showed that the Toc75, BiP and YFP markers are all present in lower density fractions (fractions 1 through 9) with the BiP and YFP peaks at a somewhat higher density than Toc75. In contrast, in the presence of $MgCl_2$, both the BiP and the CPS-YFP signals shifted substantially and coordinately to a higher density in the gradient, peaking at fractions 8-11, while the Toc75 marker moved only slightly (Figure 2.5C). Thus, data from multiple techniques confirmed that ER:CPS-YFP is correctly targeted and localized to the ER lumen, indicating that the physiological complementation observed in ER:CPS-YFP transgenic lines is not due to protein mistargeting.

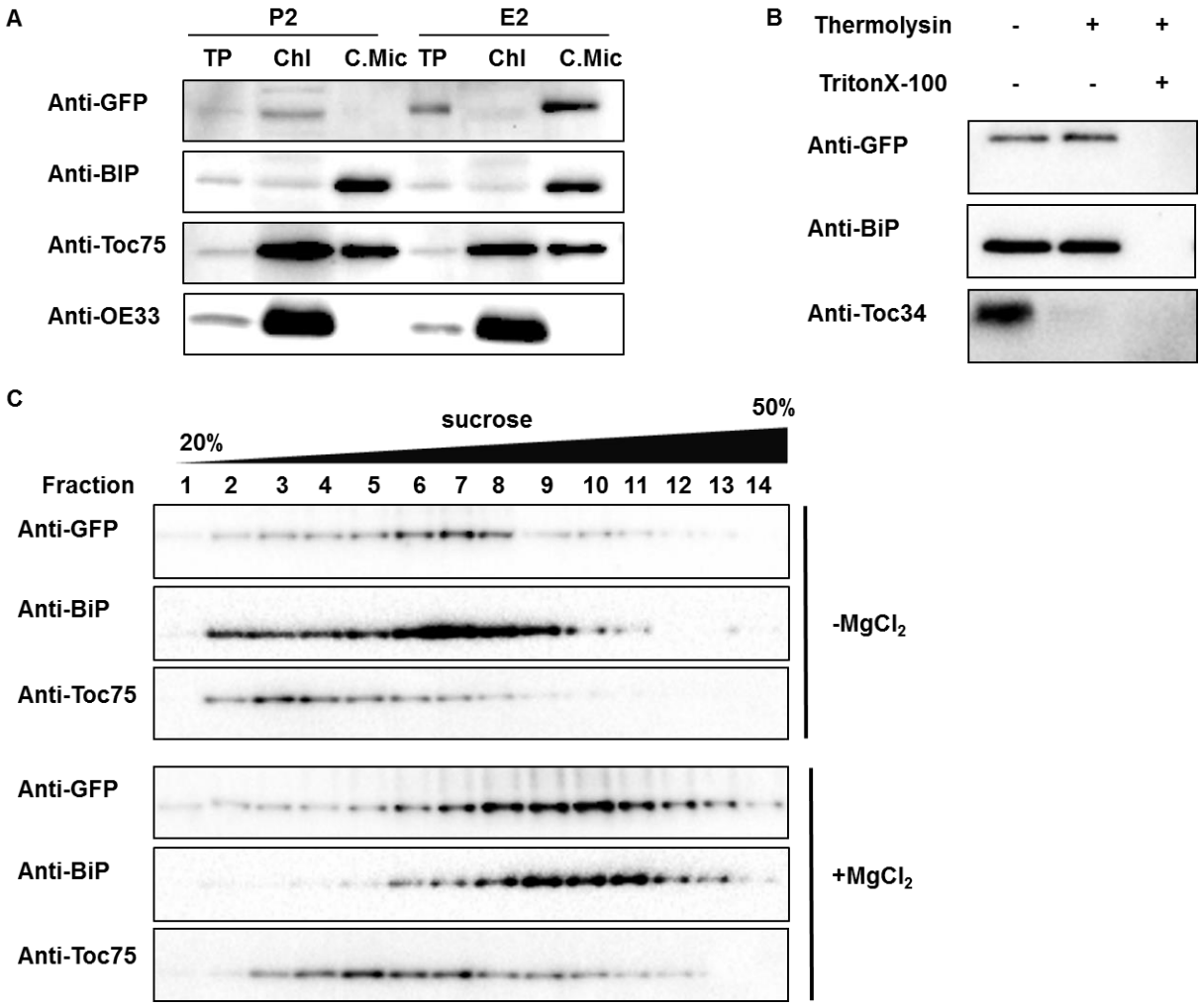


Figure 2.5: Subcellular fractionation of plastid:CPS-YFP and ER:CPS-YFP transgenic lines.

A) Immunoblots are shown for CPS-YFP, BiP and Toc75. In the P2-CPS line, CPS-YFP proteins were present in the total protein (TP) and chloroplast (Chl) fractions while in the E2-CPS line, CPS-YFP proteins were present in the TP and crude microsome (C.Mic) fractions. Enrichment of organelles in each fraction was tested by antibodies against the luminal binding protein (BiP) for ER, the translocon at the outer membrane of chloroplasts 75 (Toc75) for the OEM and the oxygen evolving complex 33 (OE33) for the thylakoid. Exposure time for all immunoblots was 30s.

B) Protease digestion of the E2-CPS crude microsome. The E2-CPS crude microsome fraction was treated with thermolysin in the presence or absence of TritonX-100 and subjected to immublotting. CPS-YFP and BiP were protected from thermolysin digestion while Toc34, an integral membrane protein with its N-terminus exposed to the cytosol, was digested. In the presence of TritonX-100 which disrupts membrane integrity, all proteins were digested. Exposure time: BiP and Toc34, 30s; GFP, 4 min.

C) CPS-YFP protein is localized to the endoplasmic reticulum (ER) in the E3-CPS transgenic line. The crude microsome fraction isolated from the E3-CPS was fractionated over continuous 20-50 % (w/w) sucrose gradient in the presence or absence of $MgCl_2$. An equal volume of each

Figure 2.5: (cont'd) fraction was analyzed by immunoblot with GFP, BiP, and Toc75 antibodies. Exposure time: GFP, 2 min; BiP, 10 s; Toc75, 1 min.

Metabolite levels in wild type, the *ga1-6* mutant and CPS-YFP complemented lines

The results presented thus far indicated that the *ga1-6* mutant and *ga1-6 ga2-1* double mutant, but not the *ga2-1* mutant, can be complemented physiologically by expression of their corresponding WT enzyme(s) in the ER lumen (Figure 2.2). This suggested that at least some level of bioactive GAs was being produced in these complemented lines. As most endogenous GA biosynthetic intermediates are too low *in vivo* to robustly quantify, we used two complementary approaches to provide insight into the biochemical consequences of complementation. First, we treated plants with paclobutrazol (PAC), an inhibitor of KO that results in the accumulation of *ent*-kaurene *in vivo*, which can then be quantified by gas chromatography-mass spectrometry (GC-MS) with selected ion monitoring. Quantification of *ent*-kaurene by this method was used as a proxy for pathway flux by the combined action of CPS and KS. In a second approach, we quantified two GA intermediates in untreated plants by stable isotope dilution with liquid chromatography-mass spectrometry (LC-MS): GA_{12}, a key intermediate produced in the ER and GA_{24}, an intermediate of the non-C13 hydroxylated GA branch that is produced in the cytosol (Figure 2.1).

Analysis of four-week-old rosettes from WT and *ga1-6* for *ent*-kaurene in the absence of PAC treatment showed the levels of *ent*-kaurene in both lines was near detection limits (approximately 2 ng g^{-1} fresh weight) and not significantly different (Figure 2.6A). When four-week-old plants were treated with PAC for four days, the levels of *ent*-kaurene in WT increased 20-fold to approximately 200 ng g^{-1} FW while in the *ga1-6* mutant, *ent*-kaurene only increased about five-fold less than WT (Figure 2.6A), consistent with the leaky nature of the *ga1-6*

mutation (Koornneef *et al.*, 1983; Sun and Kamiya, 1994). In PAC-treated plastid:CPS-YFP transgenic lines, the levels of *ent*-kaurene were correlated with protein expression. In P1-CPS, which has barely detectable fusion protein, *ent*-kaurene levels were similar to WT, while in P2-CPS and P3-CPS, which expressed much higher levels of CPS-YFP protein, *ent*-kaurene levels were approximately 55- and 200-fold higher than WT (Figure 2.6B). Protein expression levels in the three ER:CPS-YFP transgenic lines were similar to that of P2-CPS and P3-CPS, but ER:CPS-YFP lines accumulated dramatically lower levels of *ent*-kaurene (Figure 2.6B). The level of *ent*-kaurene in E1-CPS was intermediate between WT and *ga1-6* while in E2-CPS and E3-CPS, *ent*-kaurene levels were only slightly higher than WT. Quantification of the downstream GAs, GA_{12} and GA_{24} in untreated whole rosettes mimicked the trends seen for *ent*-kaurene in PAC-treated plants (Figure 2.6, C and D). The levels of both GAs in P2-CPS and P3-CPS were approximately 5-20 fold higher than WT, levels in P1-CPS and ER1-CPS were between WT and *ga1-6* while E2-CPS and E3-CPS were similar to WT (Figure 2.6, C and D).

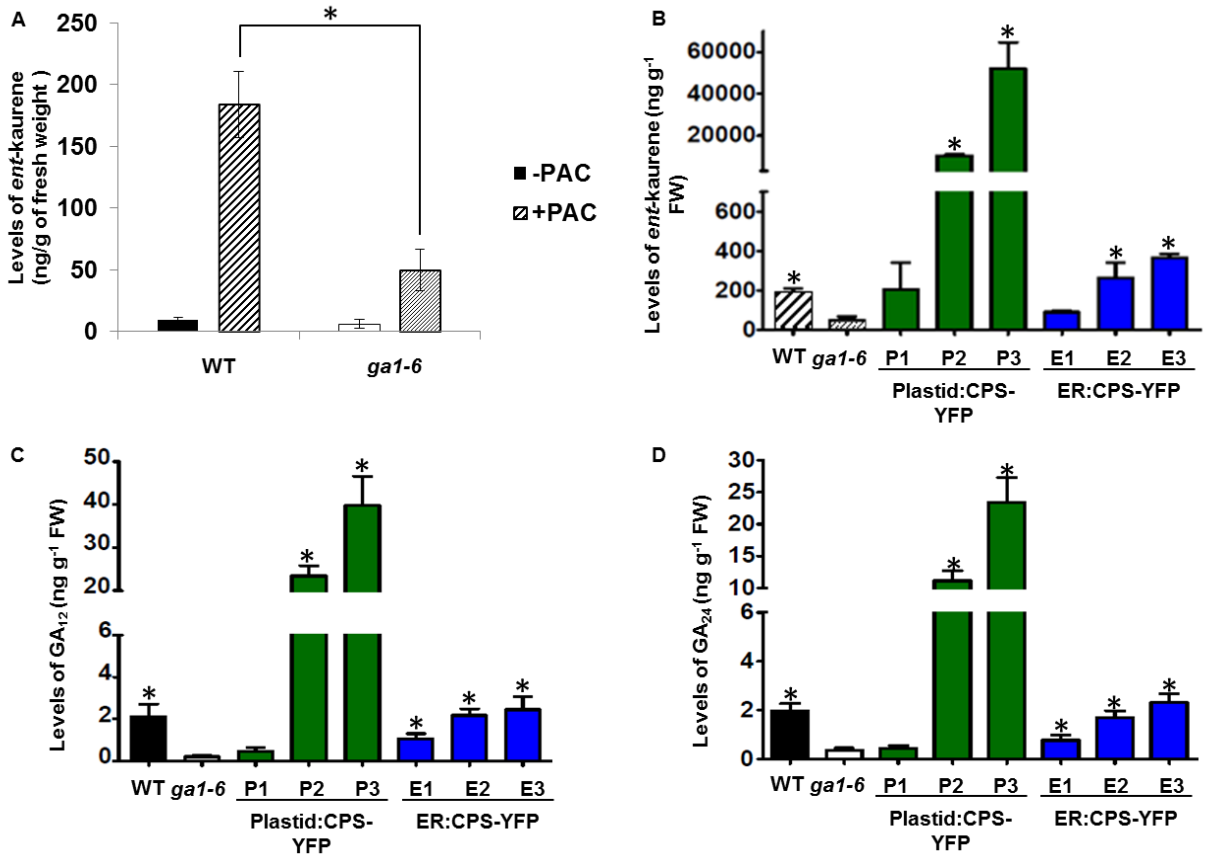


Figure 2.6: Gibberellin metabolite levels in transgenic lines.

A) Levels of *ent*-kaurene in WT and the *ga1-6* mutant. In the absence of paclobutrazol (PAC), the levels of *ent*-kaurene in four-week-old rosettes from WT and the *ga1-6* mutant were near or at detection limits and not significantly different. In the presence of PAC, the *ent*-kaurene levels in WT were five-fold higher than the *ga1-6* mutant.

B) Levels of *ent*-kaurene in the presence of paclobutrazol, a KO inhibitor in the CPS-YFP transgenic lines. Treated WT and the *ga1-6* mutant are included for reference.

C) Levels of the GA intermediate, GA_{12}, in WT, the *ga1-6* mutant, and CPS-YFP transgenic lines.

D) Levels of a non-C13 hydroxylated intermediate GA, GA_{24}, in WT, the *ga1-6* mutant, and CPS-YFP transgenic lines.

In all panels, average values are shown with standard deviation (n=3-4), * indicates significant differences from the *ga1-6* mutant at $p<0.05$.

Overexpression of GGDP synthase in the ER did not increase *ent*-kaurene production

The surprising contrast in *ent*-kaurene levels in PAC-treated plastid:CPS-YFP lines versus ER:CPS-YFP lines despite containing similar CPS levels (for example: P3-CPS and E3-CPS), suggested that the CPS substrate, GGDP, might be limiting synthesis of *ent*-kaurene in the ER lumen. In Arabidopsis, GGDP synthase 11 (GGPDS11) is the main GGDPS supplying GGDP for plastid-localized terpenoid pathways and has been well characterized (Beck *et al.*, 2013; Ruiz-Sola, Coman, *et al.*, 2016). GGDPS11 produces a full-length plastid-targeted enzyme and a truncated cytosolic enzyme, both of which are active (Ruiz-Sola, Barja, *et al.*, 2016). We retargeted the cytosolic version to the ER by adding an ER signal peptide to the N-terminus and the cyan fluorescent protein (CFP) either the N-terminus (after ER signal peptide) or C-terminus, followed in both cases by an HDEL retention signal.

To test for activity, all GGDPS constructs were expressed in a carotenoid-producing *E. coli* reporter background that lacks GGDPS activity and can only produce the reporter compound, lycopene, when active GGDPS is present (Misawa *et al.*, 1990). All versions of ER:GGDPS11 produced lycopene with most CFP fusions producing approximately 25% of the level compared to untagged GGDPS11 (Figure 2.7). C-terminus tagged ER:GGDPS11 was introduced into the E2:CPS and E3:CPS lines and resistant T1s which overexpressed both CPS-YFP and GGDPS11-CFP were selected. The levels of *ent*-kaurene in the PAC inhibitor system were then assayed in the T2 generation. GC-MS analysis showed varying levels of *ent*-kaurene ranging from five-fold lower to 1.2-fold higher than the corresponding E2-CPS or E3-CPS background (Figure 2.8). We then conducted immunoblotting analysis on these plant tissues and observed that in the low *ent*-kaurene producing lines (E2-CPS::E1-GGDPS and E3-CPS::E1-GGDPS), both CPS and GGDPS were cosuppressed, leading to little to no CPS and GGDPS11

protein expression (Figure 2.9). Even though GGDPS is overexpressed in the medium and high *ent*-kaurene producing lines, the levels of *ent*-kaurene produced were still dependent on CPS-YFP protein expression. For example, line E3-CPS::E2-GGDPS showed slightly lower levels of CPS-YFP protein expression than the E3-CPS parent line, and this resulted in 50% less *ent*-kaurene accumulated compared to its corresponding background, E3-CPS, despite GGDPS11 being highly expressed in this line. Taking all these data into consideration, this indicates that GGDP synthase activity is not limiting for production of *ent*-CDP in the ER but rather it is likely that the availability or transport of upstream precursors, isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP) is limiting for GGDP synthesis in the ER lumen.

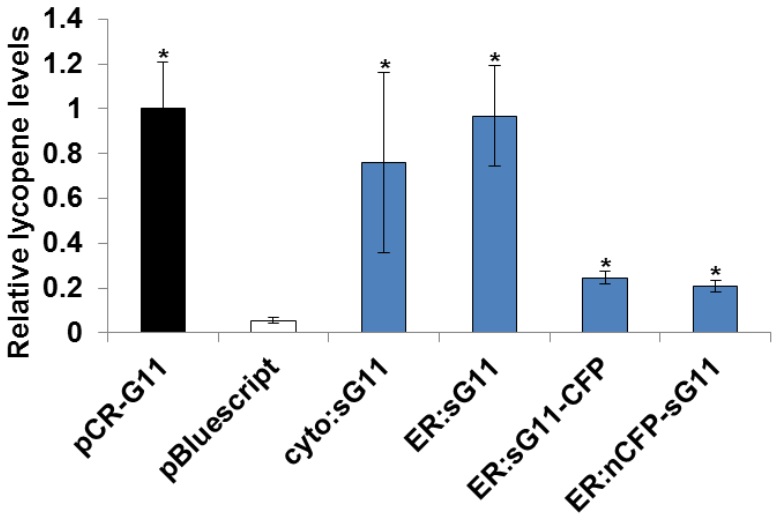


Figure 2.7: Genetic complementation in *E. coli.* Lycopene content in *E. coli* cells co-transformed with the pACCRT-BI plasmid carrying functional *crtB* (phytoene synthase) and *crtI* (phytoene desaturase) and pBluescript-GGDPS construct with no, N- or C-terminus tagged CFP or pBluescript (negative control). Lycopene levels were measured at 472 nm and normalized against bacterial growth. Relative lycopene levels are normalized to the levels in pCR-G11 (positive control). Values shown are the means and standard deviation of ten independent transformations. * indicates levels are statistically significant compared to pBluescript control at $p<0.05$.

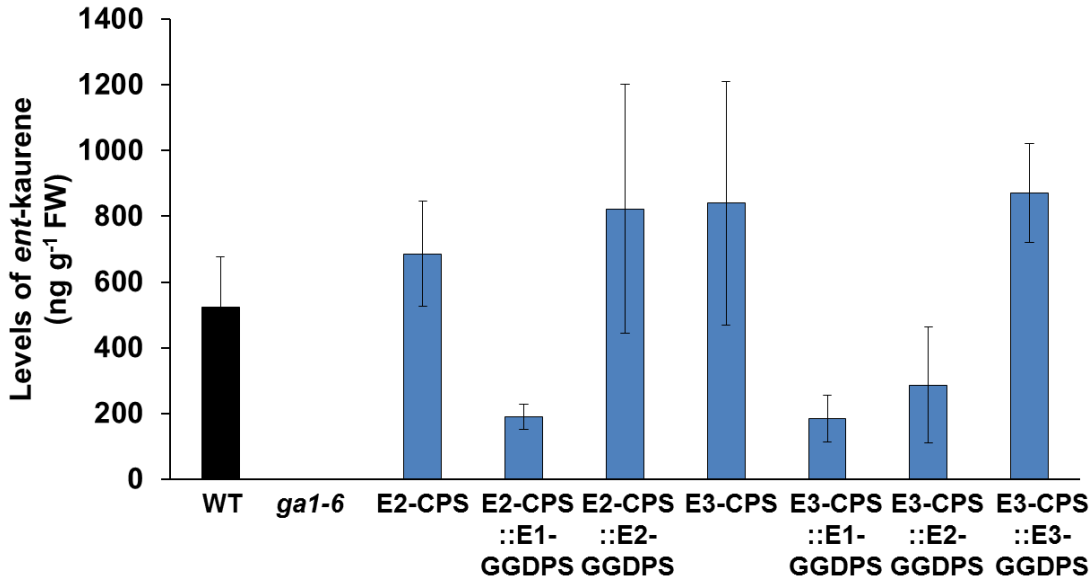


Figure 2.8: Levels of *ent*-kaurene in ER:CPS::ER:GGDPS double transgenic lines. Four-week-old segregating T2 transgenic plants overexpressing both CPS and GGDPS in the ER were treated with PAC and levels of ent-kaurene were quantified by GC-MS.

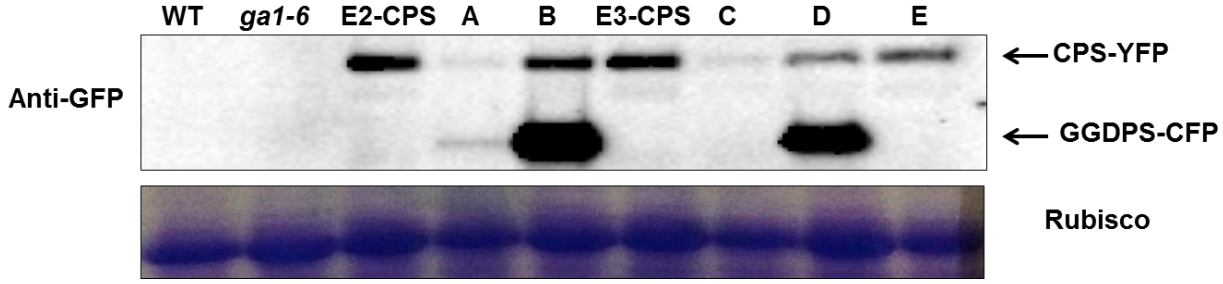


Figure 2.9: Protein levels in T2 ER:CPS::ER:GGDPS double transgenic lines. Twenty μg of total protein was loaded on 8% SDS-PAGE gel. Expression of both CPS-YFP and GGDPS-CFP were probed using antibodies against GFP (cross-reacts with YFP and CFP). Coomassie staining is shown as loading control. A: E2-CPS::E1-GGDPS; B: E2-CPS::E2-GGDPS; C: E3-CPS::E1-GGDPS; D: E3-CPS::E2-GGDPS; E: E3-CPS::E3-GGDPS.

Discussion:

The goal of this work was to probe the *in vivo* accessibility of the OEM-localized KO enzyme to the diterpene olefin *ent*-kaurene when it is produced in the ER lumen. This research, as a whole, aims to better understand how the synthesis of non-polar compounds like diterpene biosynthetic intermediates, for which no plastid envelope transporters have been identified, are transported between the plastid and ER in their normal synthesis. We selected the well-studied GA biosynthetic pathway as a model diterpene pathway for addressing these questions. All of the enzymes in the GA biosynthetic pathway have been characterized (Hedden, 2016), but how the transport of pathway intermediates such as *ent*-kaurene from the IEM to OEM, and *ent*-kaurenoic acid from the OEM to ER occurs are still unknown. Transporters of highly oxidized downstream GA metabolites such as GA_1, GA_3, and GA_4 (Kanno *et al.*, 2012; Saito *et al.*, 2015; Tal *et al.*, 2016; Kanno *et al.*, 2016) have been identified, but these transporters are plasma membrane-localized GA importers, and are likely redundant in function and transporting downstream GA metabolites as well as other plant hormones such as abscisic acid and jasmonate. Through previous transorganellar complementation work (Mehrshahi *et al.*, 2013), we hypothesized that non-polar compound(s) may be exchanged between the plastids and ER through hemifused membrane at membrane contact sites, a model that is consistent with the absence of identified transporters for the process at the plastid envelope. Therefore, we sought to test the accessibility and directionality of *ent*-kaurene at the plastid:ER contact sites by retargeting both CPS and KS into the ER lumen to generate an ER-localized pool of *ent*-kaurene and assess whether KO can access this substrate.

First, we independently retargeted CPS and KS protein into the ER lumen in mutants disrupted in these activities to serve as negative experimental controls. We reasoned that because

the product of ER-localized CPS and substrate for ER-localized KS is *ent*-CDP, a polar compound for which no transporters have been reported, therefore these transgenic lines would not complement their respective mutants (Figure 2.1). While this hypothesis was confirmed for ER:KS-mCherry, which did not complement the *ga2-1* mutation from the ER, retargeting of CPS into the ER complemented the *ga1-6* dwarf mutant phenotype. Two independent techniques were used to confirm the correct localization of ER:CPS-YFP to the ER lumen, indicating that the complementation observed is not due to protein mistargeting. This indicates that endogenous KS in the plastid must be able to access *ent*-CDP in the ER but that ER localized KS cannot access *ent*-CDP in the plastid. Complementation by ER-targeted CPS is also dependent on plastid-localized KS, as the resulting cross of the *ga1-6* E2-CPS x *ga2-1* suppressed the WT phenotype of E2-CPS (Figure 2.2C). This indicated that endogenous plastid-localized KS can access *ent*-CDP produced in the ER.

The physiological complementation of the *ga1-6* mutant by ER-targeted CPS-YFP (Figure 2.2) suggested that flux through the GA pathway was restored. This was indeed the case, as metabolite analysis of GA pathway intermediates in ER:CPS-YFP transgenic lines showed levels of *ent*-kaurene, $GA_{12,}$ and GA_{24} similar to WT (Figure 2.6, B, C, and D). There are at least two possible explanations for this result. First, *ent*-CDP produced in the ER lumen is transported from the ER into the plastid where endogenous KS can access it. This unidirectional transport of *ent*-CDP would require the presence of at least two transporters, one on the ER membrane and one on the IEM, as the OEM has transporters with broad substrate specificities. For example, outer envelope protein 24 (OEP24) allows the passage of a broad range of charged compounds including triose phosphates, dicarboxylic acids, positively or negatively charged amino acids, ATP and Pi (Pohlmeyer *et al.*, 1998). On the IEM, one candidate of such transporter is the still

unknown plastid-localized IDP (C_5) transporter that has been biochemically identified (Bick and Lange, 2003; Flügge and Gao, 2005). Bick and Lange (2003) proposed that the IDP transport is mediated by a Ca^{2+}-gated IDP/proton symporter while Flügge and Gao (2005) proposed that IDP is transported by a uniport system, with the requirement that phosphorylated compounds are present on the opposite side of the membrane. However, in Bick and Lange (2003), they found that GGDP, which is a linear, structural isomer of *ent*-CDP, was not preferentially transported across proteoliposomes reconstituted with inner envelope membrane proteins from spinach chloroplasts, whereas lower molecular weight prenyl diphosphate intermediates such as IDP, geranyl diphosphate (C_{10}) and farnesyl diphosphate (FDP, C_{15}) showed much higher rates of transport (20-, 20-, and 10-fold higher compared to GGDP). Nevertheless, in tobacco BY-2 cells, GGDP for protein prenylation in the cytosol was shown to be derived from the plastid, suggesting that GGDP can also be transported out of the plastid (Gerber *et al.*, 2009). The proposed "IDP transporter" described in Bick and Lange (2003) and Flügge and Gao (2005) may represent the same transporter, but this is difficult to test as neither genes nor proteins have been characterized to mediate this function. If *ent*-CDP is indeed transported by a putative "IDP transporter", it is likely a low-affinity transport activity, as *ent*-CDP is not known to be transported in or out of the plastid. Furthermore, no prenyl diphosphate transporter has yet been described for the ER membrane.

An alternate hypothesis to explain the access of ER-synthesized *ent*-CDP by endogenous plastid-localized KS is that *ent*-CDP is dephosphorylated in the ER lumen to its corresponding alcohol, *ent*-copalol. *Ent*-copalol is non-polar (logP 6.84) and would readily partition into the ER membrane. The activity of GGDP diphosphatases (GGDPase) in crude microsomes of rice seedlings have been biochemically characterized and occurred at a basal level that could be

further stimulated by UV-C irradiation (Nah *et al.*, 2001). Presumably, because GGDP (linear) and *ent*-CDP (bicyclic) are 20-carbon compounds, GGDPase activity could dephosphorylate *ent*-CDP to *ent*-copalol. Dephosphorylated *ent*-copalol could then be transported back into the plastid through plastid:ER contact sites, where in order for it to be a substrate for KS, it then needs to be rephosphorylated. In Arabidopsis, phytol (a linear, C_{20} alcohol) has been shown to be sequentially phosphorylated by two different plastid-localized prenyl alcohol kinases (phytol kinase and phytol-phosphate kinase, encoded by *VTE5* and *VTE6* respectively) (Valentin *et al.*, 2006; Vom Dorp *et al.*, 2015). It is an open question whether VTE5 and VTE6 have low levels of activity against geranylgeraniol (the corresponding alcohol of GGDP) or *ent*-copalol. Another potential candidate for phosphorylation of *ent*-copalol is a farnesol kinase protein (Fitzpatrick *et al.*, 2011), which is predicted to be plastid-localized. Farnesol (a linear, C_{15} alcohol) is the preferred substrate for farnesol kinase, with geraniol (a linear, C_{10} alcohol) and geranylgeraniol (a linear, C_{20} alcohol) being phosphorylated at much lower levels (Fitzpatrick *et al.*, 2011). Geranylgeraniol and geranylgeranyl monophosphate have also been shown to be phosphorylated and incorporated into chlorophyllide in the presence of exogenous ATP using broken etioplast membrane from oat seedlings (Rüdiger *et al.*, 1980). Therefore, it is conceivable that *ent*-copalol could be rephosphorylated to *ent*-CDP in the plastid, and made available as a substrate for KS. We however could not detect the presence of *ent*-copalol in either the E2-CPS or P3-CPS plant extracts (data not shown). We estimated that the detection limits of our analytical method for a diterpene alcohol at ~1 ng μL^{-1}. Since the extraction method also uses hexane as an extraction solvent, any potential *ent*-copalol conjugated to any polar compounds such as glucose or glutathione will not be extracted and detected in this analytical method. It is also conceivable that

in the P3-CPS plant extract, *ent*-CDP is efficiently converted to *ent*-kaurene, which precludes detection of *ent*-copalol.

While dephosphorylation/ rephosphorylation of *ent*-CDP cannot be eliminated, the most likely explanation for complementation of the *ga1-6* mutation by ER-CPS-YFP is that the *ent*-CDP produced is the intermediate transported. The exchange of prenyl diphosphate intermediate of various chain lengths between different subcellular compartments has been shown (Bick and Lange, 2003; Flügge and Gao, 2005; Dong *et al.*, 2015). In addition to biochemical characterization of the putative "IDP transporter" as discussed above (Bick and Lange, 2003; Flügge and Gao, 2005), a different strategy of expressing prenyl diphosphate synthase and terpene synthase in non-native subcellular compartments have also demonstrated that prenyl diphosphate intermediates, especially geranyl diphosphate (C_{10}) can also be exchanged (Gutensohn *et al.*, 2013; Dong *et al.*, 2015). Gutensohn *et al.* (2013) demonstrated increased levels of novel monoterpene production in tomato fruit cytosol by co-expressing plastid-localized snapdragon geranyl diphosphate synthase (GDPS) and cytosol-localized α-zingiberene synthase, which possesses both mono- and sesquiterpene synthase activity. The authors demonstrated that α-zingiberene synthase is localized in the cytosol, and hypothesized that the observed production of novel monoterpenes is due to channeling of plastidic GDP to the cytosol (Gutensohn et al., 2013). Dong *et al*. (2015) coexpressed different combinations of GDP synthase and geraniol synthase in the plastid, cytosol, or mitochondria and used the resulting production of geraniol or geraniol-derived products as reporters to show that there was a difference in the levels of GDP exchanged between these different subcellular compartments. These lines of evidence point towards a yet to be characterized transport mechanism for

exchanging prenyl diphosphate intermediates between plastids and other subcellular compartments.

Overexpression of the CPS protein in the plastid, its native organelle, led to massive increases in the levels of GA pathway intermediates such as *ent*-kaurene, $GA_{12,}$ and GA_{24}, which is consistent with the findings by Fleet *et al.* (2003) which showed an increase of up to 184-fold, 13-fold, and 4-fold higher than WT levels of *ent*-kaurene, $GA_{12,}$ and GA_{24}, respectively when CPS was overexpressed in the plastid. However, in this work, similar levels of CPS protein overexpression in the ER only led to metabolite levels similar to WT (Figure 2.6, B, C, and D). One explanation for this difference could be that insufficient levels of GGDP precursor are available in the ER lumen and that it is the availability of substrate, rather than total CPS activity, which limits the level of complementation. In Arabidopsis, GGDPS3 and 4 are targeted to the ER but their expression is confined to specific organs such as siliques, flowers and roots and they are only weakly expressed in leaves, the focus tissue of this work (Beck *et al.*, 2013). Overexpression of both GGDPS and CPS in the ER, however, did not lead to any significant increase in the levels of *ent*-kaurene (Figure 2.9) suggesting the limited production of diterpenes in the ER is due to limited availability and/or transport into the ER lumen of IDP and DMADP from the cytosolic MVA pathway.

The availability of IDP and DMADP, along with other MVA-derived terpenoids, is controlled by 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase, which is subject to tight regulation at the transcriptional, translational and post-translational levels (Hemmerlin, 2013). The multilevel regulation of HMG-CoA reductase might be circumvented by expressing a truncated, unregulated form of the protein, which has been shown to increase the levels of triterpenes in the cytosol (C30, sterols) (Harker *et al.*, 2003). This strategy was used by Wu *et al*.

(2008) in determining the maximum capacity of the cytosol to produce patchoulol (a sesquiterpene). They stably expressed the combination of farnesyl diphosphate synthase, patchoulol synthase, and the truncated form of HMG-CoA reductase in the cytosol, and could achieve an accumulation of 10-12 μg g^{-1} FW of patchoulol, which they concluded may represent the upper limits for producing sesquiterpenes in the cytosol of tobacco (Wu *et al.*, 2006). This level of patchoulol produced in the cytosol is substantial but about 5-fold lower than the levels of *ent*-kaurene accumulated in the plastid:CPS-YFP lines (maximum of 52 μg g^{-1} FW) and about 25-fold higher than the ER:CPS-YFP lines (maximum of 400 ng g^{-1} FW). We hypothesize that this strategy of overexpressing the truncated form of HMG-CoA reductase, along with GGDPS and CPS in the ER could potentially increase the levels of *ent*-kaurene if the issue was the limited availability of IDP and DMADP in the ER. However, if the transport of IDP and DMADP into the ER lumen is limiting, then the levels of *ent*-kaurene will stay the same, while the levels of cytosolic-derived terpenoids such as sterols and sesquiterpenes will increase.

Another possibility for the lower levels of GA metabolites in ER:CPS-YFP transgenic lines could be attributed to limited activity of the CPS protein in the ER. Characterization of recombinant CPS protein in an *in vitro* system showed that the optimal pH for CPS enzymatic activity is between pH 7.5 to 8.0 (Prisic and Peters, 2007), which is close to the pH of the stroma during illumination (Höhner *et al.*, 2016). The pH of the ER lumen is slightly lower at pH 7.1 (Shen *et al.*, 2013). In addition to GGDP levels, CPS activity is also influenced by Mg^{2+} concentration (Prisic and Peters, 2007). At high GGDP and Mg^{2+} concentrations (>1 μM and 1 mM, respectively), the k_{cat} of CPS decreased two-fold compared to the k_{cat} at an optimum Mg^{2+} concentration of 0.1 mM. During the transition of proplastids to chloroplasts, the levels of Mg^{2+} in this organelle increase from submillimolar to millimolar concentrations, when GGDP levels

are also proposed to increase with the strong expression of plastidic GGDPS11 (Fleet *et al.*, 2003; Beck *et al.*, 2013). It is unlikely that the ER luminal pH would substantially impact ER:CPS activity, but it is impossible to assess whether the other two components, GGDP and Mg^{2+} for their effects on activity, as there is no reliable estimate on their concentrations in the ER lumen.

Our work has also shown that in contrast to the ability of plastid-localized KS to access *ent*-CDP produced in the ER by ER:CPS-YFP. The converse, that ER-targeted KS can access plastid-localized *ent*-CDP substrate produced by the endogenous, plastid-localized CPS, is not true as evident by the lack of physiological complementation of the ER:KS-mCherry transgenic lines (Figure 2.2 B). This clearly indicates that the transport of *ent*-CDP is unidirectional from the ER back into the plastid. We also concluded that OEM-localized KO can access its non-polar substrate, *ent*-kaurene synthesized in the ER by the combined activities of ER-targeted CPS and KS in the *ga1ga2* double mutant by the recovery of the WT phenotype. As endogenous CPS acts as the "gatekeeper" in the synthesis of GA (Silverstone *et al.*, 1997; Fleet *et al.*, 2003), we hypothesized that the levels of GA metabolites in the ER-targeted CPS and KS in the *ga1ga2* double mutant background will be similar to the levels of ER:CPS-YFP (Figure 2.6, B, C, and D). The levels of early GA metabolites (*ent*-kaurene, $GA_{12,}$ and GA_{24}) have been shown increased synergistically when both CPS and KS were overexpressed in the plastids, but levels were similar to WT when only KS was overexpressed (Fleet *et al.*, 2003). If this is indeed true, this work shows that WT levels of *ent*-kaurene synthesized in the ER could allow for access of *ent*-kaurene by OEM-localized KO, and expand the repertoire of non-polar compound access across the plastid:ER contact sites to include the diterpene olefin, *ent*-kaurene.

In the endogenous situation with plastid-localized CPS and KS, the *ent*-kaurene produced in the stroma would likely partition into the IEM due to its hydrophobicity. Subsequent transport of *ent*-kaurene from the IEM to OEM could be further mediated by the contact sites between the IEM and OEM, in which large protein complexes that traverse the IEM and OEM such as the protein import machinery (TIC-TOC) and TGD complex could be localized (LaBrant *et al.*, 2018). In addition, *in vitro* studies have also shown that the N-terminus portion of digalactosyldiacylglycerol synthase 1 (DGD1) which is an integral OEM protein, could induce vesicle fusion in the presence of phosphatidic acid (Kelly *et al.*, 2016). The vesicle fusion between the IEM and OEM could transport *ent*-kaurene to the OEM, where KO could convert it to *ent*-kaurenoic acid. On the other hand, in the retargeted CPS and KS, where *ent*-kaurene is synthesized in the ER, plastid:ER contact sites likely mediate the transport of *ent*-kaurene from the ER membrane to the OEM for conversion to *ent*-kaurenoic acid.

Subsequent reaction steps in the GA pathway, post-formation of *ent*-kaurenoic acid, involve ER-localized cytochrome P450s. The transport of *ent*-kaurenoic acid to the ER could also be mediated by plastid:ER contact sites, and channeled directly to kaurenoic acid oxidases (KAO1 and KAO2) localized on the ER membrane (Helliwell *et al.*, 2001). GA_{12} and GA_{53}, which is a C13-hydroxylated GA intermediate, are further converted to bioactive GAs in the cytosol by soluble 2-oxoglutarate dependent dioxygenases. Recently, elucidation of the biosynthesis of crocin (a class of apocarotenoid glycosides) showed the consecutive enzymes catalyzing the first, second and third step are localized in the plastid, ER, and cytosol, respectively (Demurtas *et al.*, 2018). Demurtas *et al.* (2018) transiently expressed the pathway in *Nicotiana benthamiana* and proposed that the intermediate, crocetin dialdehyde is

dehydrogenated at the plastid:ER contact sites and the ER acts as a “transit center” in the synthesis of crocins before they are glycosylated and transported into the vacuole.

In conclusion, this work has shown that *ent*-kaurene could be successfully produced in the ER lumen and that it is accessible by OEM-localized KO enzyme, an orientation of movement that does not occur in the native pathway. In addition, we have also shown the surprising unidirectional transport of *ent*-CDP from the ER back into the plastid. These two different transport activities would most likely occur at the plastid:ER contact sites and would be mediated by a different transporter or transport mechanism.

Materials and Methods:

Vector construction

A construct encoding plastid-targeted CPS-YFP was generated by subcloning the full length CPS cDNA (Sun and Kamiya, 1994) into a pUC57 entry vector between the cauliflower mosaic virus (CaMV) 35S promoter and a C-terminus yellow fluorescent protein (YFP) tag. The construct encoding ER-targeted CPS-YFP was generated by truncating 189-bp (63 amino acids) of CPS CDS from the N-terminus, replaced it with the coding sequence for an ER signal peptide (Matsuoka *et al.*, 1995) and then fusing to the C-terminus a YFP tag and HDEL retention signal. Full-length coding sequence for KS was codon-optimized and synthesized by GenScript (Piscataway, New Jersey). For ER-targeting of KS, 99-bp (33 amino acids) of the KS CDS was truncated from the N-terminus, replaced it with the coding sequence for an ER signal peptide (Matsuoka *et al.*, 1995) and then fusing to a C-terminus a mCherry-FLAG tag and HDEL retention signal. The CPS constructs were then subcloned into pMLBart and the KS constructs were subcloned into pART27 and transformed into the *ga1-6* and *ga2-1* mutant background, respectively. All Arabidopsis transformations were done by floral dipping (Clough and Bent, 1998). CPS T_1 resistant seedlings were selected on soil by spraying with 300 μM glufosinate-ammonium (Ignite®, Bayer), while KS T_1 resistant seedlings were selected on Murashige and Skoog (MS) plates containing 50 μg/μL kanamycin. Single insert homozygous transgenic lines were established and used for subsequent experiments.

Plant transformation and growth condition

ga1-6 and *ga2-1* seeds were obtained from the Arabidopsis Biological Resource Center, Ohio State University. Mutant seeds were sown on 1x MS plates supplemented with 10 μM GA_3 (PhytoTechnology Laboratories) and stratified at 4°C for two days in the dark to aid germination.

Ten-day-old seedlings were then transplanted to soil and grown under standard conditions (120-200 μmole m^{-2} s^{-1}, 16-hour day at 22°C/ 8-hour dark at 18°C). For metabolite analysis, 4-week-old soil grown plants were utilized, while for subcellular fractionation, 3-week-old plate grown seedlings were utilized.

Genotyping information for the GA biosynthetic mutants

Derived cleaved amplified polymorphic sequences (dCAPS) markers were used in order to genotype the respective mutant backgrounds. Genotyping for *ga1-6* mutant was done using PCR primers (Forward 5'-CGTGAAATGGATCGCCGAGAACCAACTTTCCGATGGTC-3'; Reverse 5' CACTAAAAAAGTTTGGATTTTCTGCAAAAAGTTAAGATTTTC-3') and restriction digestion with the AvaII enzyme (New England Biolabs). Genotyping of *ga2-1* mutant was done using PCR primers (Forward 5'-CTCGTGAGCACTATGGGTCGTCTTCTAAATGGGATC-3'; Reverse 5'-CAATGAAACCGCATTCAGC-3') and restriction digestion with the BamHI enzyme (New England Biolabs).

Transient expression in tobacco

Three-week-old *Nicotiana tabacum* plants were grown under standard conditions for transient expression. Single colonies of *A. tumefaciens* harboring either Pl:KS-mCherry or ER:KS-mCherry were grown individually in Luria-Bertani media supplemented with appropriate antibiotics (rifampicin, 25 μg/mL; gentamycin, 10 μg/mL; spectinomycin, 100 μg/mL) for 16 hours at 28°C. Bacterial optical density was adjusted to OD_{600} 0.05 and transiently expressed in *N. tabacum* using the methods described in Sparkes *et al*., 2006, with the MES stock solution pH adjusted to 5.8. *N. tabacum* plants were returned to growth chamber post *A. tumefaciens* infiltration under standard conditions.

Confocal microscopy

Confocal imaging was performed on either stable Arabidopsis leaves expressing a given transgene or tobacco leaves infiltrated with *Agrobacterium* using an inverted scanning confocal microscope Zeiss LSM510 META. CFP was excited using 438 nm and emission collected at 465-510 nm; YFP was excited using 514 nm and emission collected at 520-555 nm, mCherry was excited at 543 nm and emission collected at 560-615 nm while chlorophylls were excited either at 543 or 594 nm and emission collected at 650-736 nm.

Cell fractionation

Four-week-old seedlings (~10g fresh weight) grown on plates were harvested and homogenized in approximately 50 mL 1x grinding buffer (0.3 M sorbitol, 5 mM $MgCl_2$, 5 mM EGTA, 20 mM HEPES-KOH, pH7.3, 4 mM EDTA, 0.1 % BSA) in several 5-s bursts and filtered through two layers of Miracloth (Calbiochem). Homogenization and filtration were repeated again with 30 mL 1x grinding buffer. Filtrates were pooled and an aliquot was removed to serve as the total protein fraction. Pooled filtrates were centrifuged at 4000 RPM for 5 minutes using an HB4 rotor.

For chloroplast isolation, the resulting pellet (crude chloroplast fraction) was resuspended with 2 mL 1x grinding buffer (50 mM HEPES-KOH, pH 8, 0.33 M sorbitol) and carefully loaded on top of a 20 mL 30% (v/v) Percoll cushion in 1x import buffer. The Percoll cushion was then centrifuged at 3000 RPM for 10 minutes using a swing out rotor in Eppendorf 5810R table top centrifuge. The Percoll was removed and intact chloroplasts were recovered as a pellet at the bottom of the tube. The intact chloroplast pellets were washed with 10 mL 1x import buffer and centrifuged at 3000 RPM for 5 minutes using a swing out rotor. Washed chloroplast pellets were resuspended in 20 mL 1x grinding buffer and aliquots of 20, 40 and 60 μL were used to quantify

chlorophyll. All chloroplast suspensions were resuspended to a concentration of 1 mg chlorophyll/mL.

For crude microsome isolation, the supernatant from the initial HB4 centrifugation to harvest intact chloroplasts was centrifuged at 10000 RPM for 20 minutes using an HB6 rotor to remove most cell debris and thylakoid membranes. The resulting supernatant was then centrifuged at 32000 RPM for 1 hour using a 70Ti rotor to pellet crude microsomes. The resulting pellet (crude microsome pellet) was resuspended in 1 mL 14% (w/w) sucrose in 10 mM Tris-HCl, pH 7.5 with or without 5 mM $MgCl_2$ and homogenized with Dounce homogenizer. The homogenized supernatant was centrifuged at 13000 RPM for 1 minute using a microcentrifuge. 700 μL of the supernatant is then loaded onto a 7.6 mL 20-50% (w/w) sucrose gradient in 10 mM Tris-HCl, pH 7.5 with or without 5 mM $MgCl_2$ and centrifuged at 24000 RPM for 16.5 hours using an SW32.1Ti rotor. 500 μL gradient fractions were collected post centrifugation and subjected to immunoblotting analysis. All centrifugation steps were performed between 4-6°C.

Protease digestion of crude microsome

Crude microsomes isolated as described above and resuspended in 1 mL of 1x import buffer using a Dounce homogenizer and digested following the protocol described in Froehlich, 2011. Briefly, 250 μg of crude microsome protein was treated with 6.25 μg thermolysin in the presence or absence of 1% (v/v) of TritonX-100 for 30 minutes on ice. The reaction was quenched by addition of 60 mM EDTA in 1x import buffer and further incubated for 5 minutes on ice. For samples containing TritonX-100, proteins were precipitated by addition of 1 mL of cold acetone followed by incubation for 30 minutes on ice. All samples were recovered by centrifugation at

13000 RPM for 30 minutes at 4°C and resuspended in 40 μL of 2X Laemmli buffer, boiled at 95°C for 5 minutes and 10 μL were separated on 10% SDS-PAGE gel prior to immunoblotting.

Immunoblotting analysis

For immunoblotting analysis on fractions collected from the linear sucrose gradient, 20 μL of each fraction was loaded. For immunoblotting analysis on the total protein, chloroplast, and crude microsome fractions, protein concentrations were determined using RC DC™ protein assay (Bio-Rad, Hercules, CA). Ten μg of proteins dissolved in 6x Laemmli buffer were separated on a 10% SDS-PAGE gel and transferred onto polyvinylidene fluoride (PVDF) membranes (GE Healthcare, Chicago, IL). Membranes were blocked overnight with West-Ezier Super Blocking buffer (GenDEPOT, Katy, TX) at 4°C, probed with primary antibody for either 3 hours (for BiP) or 1 hour (GFP, TOC75, TOC34, TIC 110 and OE33) at room temperature. Membranes were subsequently washed and incubated for 1 hour at room temperature with horseradish peroxidase-conjugated goat anti-rabbit (Jackson ImmunoResearch Laboratories, West Grove, PA) diluted 1:20000. Membranes were washed and detection was done using the Clarity Western ECL Blotting substrates (Bio-Rad, Hercules, CA). Immunoblots were incubated with the following antibodies (sources and dilutions are indicated): GFP antibody (1:2500, Life Technologies); TOC75 (1:5000) and TIC100 (1:2500) antibodies were provided by John Froehlich, Michigan State University; BiP antibody (1:5000, Santa Cruz Biotechnology, Dallas, TX) and TOC34 (1:10000) and OE33 antibodies (1:5000) were from Agrisera (Sweden).

***ent*-Kaurene analysis**

~0.5 g of frozen powdered Arabidopsis whole rosette that had been treated with 120 μM paclobutrazol (PhytoTechnology Laboratories, Shawnee Mission, KS) for 4 days were extracted with 2 mL hexane containing 0.25 μg/ mL methyl heptadecanoate as an internal standard

(Sigma-Aldrich, St. Louis, MO) and partitioned against 2 mL methanol for 8 mins to remove more polar compounds using a paint shaker. To ensure thorough separation of the hexane and methanol phase, 1.5 mL H_2O was added and the mixture again shaken using a paint shaker for an additional 2 mins. Samples were then centrifuged at 4°C and 3750 RPM for 10 min to pellet plant debris and ensure phase separation. 1.5 mL of the hexane layer (top layer) was transferred to a disposable borosilicate 13 mm x 100 mm test tube and evaporated to completeness under nitrogen gas. The dried extract was then resuspended in 50 μL hexane. A standard calibration curve was constructed using an *ent*-kaurene standard (gift of Dr. Reuben Peters, Iowa State University) with the methyl heptadecanoate internal standard over a range of concentrations from 12.2 pg/μL to 200 ng/μL. Quantification was performed on an Agilent 5975 GC/ single quadrupole mass spectrometer with an Agilent J&W VF-5ms column (30 m x 0.25 mm x 0.25 μm) (Agilent, Santa Clara, CA). 1 μL of the sample was injected in splitless mode with an injector temperature of 250°C and a flow rate of 1.0 mL/min helium. Initial oven temperature was 40°C, held for 2 min, then raised by 40°C/min to 210°C, 4°C/min to 233°C, 50°C/min to 340°C and held at 340°C for 9.36 min. Ionization employed 70 eV electron ionization. QuanLynx (Waters Corporation, Milford, MA) was used for data analysis after exporting the original ChemStation data file into NetCDF format, then converting into Waters .raw file using DataBridge (Waters Corporation, Milford, MA). Quantification was based on integration of the extracted ion chromatogram for mass-to-charge ratio (m/z) 229 for *ent*-kaurene and 284 for methyl heptadecanoate.

GA analysis

~0.5 g of frozen powdered Arabidopsis whole rosette was extracted with 2 mL acetonitrile: isopropanol: H_2O, pH 8 (3:3:2, v/v/v) containing 2.5 ng/ mL deuterium-labeled GAs (internal

standard) for 10 minutes using a paint shaker. Samples were then centrifuged at 4°C and 3750 RPM for 10 min to pellet plant debris. 1.75 mL of supernatants were then loaded onto preconditioned (3 mL methanol, followed by 3 mL H_2O) 6cc, 150 mg sorbent Oasis WAX column (Waters Corporation, Milford, MA). Liquid was eluted dropwise from the column by gravity to ensure maximum interaction and binding of GAs to the column. The column was similarly washed with 3 mL acetonitrile: H_2O, pH 8 (9:1, v/v) and the GAs were eluted with 2 mL acetonitrile:H_2O:NH_4OH (7:2.5:0.5, v/v/v). The elution fraction was then evaporated to completeness in a speed vac concentrator and resuspended in 100 μL acetonitrile:H_2O (1:1, v/v). Standard calibration curves were constructed using gibberellin A_{12} (GA_{12}) and gibberellin A_{24} (GA_{24}) (OlChemIm, Czech Republic) with deuterium-labeled GA_{12} and GA_{24} (Dr. Susanne Hoffmann-Benning, Michigan State University) over a range of concentrations from 0 to 1 μM. Quantification was done on a Waters Quattro Premier XE UPLC/MS/MS (Waters Corporation, Milford, MA). 10 μL of the extract was then analyzed using a 5-min gradient method on an Ascentis® Express C18 HPLC column (5 cm x 2.1 mm, 2.7 μm) with mobile phase consisting of 0.1% formic acid in H_2O (solvent A) and acetonitrile:isopropanol (1:1, v/v) (solvent B). The 5-min gradient method was: 5% B at 0.00 to 0.50 min, linear gradient to 99% B from 0.50 min to 3.00 min, hold at 99% B from 3.00 min to 4.00 min, then return to 5% B from 4.01 min to 5.00 min. The flow rate was 0.3 mL/min and the column temperature was held at 50°C. The mass spectrometer was used in negative electrospray ionization mode and data were collected in MRM channels GA_{12} 331>269 [Cone voltage (CV): 40 V; collision energy (CE): 34 eV]; d_2-GA_{12} 333>271 (CV: 40 V; CE: 28 eV); GA_{24} 345>301 (CV: 40 V; CE: 22 eV); d_2-GA_{24} 347>303 (CV: 40 V, CE: 22 eV).

In vivo bacteria complementation

Constructs to express either soluble or ER-targeted GGDPS11 were generated as described in Table 1. pBluescript-sG11 vectors were constructed by restriction ligation into the SpeI and KpnI of pBluescript SK sites. The resulting construct consisted of either N- or C-terminus tagged or non-tagged versions of ER-targeted sGGDPS11 with or without the ER signal peptide. *E. coli* Top10 cells were cotransformed with pACCRT-BI plasmid (gift of Dr. Manuel Rodriguez-Concepcion, CRAG Barcelona, Spain) carrying genes encoding for phytoene synthase (CrtB) and phytoene desaturase (CrtI) and the pBluescript-sG11. Positive control, pCR-G11 (described in (Ruiz-Sola, Barja, *et al.*, 2016), was provided by Dr. Manuel Rodriguez-Concepcion (CRAG Barcelona, Spain) and the empty vector, pBluescript SK were used as controls in this experiment. Positive transformants were selected on Luria-Bertani plates with carbenicillin (100 $\mu g\ mL^{-1}$) and chloramphenicol (35 $\mu g\ mL^{-1}$). Ten individual colonies were inoculated in liquid LB overnight at 37°C. The following day, 1 mL of the overnight cultures were used to inoculate 10 mL of fresh liquid LB with antibiotics and 1 mM IPTG to induce lycopene production and grown at 20°C for 72 hours in the dark. 200 μL of the culture was diluted 5-fold and used to measure cell density at OD600. The cultures were then centrifuged at 4°C and 4000 RPM for 10 min. The bacterial pellet was then resuspended in a small volume of liquid LB and transferred to a 2 mL microtube. The culture was then centrifuged at room temperature and 13300 RPM for 2 min and remaining supernatant were discarded. The bacterial pellet was then extracted with 700 μL acetone and mixed at 55°C and 1400 RPM for 20 min. After mixing, the extracts were then centrifuged at room temperature and 13300 RPM for 5 min. Absorbance were then measured at 443, 472 and 502 nm and quantification was done using absorbance at 472 nm.

	Targeting	Tag	Construct	Plasmid backbone	Cloning method
In vivo GGDPS activity assay	Soluble	Non-tagged	Cyto:sG11	pBluescript SK	Restriction digestion at SpeI and KpnI
		N-terminus CFP	ER:nCFP-sG11		
	ER	C-terminus CFP	ER:sG11-cCFP		
		Non-tagged	ER:sG11		

Table 2: Constructs for testing GGDP synthase activity in *E.*coli.

REFERENCES

REFERENCES

Baginsky, S. and Gruissem, W. (2009) The Chloroplast Kinase Network: New Insights from Large-Scale Phosphoproteome Profiling. *Mol. Plant*, **2**, 1141–1153.

Beck, G., Coman, D., Herren, E., Ruiz-Sola, M.A., Rodríguez-Concepción, M., Gruissem, W. and Vranová, E. (2013) Characterization of the GGPP synthase gene family in Arabidopsis thaliana. *Plant Mol. Biol.*, **82**, 393–416.

Bick, J.A. and Lange, B.M. (2003) Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch. Biochem. Biophys.*, **415**, 146–154.

Brandizzi, F., Irons, S.L., Johansen, J., Kotzer, A. and Neumann, U. (2004) GFP is the way to glow: Bioimaging of the plant endomembrane system. *J. Microsc.*, **214**, 138–158.

Breuers, F.K.H., Bräutigam, A. and Weber, A.P.M. (2011) The Plastid Outer Envelope - A Highly Dynamic Interface between Plastid and Cytoplasm. *Front. Plant Sci.*, **2**, 97.

Chen, F., Tholl, D., Bohlmann, J. and Pichersky, E. (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.*, **66**, 212–229.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.*, **16**, 735–43.

Demurtas, O.C., Frusciante, S., Ferrante, P., et al. (2018) Candidate Enzymes for Saffron Crocin Biosynthesis Are Localized in Multiple Cellular Compartments. *Plant Physiol.*, **177**, 990–1006.

Dong, L., Jongedijk, E., Bouwmeester, H. and Krol, A. Van Der (2015) Monoterpene biosynthesis potential of plant subcellular compartments Monoterpene biosynthesis potential of plant subcellular compartments. , 679–690.

Dorp, K. Vom, Hölzl, G., Plohmann, C., Eisenhut, M., Abraham, M., Weber, A.P.M., Hanson, A.D. and Dörmann, P. (2015) Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. *Plant Cell*, **27**, 2846–59.

Facchinelli, F. and Weber, A.P.M. (2011) The Metabolite Transporters of the Plastid Envelope: An Update. *Front. Plant Sci.*, **2**, 1–18.

Ferro, M., Brugière, S., Salvi, D., et al. (2010) AT_CHLORO, a Comprehensive Chloroplast Proteome Database with Subplastidial Localization and Curated Information on Envelope Proteins. *Mol. Cell. Proteomics*, **9**, 1063–1084.

Fischer, K., Weber, A.P.M.. and Kunz, H.-H. (2016) The Transporters of Plastids-New Insghts into an Old Field. In H. Kirchhoff, ed. *Chloroplasts: Current Research and Future*

Trends. Norfolk: Caister Academic Press, pp. 209–240.

Fitzpatrick, A.H., Bhandari, J. and Crowell, D.N. (2011) Farnesol kinase is involved in farnesol metabolism, ABA signaling and flower development in Arabidopsis. *Plant J.*, **66**, 1078–88.

Fleet, C.M., Yamaguchi, S., Hanada, A., Kawaide, H., David, C.J., Kamiya, Y. and Sun, T. (2003) Overexpression of AtCPS and AtKS in Arabidopsis Confers Increased ent -Kaurene Production But No Increase in Bioactive Gibberellins 1. *Plant Physiol.*, **132**, 830–839.

Flügge, U. and Gao, W. (2005) Transport of Isoprenoid Intermediates Across Chloroplast Envelope Membranes. *Plant Biol.*, **7**, 91–97.

Flügge, U.I. (2000) Transport in and out of plastids: does the outer envelope membrane control the flow? *Trends Plant Sci.*, **5**, 135–7.

Froehlich, J. (2011) Studying Arabidopsis Envelope Protein Localization and Topology Using Thermolysin and Trypsin Proteases. In *Methods in Molecular Biology*. pp. 351–367.

Froehlich, J.E., Wilkerson, C.G., Ray, W.K., McAndrew, R.S., Osteryoung, K.W., Gage, D.A. and Phinney, B.S. (2003) Proteomic study of the Arabidopsis thaliana chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *J. Proteome Res.*, **2**, 413–25.

Gerber, E., Hemmerlin, A., Hartmann, M., et al. (2009) The plastidial 2-C-methyl-D-erythritol 4-phosphate pathway provides the isoprenyl moiety for protein geranylgeranylation in tobacco BY-2 cells. *Plant Cell*, **21**, 285–300.

Gutensohn, M., Orlova, I., Nguyen, T.T.H., Davidovich-Rikanati, R., Ferruzzi, M.G., Sitrit, Y., Lewinsohn, E., Pichersky, E. and Dudareva, N. (2013) Cytosolic monoterpene biosynthesis is supported by plastid-generated geranyl diphosphate substrate in transgenic tomato fruits. *Plant J.*, **75**, 351–363.

Harker, M., Holmberg, N., Clayton, J.C., Gibbard, C.L., Wallace, A.D., Rawlins, S., Hellyer, S.A., Lanot, A. and Safford, R. (2003) Enhancement of seed phytosterol levels by expression of an N-terminal truncated Hevea brasiliensis (rubber tree) 3-hydroxy-3-methylglutaryl-CoA reductase. *Plant Biotechnol. J.*, **1**, 113–21.

Hashimoto, A., Ettinger, W.F., Yamamoto, Y. and Theg, S.M. (1997) Assembly of Newly Imported Oxygen-Evolving Complex Subunits in Isolated Chloroplasts: Sites of Assembly and Mechanism of Binding. *Plant Cell*, **9**, 441–452.

Hedden, P. (2016) Gibberellin biosynthesis in higher plants. In *Annual Plant Reviews, Volume 49*. Chichester, UK: John Wiley & Sons, Ltd, pp. 37–72.

Helliwell, C.A., Sullivan, J.A., Mould, R.M., Gray, J.C., Peacock, W.J. and Dennis, E.S. (2001) A plastid envelope location of Arabidopsis ent-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway. *Plant J.*, **28**, 201–8.

Hemmerlin, A. (2013) Post-translational events and modifications regulating plant enzymes

involved in isoprenoid precursor biosynthesis. *Plant Sci.*, **203**–**204**, 41–54.

Höhner, R., Aboukila, A., Kunz, H.-H. and Venema, K. (2016) Proton Gradients and Proton-Dependent Transport Processes in the Chloroplast. *Front. Plant Sci.*, **7**, 1–7.

Jarvis, P. and López-Juez, E. (2013) Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.*, **14**, 787–802.

Kanno, Y., Hanada, A., Chiba, Y., Ichikawa, T., Nakazawa, M., Matsui, M., Koshiba, T., Kamiya, Y. and Seo, M. (2012) Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. *Proc. Natl. Acad. Sci.*, **109**, 9653–9658.

Kanno, Y., Oikawa, T., Chiba, Y., et al. (2016) AtSWEET13 and AtSWEET14 regulate gibberellin-mediated physiological processes. *Nat. Commun.*, **7**, 13245.

Keeling, C.I., Madilao, L.L., Zerbe, P., Dullat, H.K. and Bohlmann, J. (2011) The primary diterpene synthase products of Picea abies levopimaradiene/abietadiene synthase (PaLAS) are epimers of a thermally unstable diterpenol. *J. Biol. Chem.*, **286**, 21145–53.

Keeling, C.I., Weisshaar, S., Ralph, S.G., Jancsik, S., Hamberger, B., Dullat, H.K. and Bohlmann, J. (2011) Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (Picea spp.). *BMC Plant Biol.*, **11**, 43.

Kelly, A.A., Kalisch, B., Hölzl, G., Schulze, S., Thiele, J., Melzer, M., Roston, R.L., Benning, C. and Dörmann, P. (2016) Synthesis and transfer of galactolipids in the chloroplast envelope membranes of Arabidopsis thaliana. *Proc. Natl. Acad. Sci.*, **113**, 10714–10719.

Koornneef, M., Eden, J. Van, Hanhart, C.J. and Jongh, A.M.M. De (1983) Genetic fine-structure of the GA-1 locus in the higher plant Arabidopsis thaliana (L.) Heynh. *Genet. Res.*, **41**, 57.

Koornneef, M. and Veen, J.H. van der (1980) Induction and analysis of gibberellin sensitive mutants in Arabidopsis thaliana (L.) heynh. *Theor. Appl. Genet.*, **58**, 257–263.

LaBrant, E., Barnes, A.C. and Roston, R.L. (2018) Lipid transport required to make lipids of photosynthetic membranes. *Photosynth. Res.*, **0**, 0.

Lopez-Juez, E. and Pyke, K.A. (2005) Plastids unleashed: Their development and their integration in plant development. *Int. J. Dev. Biol.*, **49**, 557–577.

Lord, J. (1987) Isolation of endoplasmic reticulum: General principles, enzymatic markers, and endoplasmic reticulum-bound polysomes. In *Methods in Enzymology*. pp. 576–584.

Martin, D.M., Fäldt, J. and Bohlmann, J. (2004) Functional Characterization of Nine Norway SpruceTPS Genes and Evolution of GymnospermTerpene Synthases of the TPS-d Subfamily. *Plant Physiol.*, **135**, 1908–1927.

Matsuoka, K., Watanabe, N. and Nakamura, K. (1995) O-glycosylation of a precursor to a sweet potato vacuolar protein, sporamin, expressed in tobacco cells. *Plant J.*, **8**, 877–89.

Mehrshahi, P., Johnny, C. and DellaPenna, D. (2014) Redefining the metabolic continuity of chloroplasts and ER. *Trends Plant Sci.*, **19**, 501–7.

Mehrshahi, P., Stefano, G., Andaloro, J.M., Brandizzi, F., Froehlich, J.E. and DellaPenna, D. (2013) Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 12126–31.

Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K. and Harashima, K. (1990) Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli. *J. Bacteriol.*, **172**, 6704–12.

Nah, J., Song, S.J. and Back, K. (2001) Partial characterization of farnesyl and geranylgeranyl diphosphatases induced in rice seedlings by UV-C irradiation. *Plant Cell Physiol.*, **42**, 864–867.

Oliver, S.C., Venis, M.A., Freedman, R.B. and Napier, R.M. (1995) Regulation of synthesis and turnover of maize auxin-binding protein and observations on its passage to the plasma membrane: comparisons to maize immunoglobulin-binding protein cognate. *Planta*, **197**, 465–74.

Peters, R.J., Flory, J.E., Jetter, R., Ravn, M.M., Lee, H.J., Coates, R.M. and Croteau, R.B. (2000) Abietadiene synthase from grand fir (Abies grandis): Characterization and mechanism of action of the "pseudomature" recombinant enzyme. *Biochemistry*, **39**, 15592–15602.

Pohlmeyer, K., Soll, J., Grimm, R., Hill, K. and Wagner, R. (1998) A high-conductance solute channel in the chloroplastic outer envelope from Pea. *Plant Cell*, **10**, 1207–16.

Pottosin, I. and Dobrovinskaya, O. (2015) Ion Channels in Native Chloroplast Membranes: Challenges and Potential for Direct Patch-Clamp Studies. *Front. Physiol.*, **6**, 396.

Prisic, S. and Peters, R.J. (2007) Synergistic substrate inhibition of ent-copalyl diphosphate synthase: a potential feed-forward inhibition mechanism limiting gibberellin metabolism. *Plant Physiol.*, **144**, 445–54.

Pyke, K. (2007) Plastid biogenesis and differentiation. In R. Bock, ed. *Cell and Molecular Biology of Plastids*. Heidelberg: Springer Berlin Heidelberg, pp. 1–28.

Rüdiger, W., Benz, J. and Guthoff, C. (1980) Detection and partial characterization of activity of chlorophyll synthetase in etioplast membranes. *Eur. J. Biochem.*, **109**, 193–200.

Ruiz-Sola, M.Á., Barja, M.V., Manzano, D., Llorente, B., Schipper, B., Beekwilder, J. and Rodriguez-Concepcion, M. (2016) A Single Arabidopsis Gene Encodes Two Differentially Targeted Geranylgeranyl Diphosphate Synthase Isoforms. *Plant Physiol.*, **172**, 1393–1402.

Ruiz-Sola, M.Á., Coman, D., Beck, G., et al. (2016) Arabidopsis GERANYLGERANYL DIPHOSPHATE SYNTHASE 11 is a hub isozyme required for the production of most photosynthesis-related isoprenoids. *New Phytol.*, **209**, 252–264.

Saito, H., Oikawa, T., Hamamoto, S., et al. (2015) The jasmonate-responsive GTR1 transporter is required for gibberellin-mediated stamen development in Arabidopsis. *Nat. Commun.*, **6**, 6095.

Schaller, G.E. (2017) *Isolation of Plant Organelles and Structures* N. L. Taylor and A. H. Millar, eds., New York, NY: Springer New York.

Shen, J., Zeng, Y., Zhuang, X., Sun, L., Yao, X., Pimpl, P. and Jiang, L. (2013) Organelle pH in the arabidopsis endomembrane system. *Mol. Plant*, **6**, 1419–1437.

Silverstone, A.L., Chang, C., Krol, E. and Sun, T.P. (1997) Developmental regulation of the gibberellin biosynthetic gene GA1 in Arabidopsis thaliana. *Plant J.*, **12**, 9–19.

Sparkes, I.A., Runions, J., Kearns, A. and Hawes, C. (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.*, **1**, 2019–2025.

Sun, T., Goodman, H.M. and Ausubel, F.M. (1992) Cloning the Arabidopsis GA1 Locus by Genomic Subtraction. *Plant Cell*, **4**, 119–128.

Sun, T.P. and Kamiya, Y. (1994) The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell*, **6**, 1509–18.

Sveshnikova, N., Soll, J. and Schleiff, E. (2000) Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 4973–4978.

Tal, I., Zhang, Y., Jørgensen, M.E., et al. (2016) The Arabidopsis NPF3 protein is a GA transporter. *Nat. Commun.*, **7**, 11486.

Tranel, P.J., Froehlich, J., Goyal, A. and Keegstra, K. (1995) A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. *EMBO J.*, **14**, 2436–46.

Valentin, H.E., Lincoln, K., Moshiri, F., et al. (2006) The Arabidopsis vitamin E pathway gene5-1 mutant reveals a critical role for phytol kinase in seed tocopherol biosynthesis. *Plant Cell*, **18**, 212–24.

Vogel, B.S., Wildung, M.R., Vogel, G. and Croteau, R. (1996) Abietadiene Synthase from Grand Fir (Abies grandis). *J. Biol. Chem.*, **271**, 23262–23268.

Wang, Z., Xu, C. and Benning, C. (2012) TGD4 involved in endoplasmic reticulum-to-chloroplast lipid trafficking is a phosphatidic acid binding protein. *Plant J.*, **70**, 614–623.

Wu, S., Schalk, M., Clark, A., Miles, R.B., Coates, R. and Chappell, J. (2006) Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nat. Biotechnol.*, **24**, 1441–1447.

Yamaguchi, S., Sun, T. p, Kawaide, H. and Kamiya, Y. (1998) The GA2 locus of Arabidopsis thaliana encodes ent-kaurene synthase of gibberellin biosynthesis. *Plant Physiol.*, **116**, 1271–8.

Zeevaart, J.A.D. and Talon, M. (1992) Gibberellin mutants in Arabidopsis thaliana. In *Progress in Plant Growth Regulation*. pp. 34–42.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q. and Wijk, K.J. van (2008) Sorting Signals, N-Terminal Modifications and Abundance of the Chloroplast Proteome. *PLoS One*, **3**, e1994.

CHAPTER 3: Expression of a non-endogenous diterpene pathway, the diterpene resin acid pathway, in Arabidopsis.

Abstract:

Diterpene resin acids (DRAs) are produced in the resin ducts of conifers by the action of two classes of enzymes: bifunctional, plastid-localized diterpene synthases that cyclize the precursor geranylgeranyl diphosphate (GGDP) into various tricyclic diterpene olefins which are then oxidized to their corresponding DRAs by ER-localized cytochrome P450s in the family CYP720B. These enzymes are exclusively found in gymnosperms where they produce defense compounds against various biotic stressors. In this work, I attempted to reconstruct the DRA pathway in Arabidopsis by heterologous expression of diterpene synthases, *Abies grandis* abietadiene synthase (AgAS) and *Picea sitchensis* isopimaradiene synthase (PsISO) in the plastid and *Picea sitchensis* CYP720B4 in the ER. Generation of stable Arabidopsis transgenic lines containing both the activities of diterpene synthases in the plastid and CYP720B4 in the ER would allow for the *in vivo* study of non-native diterpene olefin transport from the plastid into the ER. Expression of yellow fluorescent protein (YFP)-tagged AgAS and PsISO in Arabidopsis only resulted in low-levels of the corresponding diterpene olefins, abietadiene and isopimaradiene at 16 to 540 ng g^{-1} FW. Transcript analysis showed unexpected alternatively spliced variants in the YFP region of the coding sequence, which resulted in only half of the transcripts encoding full-length YFP-tagged AgAS and PsISO. Possible explanations for low-levels of diterpene olefins produced are discussed.

Introduction:

Photosynthetic eukaryotes have a long evolutionary history with vascular and non-vascular plants (e.g.: mosses and lycophytes) diverging approximately 500 million years ago and the angiosperm/gymnosperm split occurring about 360 million years ago (Palmer *et al.*, 2004; Troitsky *et al.*, 1991). Within this evolutionary timeline, plant species are able to occupy ecological niches partly by producing specialized metabolites that allowed interaction with and accommodation of various biotic and abiotic stressors. One major group of such specialized metabolites are the terpenoids, which are made up of 5-carbon building blocks, IDP and DMADP, that are substrates for a family of prenyl diphosphate synthases that produce 10-carbon (geranyl diphosphate), 15-carbon (farnesyl diphosphate) and 20-carbon (geranylgeranyl diphosphate, GGDP) prenyl diphosphates. These prenyl diphosphates are substrates for terpene synthases that produce the corresponding mono- (10-carbon), sesqui- (15-carbon) and di- (20-carbon) terpene hydrocarbons (Tholl and Lee, 2011), which are further modified by oxidation, hydroxylation, glycosylation and cleavage reactions, which further add to the structural diversity of terpenoids produced (Tholl and Lee, 2011).

In conifers (order Coniferales), diterpene resin acids (DRAs), along with monoterpenes and to a lesser extent sesquiterpenes make up oleoresin, which is a chemical defense against herbivores and insect pathogens such as bark beetles (Keeling and Bohlmann, 2006). Oleoresin production occurs in specialized structures in conifers called resin ducts and can be induced during pathogen attack, or synthesized constitutively and stored (Bannan, 1936; Keeling and Bohlmann, 2006; Martin *et al.*, 2002). Formation of DRAs is catalyzed by a class of bifunctional diterpene synthases (diTPS) from the subfamily TPS-d3 that produce various tricyclic diterpene olefins (Figure 3.1; Martin *et al.*, 2004; Chen *et al.*, 2011). Bifunctional diTPS, which are only

found in non-seed plants and gymnosperms, retain both class I and II active sites. The class II active site (DXDD) of diTPS catalyzes the protonation-induced cyclization of GGDP to form the intermediate (+)-copalyl diphosphate [(+)-CDP]. The class I active site (DDXXD/NTE) then facilitates ionization of the diphosphate group of (+)-CDP to form reactive carbocation intermediates that are then rearranged and cyclized to form various tricyclic diterpene structures (Christianson, 2006).

In *Abies grandis* (grand fir), abietadiene synthase (AgAS) catalyzes the formation of three major diterpene olefins [levopimaradiene (34%), abietadiene (31%), neoabietadiene (28%)] and three minor diterpene olefins [pimara-8(14),15-diene (3%), palustradiene (2%) and sandaracopimaradiene (2%)] (Table 2, Vogel *et al.*, 1996; Peters *et al.*, 2000). Another bifunctional diTPS, isopimaradiene synthase (ISO), from Norway spruce (*Picea sitchensis*) catalyzes the formation of isopimaradiene as its major compound (98%) and sandaracopimaradiene as its minor compound (2%) (Table 2; Keeling *et al.*, 2011). These diterpene olefins are members of the abietane and pimarane structural groups, which are characterized by their tricyclic parent structures (Figure 3.1). In *P. sitchesis* a conifer-specific cytochrome P450, CYP720B4, is also highly expressed in resin duct cells and encodes a multifunctional enzyme that catalyzes the three-step oxidation of C-18 of abietane, pimarane, and dehydroabietane diterpene olefins converting them into their corresponding DRAs in the endoplasmic reticulum (ER) (Hamberger *et al.*, 2011). Thus, these two diterpene synthases and CYP720B4 represent a simple two-step, plastid:ER spanning conifer pathway that is absent in dicots, including the model organism *A. thaliana*.

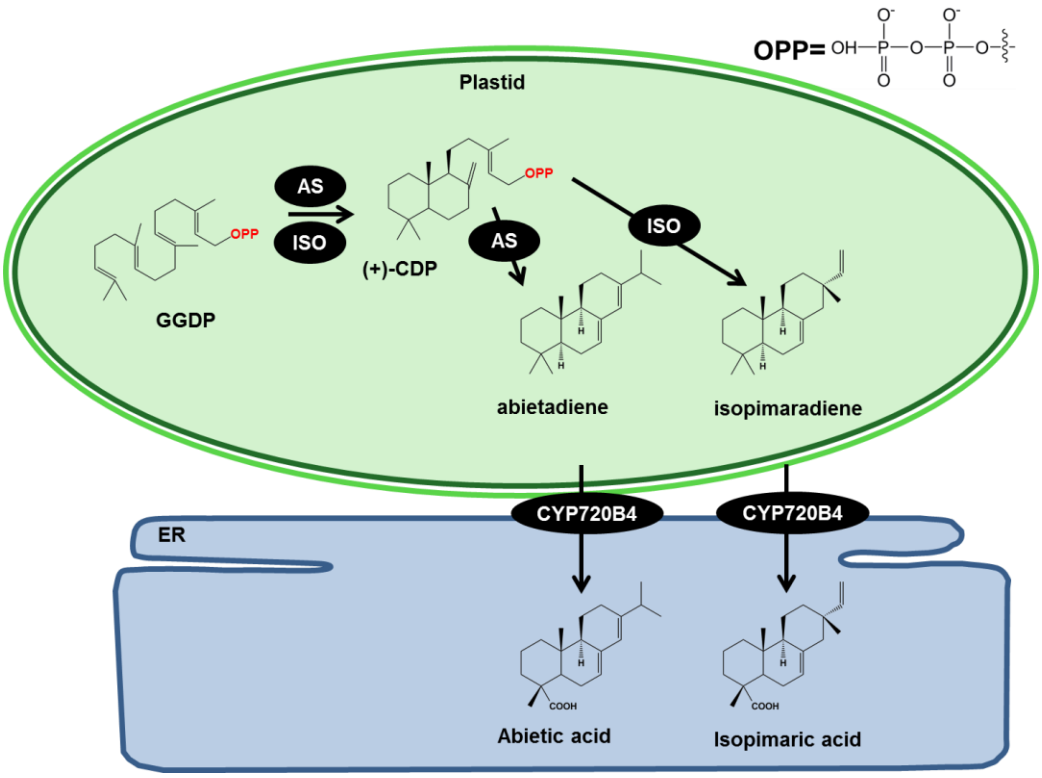


Figure 3.1: Diterpene resin acid (DRA) pathway in conifers. The plastid-localized bifunctional diterpene synthases [abietadiene synthase (AS) and isopimaradiene synthase (ISO)] catalyze the conversion of geranylgeranyl diphosphate (GGDP) into the intermediate (+)-copalyl diphosphate [(+)-CDP] which is then cyclized into either abietadiene (by AS) or isopimaradiene (by ISO) as the main diterpene olefin products of these enzymes. These diterpene olefins are then oxidized by cytochrome P450, CYP720B4 into their corresponding diterpene resin acids in the ER. Only the major products of AS and ISO are shown; their respective full product profiles are shown in Table 2. **OPP** represents the diphosphate group.

In this study, my goal was to heterologously express plastid-localized conifer-specific bifunctional diTPSs (AgAS and PsISO), together with the ER-localized CYP720B4 in Arabidopsis to investigate whether the diTPSs can access and produce high levels of their respective products from the GGDP pool in the plastid and if so, whether existing systems in Arabidopsis can transport the resulting diterpene olefins to the ER for production of DRAs. Dicots, including Arabidopsis, do not produce this class of diterpenes, with the notable exception being dehydroabietadienal, the aldehyde of dehydroabietadiene, which it is produced at extremely low levels in response to pathogen attack where it is transported in the vasculature to

distal tissues and induces systemic acquired resistance (Chaturvedi *et al.*, 2012). In the absence of pathogen challenge, dicots lack dehydroabietadienal, thus Arabidopsis should provide a clean background for measuring abietane and pimarane diterpenes and their corresponding acids to test for transport of these novel diterpenes across the plastid and ER membranes. Since Arabidopsis does not produce these compounds, nor are any homologous genes for these diterpene synthase found in the Arabidopsis genome (Chen *et al.*, 2011), it is reasonable to assume that an abietadiene or isopimeradiene transporter is unlikely to exist for these intermediates in the DRA pathway. In addition, the DRA pathway is exclusively found in gymnosperms, and since gymnosperms and angiosperms split occurred 360 million years ago, presumably any transporters that might mediate transport of diterpene olefins (none have been identified to date) will have diverged in gymnosperms and angiosperms, as have enzymes for the DRA pathway (Martin *et al.*, 2004; Keeling, Weisshaar, *et al.*, 2011; Hamberger *et al.*, 2011; Geisler *et al.*, 2016). In addition, since these DRAs are not part of general metabolism, it is unlikely that production of these DRAs will interfere with general metabolisms in Arabidopsis, in contrast to the dramatic phenotypes of GA biosynthetic mutants reported in Chapter 2 of this dissertation.

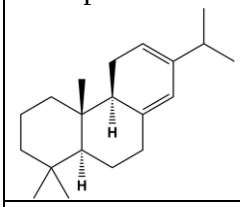
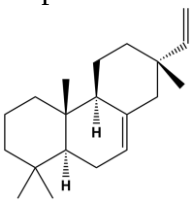
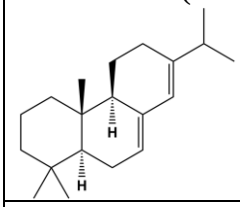
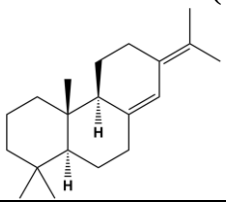
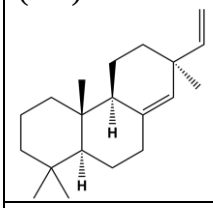
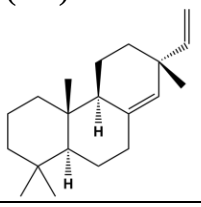
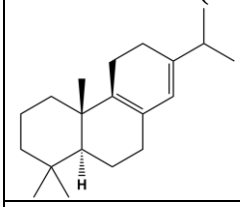
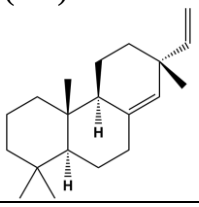
	Abietadiene synthase	Isopimaradiene synthase
Major products	Levopimaradiene (34%) 	Isopimaradiene (98%) 
	Abietadiene (31%) 	
	Neoabietadiene (28%) 	
Minor products	Pimara-8(14),15-diene (3%) 	Sandaracopimaradiene (2%) 
	Palustradiene (2%) 	
	Sandaracopimaradiene (2%) 	

Table 3: Products of bifunctional diterpene synthases. In conifers, plastid-localized bifunctional diterpene synthases (abietadiene synthase and isopimaradiene synthase) contain both class I and II active sites, converting the substrate geranylgeranyl diphosphate (GGDP) into the intermediate (+)-copalyl diphosphate [(+)-CDP] which is then cyclized into various diterpene olefins.

Results:

Reconstructing the DRA pathway in Arabidopsis by expression of bifunctional conifer diterpene synthases in plastids and a cytochrome P450 in the ER

The overall goal of this study is to create a pool of novel diterpene(s) in the plastid and assess its *in vivo* access in the ER via the generation of oxidized product(s) by the subsequent ER-resident P450. In this study, to independently probe for *in vivo* transport, I chose to express two different bifunctional conifer diterpene synthases, AgAS and PsISO, in Arabidopsis in their native organelle, the plastid. The rationale for this is that it is well known that the plastid has a large flux to GGDP to support the synthesis of carotenoids, chlorophylls, GAs and tocopherols (Beck *et al*., 2013) that should provide ample GGDP to support the synthesis of unusual diterpenes like abietadiene and isopimaradiene. I also generated another Arabidopsis transgenic line expressing *P. sitchensis* CYP720B4 (CYP720B4) in its native organelle, the ER, to allow for *in vivo* detection of transport by the production of the corresponding DRAs in transgenic lines crossed to contain both activities.

The cDNAs encoding conifer AgAS, PsISO, and CYP720B4 were codon-optimized for expression in Arabidopsis and expressed under the control of cauliflower mosaic virus (CaMV) 35S promoter with their C-termini tagged with Venus yellow fluorescent protein (vYFP). Initially, to test for functional expression of all the YFP-tagged proteins, I infiltrated four-week-old *Nicotiana tabacum* plants with *Agrobacterium tumefaciens* harboring these constructs. In

tobacco leaves infiltrated with the CYP720B4-YFP construct, the YFP signal was seen in the reticulated structure of ER and did not colocalize with chlorophyll autofluorescence, consistent with correct localization to its native organelle, the ER (Figure 3.2A). On the other hand, in tobacco leaves infiltrated either with AgAS-YFP or PsISO-YFP, YFP signal overlapped with chlorophyll autofluorescence, again consistent with the correct localization to their native organelle, the plastid (Figure 3.2, B and C). I then established stable transgenic lines for AgAS-YFP, PsISO-YFP, and CYP720B4-YFP in the WT *Landsberg* background. The resistant T_1 plants expressing AgAS-YFP, PsISO-YFP or CYP720B4-YFP appeared phenotypically indistinguishable from WT (data not shown). However, when attempting to determine the protein expression levels in these T_1 transgenic lines by immunoblotting with antibodies to YFP, I could not demonstrate detectable levels of fusion proteins. The T_1 plants were allowed to set seeds and single insert lines (segregating 3:1 for their resistance marker) were grown under standard conditions. A total of four, ten and five single insert transgenic lines were isolated for AgAS-YFP, PsISO-YFP, and CYP720B4-YFP, respectively. I again attempted to determine protein expression levels of resistant T2 lines by immunoblotting but was unsuccessful (Figure 3.3). A band with a molecular weight 125 kDa, which is the predicted size for YFP-tagged AgAS and PsISO, was observed in all transgenic lines but was absent in the WT. This band, however, was most likely a non-specific band, as it was also present in CYP720B4-YFP transgenic lines, instead of the predicted size of 83 kDa. I also attempted confocal imaging on these stable T2

Arabidopsis transgenic lines but was unable to observe any YFP fluorescent signal. These data are in sharp contrast to the strong signals obtained in the transient expression assays.

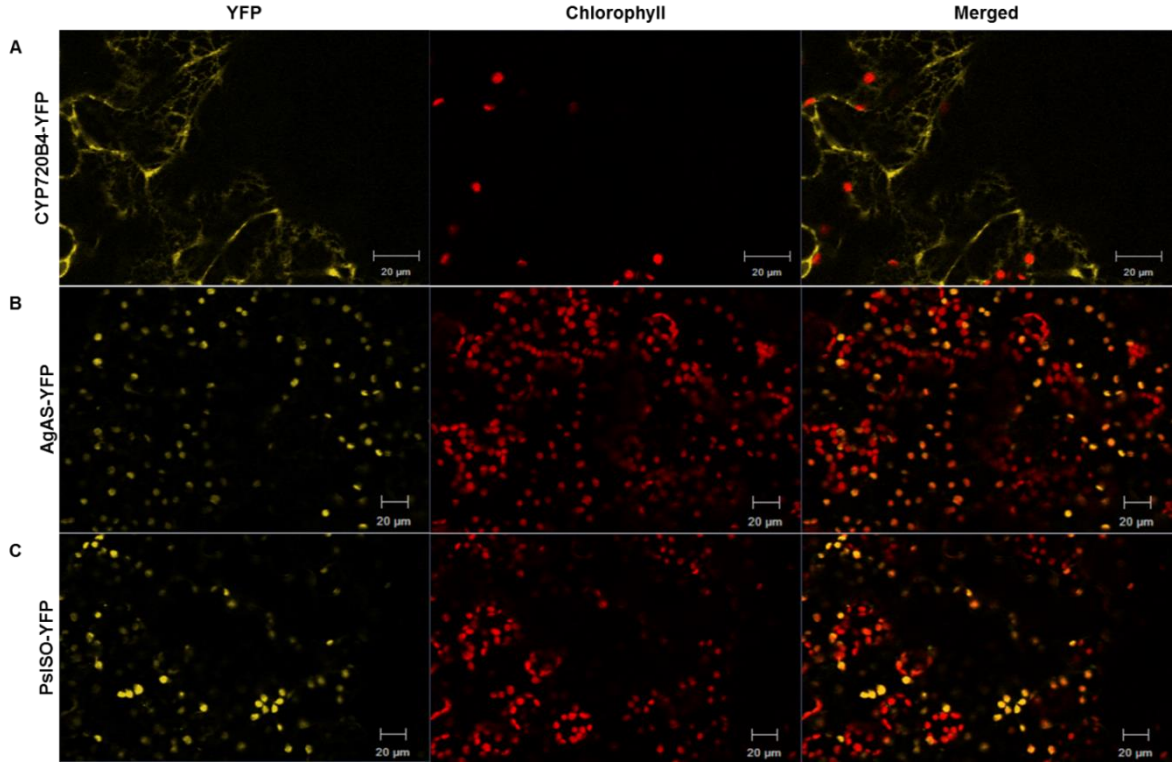


Figure 3.2: Subcellular localization of transiently-expressed YFP-tagged CYP720B4, AgAS and PsISO. Confocal microscopy of *N. tabacum* leaves infiltrated with *Agrobacterium* harboring plasmid expressing either CYP720B4-YFP, AgAS-YFP or PsISO-YFP. The first column shows the yellow fluorescent protein (YFP) signal in *yellow*, the second shows chlorophyll autofluorescence in *red* and the third column shows an overlay of these two channels. Scale bar is 20 µm.

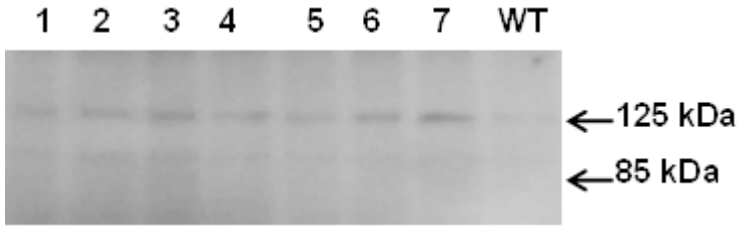


Figure 3.3: Protein expression in transgenic lines. Twenty microgram of total proteins of resistant T2 lines were loaded on 8% SDS-PAGE gel. YFP-tagged proteins were assessed with immunoblotting against anti-GFP antibodies (cross reacts with YFP). Estimated sizes for YFP-tagged proteins: AgAS-YFP 127 kDa, PsISO-YFP 128kDa, and CYP720B4-YFP 83 kDa. Lane 1-3 are independent lines for PsISO-YFP, lane 4-5 for AgAS-YFP and lane 6-7 for CYP720B4-YFP. The estimated locations for the positive signal are indicated by arrows.

Metabolite analysis of transgenic lines

Although I could not detect fusion protein in the transgenic lines with YFP antibody, low levels of the enzyme could still be produced and generate the diterpene olefin products and this might be a more sensitive measure of the enzyme activity. I measured the levels of diterpenes in the segregating T2 generation of both the AgAS-YFP and PsISO-YFP lines. Since the trait for AgAS and PsISO is dominant, ¾ of the segregating T2 plants would still be expected to produce the diterpene olefins.

I analyzed four-week-old pooled rosette leaves from the T2 transgenic lines and WT using gas chromatography-mass spectrometry and monitored the unique mass-over-charge (m/z) of 229 for these diterpene olefins. A peak was detected at 11.66 min in extracts of AgAS-YFP transgenic lines that was absent in the WT extract (Figure 3.4A). Similarly, a peak was detected at the same retention time in *N. tabacum* leaves infiltrated with the AgAS-YFP construct, which was absent in no-infiltration negative controls (Figure 3.4A). The mass spectrum of the 11.66 min peak matched the published mass spectrum of abietadiene (Keeling, Madilao, *et al.*, 2011), confirming that the 11.66 min peak is indeed abietadiene. Using the same m/z 229, I also detected a peak at 10.77 min that was present in PsISO-YFP transgenic lines (Figure 3.4A) that was also absent in the WT extract. The mass spectrum of 10.77 min peak also matched the published mass spectrum for isopimaradiene (Gnanasekaran *et al.*, 2015). Using an *ent*-kaurene standard curve (another diterpene olefin with similar ionization characteristics), the levels of abietadiene present in the transgenic lines were estimated to range from 110 to 540 ng g^{-1} fresh weight (FW) (Figure 3.4C) while the levels of isopimaradiene in the PsISO-YFP transgenic lines were much lower at 14 to 62 ng g^{-1} FW (Figure 3.4D). The levels of abietadiene and isopimaradiene detected in these transgenic lines were 100-1000 fold lower than the levels of

ent-kaurene in plastid:CPS-YFP lines reported in Chapter 2 (Figure 2.6B). These extremely low levels of abietadiene and isopimaradiene in the AgAS-YFP and PsISO-YFP transgenic lines compared to *ent*-kaurene levels in plastid:CPS-YFP overexpression lines occurred, despite the fact that all diterpene synthases were expressed using the same CaMV 35S promoter.

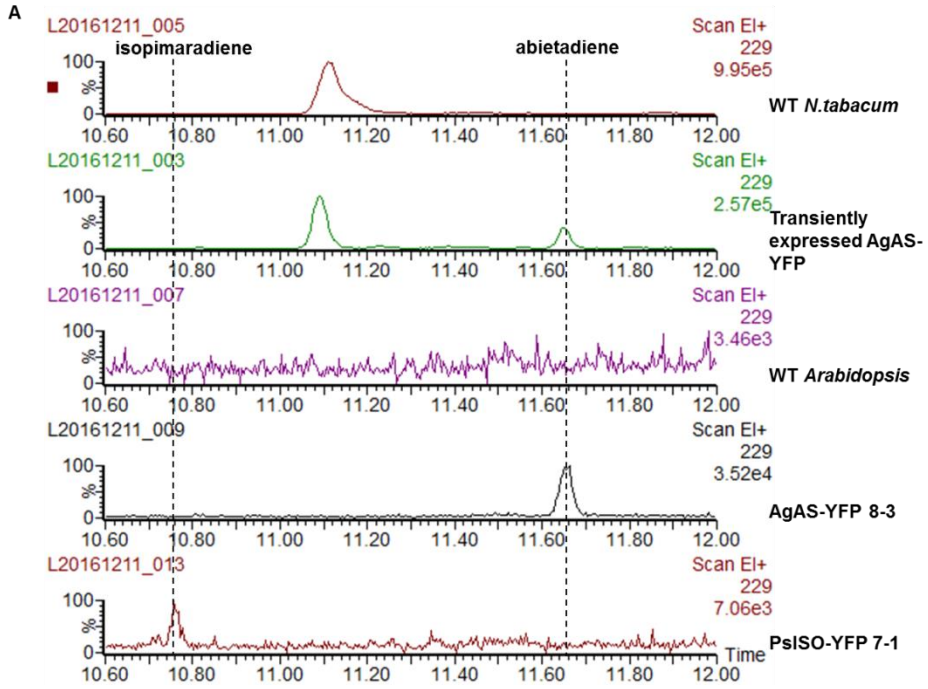
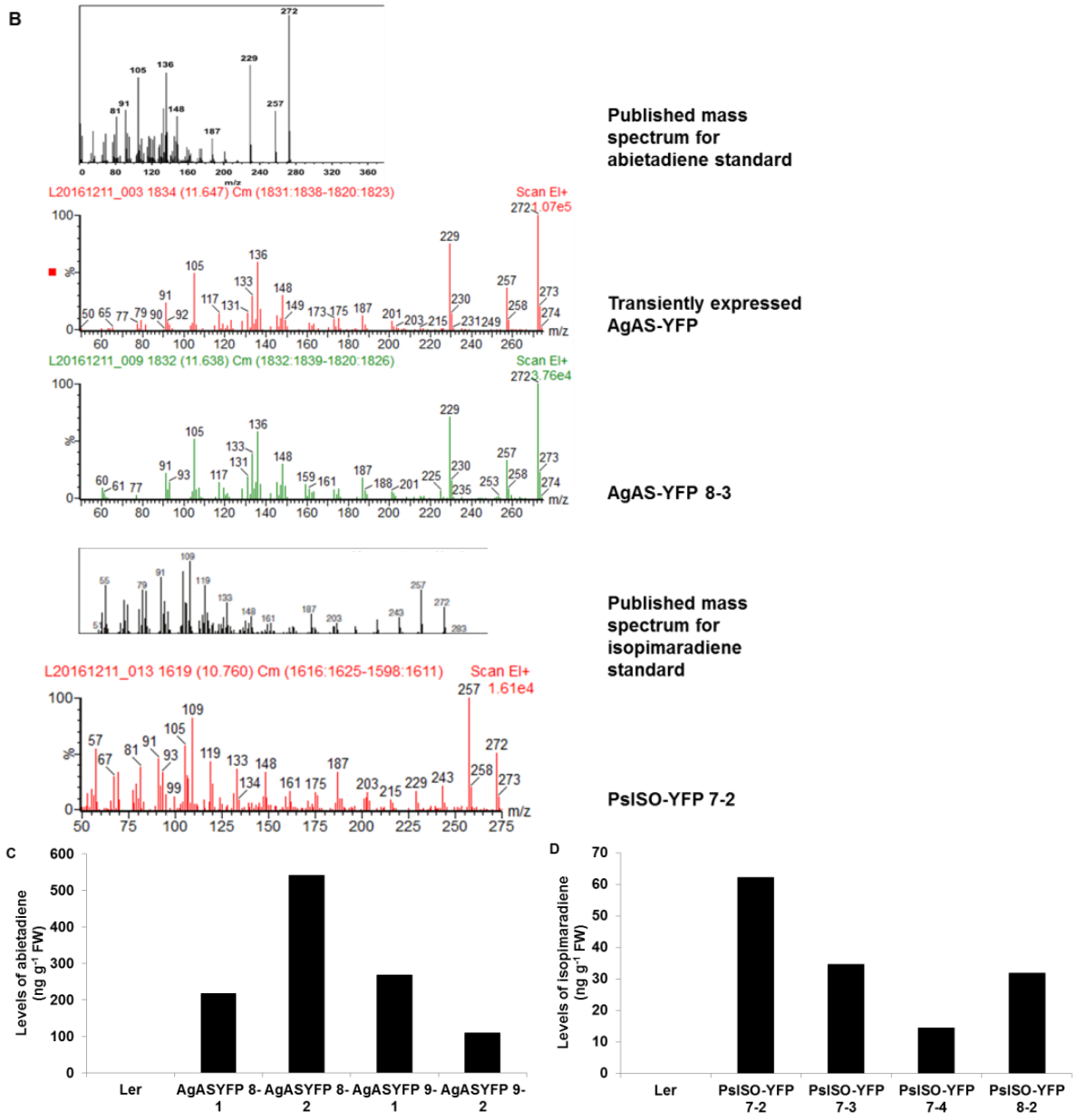


Figure 3.4: Metabolite analysis of transient and stable AgAS-YFP and PsISO-YFP transgenic lines. A) Extracted ion chromatograms at m/z 229 for the detection of abietadiene (at 11.66 min) and isopimaradiene at 10.75 min using GC-MS analysis.

Figure 3.4: (cont'd)



B) Mass spectrum of the abietadiene peak at 11.67 min in AgAS-YFP samples and isopimaradiene peak at 10.75 min in PsISO-YFP samples. The published mass spectrum for abietadiene was obtained from Keeling, Mandilao *et al.*, 2011 while the published mass spectrum for isopimaradiene peak was obtained from Gnanasekaran *et al.*, 2015. **C)** Levels of abietadiene in four independent AGAS-YFP transgenic lines **D)** Levels of isopimaradiene in four independent PsISO-YFP lines.

Transcript analysis of transgenic lines

The fact that abietadiene and isopimaradiene could be detected in transgenic lines, albeit at quite low levels (Figure 3.4, C and D), despite the inability to detect fusion protein signals for the introduced enzymes by immunoblotting (Figure 3.3) and confocal microscopy led me to test whether there were problems with expression from the constructs. I performed a transcript analysis by reverse-transcriptase PCR (RT-PCR) on mRNA isolated from bulked T2 seedlings.

Full-length transcripts of the appropriate sizes could be amplified in the transgenic lines using a forward primer on the TMV Ω enhancer and a reverse primer 3' of the stop codon, both of which are found in all constructs. A 3.2-kb band was detected for both AgAS-YFP and PsISO-YFP transgenic lines while a 2.3-kb band was detected in two independent CYP720B4-YFP transgenic lines (Figure 3.5A). The expression levels of both AgAS-YFP and CYP720B4-YFP were high, while PsISO-YFP showed a moderately lower expression (Figure 3.5, A and B). No amplified product was present in WT cDNA using this same primer pair (Figure 3.5B).

Amplification of smaller fragments across the transcripts, however demonstrated heterogeneity in the YFP portion of each transcript, which led to me clone these PCR fragments into the pGEMT-Easy vector followed by *E. coli* transformation (Figure 3.6). I then performed sequencing on twelve individual colonies containing PCR fragments spanning the YFP portion of each transcript and found that all constructs have alternatively spliced versions present that resulted in some transcripts missing either 87, 67 or 84 bp (A1, A2, A3, respectively; Figure 3.6). Transcript analysis of the YFP region in AgAS-YFP, PsISO-YFP, and CYP720B4-YFP showed that 50-60% of these transcripts were alternatively spliced while 40-50% encoded the full-length, unaltered YFP-tagged transcript. Analysis of this YFP region with an Arabidopsis alternative

splicing prediction tool (www.cbs.dtu.dk/services/NetPGene/) did not predict these regions to be alternatively spliced.

The transcript analysis showed that there were four combinations of alternatively spliced transcripts, with the deletion of the A3 region being the most predominant spliced variants (44%, Table 3). The deletion of the A3 region resulted in an in frame deletion of amino acids 127 to 154 of the YFP (Figure 3.7). Other combinations included deletion of A1+A2 (2.8%), A1+A2+A3 (2.8%), and A1+A3 (5.6%, Table 3). The deletion of both A1+A2 and A1+A2+A3 resulted in a 28 amino acid in frame deletion (amino acids 6-34) followed by a second out of frame deletion that resulted in 5 missense amino acids from amino acids 51-55 followed by a premature stop (Figure 3.7). Translation of these spliced variants will produce a truncated protein composed of 26 amino acids. The deletion of A1+A3, on the other hand, resulted in the same 28 amino acid in frame deletion and a second in frame deletion from amino acids 127 to 154 followed by the rest of the mature sequence (Figure 3.7). Since the alternative splicing only occurred in the YFP region, all proteins produced by alternatively spliced transcripts should contain intact diterpene synthase coding sequence and retain their enzymatic activities. Most of the altered proteins would also contain large portions of the YFP coding region, which would result in no fluorescence, but would still contain antigenic region that should still react with the polyclonal GFP antibody used.

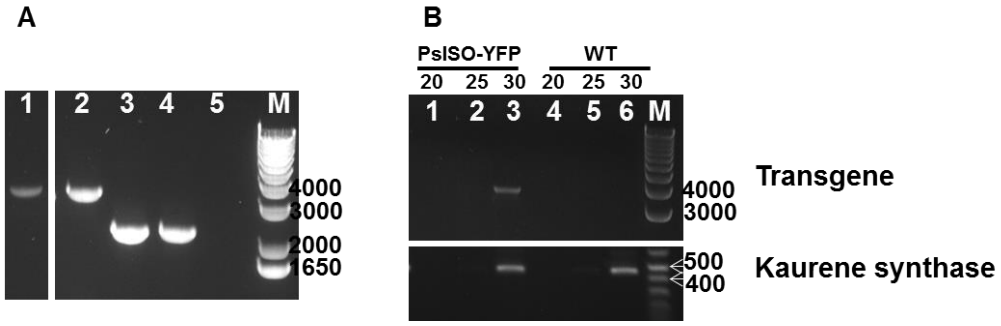


Figure 3.5: Agarose gel electrophoresis detecting full-length transcripts of the transgene. A) Full-length transcripts of PsISO-YFP (lane 1), AgAS-YFP (lane 2) and CYP720B4-YFP (Lane 3 and 4) were amplified for 30 cycles from mRNA isolated from resistant transgenic lines. The expected size for a full-length transcript of PsISO-YFP and AgAS-YFP is 3.4 kbp, while CYP720B4-YFP is 2.3 kbp. Lane 5 is water only negative control. **B)** Full-length transcripts of the transgene were determined using the same primer pair as **A)** at increasing cycle numbers (20, 25, or 30) in PsISO-YFP transgenic line and a WT control. As a control for cDNA, a primer pair amplifying 443 bp of the kaurene synthase gene was used.

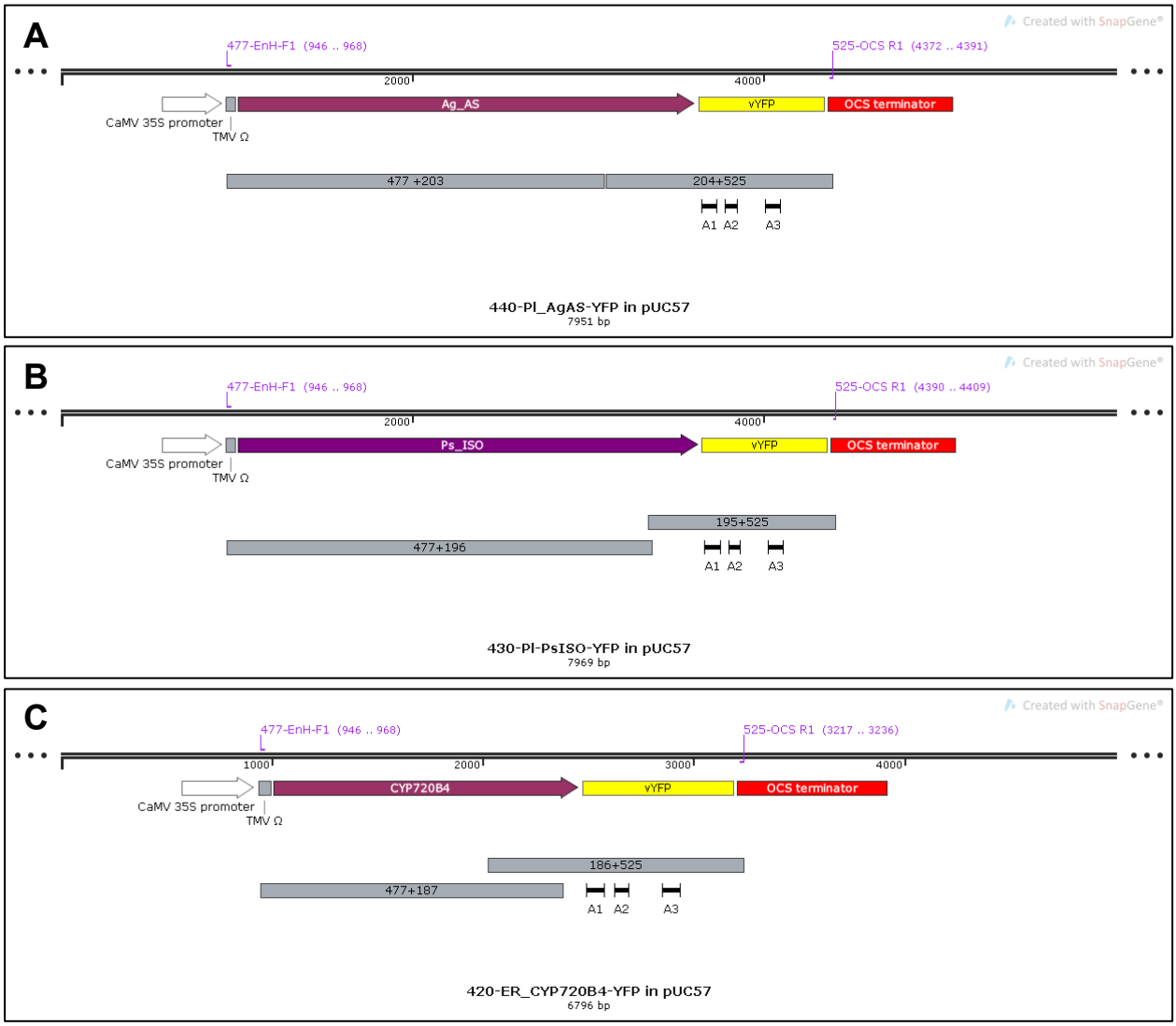


Figure 3.6: Transcript analysis on segments of the transgene. Vector maps showing primer pair (477 and 525) used to amplify full-length transcript in **A)** AgAS-YFP, **B)** PsISO-YFP, and **C)** CYP720B4-YFP. Two regions of AgAS-YFP, PsISO-YFP and CYP720B4-YFP were amplified and ligated into the pGEMT-Easy vector. A1 (87 bp), A2 (67 bp) and A3 (84 bp) were alternatively spliced regions in the YFP region.

YFP variant	AgAS-YFP	PsISO-YFP	CYP720B4-YFP	Percentage
Full-length YFP	5	7	4	44.4%
A1+A2 deleted	0	1	0	2.8%
A1+A2+A3 deleted	0	1	0	2.8%
A1+A3 deleted	1	1	0	5.6%
A3 deleted	6	2	8	44.4%

Table 4: Distribution of YFP variants in each construct. A total of twelve individual clones from were sequenced and aligned to the YFP reference sequence. A1, A2, and A3 refer to the alternatively spliced regions shown in Figure 3.6.

Figure 3.7: Protein sequence for alternatively spliced transcripts. The sequence of full-length YFP is shown on the top line as a reference. Deletion of A1+A2 and A1+A2+A3 result in a 28 amino acid in frame deletion (amino acids 6-34) followed by a second out of frame deletion that results in 5 missense amino acids (in black rectangle) from amino acids 51-55 followed by a premature stop. Deletion of A1+A3 results in the same 28 amino acid in frame deletion and a second in frame deletion from amino acids 127 to 154 followed by the rest of the mature sequence. Deletion A3 only results in frame deletion of amino acids 127 to 154.

Discussion:

The goal of this work was to install a novel diterpene pathway that is not endogenous to Arabidopsis to test for the ability to transport and access novel diterpenes across the plastid:ER interface. As DRAs have specifically evolved in conifers which are separated evolutionary from Arabidopsis by 360 million years ago, it is unlikely that transporters for these diterpene olefins exist in Arabidopsis. Prior to initiating the work in this chapter, the enzymatic reactions of bifunctional diTPSs, AgAS and PsISO had been elucidated using recombinant proteins produced in *E. coli* (Peters *et al.*, 2000; Keeling, Weisshaar, *et al.*, 2011). Subsequently, Brückner and Tissier (2013) showed that *Agrobacterium*-mediated transient expression of plant diTPSs in *Nicotiana benthamiana*, such as cembratrienol synthase from *Nicotiana sylvestris*, casbene synthase from *Ricinus communis* and levopimaradiene synthase from *Ginkgo biloba,* allowed for the accumulation of the respective diterpene products: cembratrienol, casbene, and levopimaradiene in the heterologous plant host. The expression of these diTPSs resulted in differing amounts of product accumulated, the lowest being levopimaradiene at 100 ng/cm^2 leaf infiltrated and the highest being cembratrienol at 700 ng/cm^2 leaf infiltrated (Brückner and Tissier, 2013). Assuming that the average leaf density of young *N. benthamiana* is 0.05 g/cm^2 (Robert *et al.*, 2013), the levels reported in Brückner *et al.* (2013) ranged from 2000 ng g^{-1} FW (levopimaradiene) to 14000 ng g^{-1} FW (cembratrienol). In this work, the expression of AgAS-YFP and PsISO-YFP in stable Arabidopsis transgenic lines showed a much lower accumulation of diterpene olefin products accumulated, which ranged from 14 to 540 ng g^{-1} FW. Brückner and Tissier (2013) also demonstrated that the levels of diterpene produced in *N. benthamiana* increased at 3.5-fold with the combined transient expression of diTPSs with enzymes upstream in the MEP pathway such as 1-deoxy-D-xylulose-5-phosphate synthase (DXS) and GGDP

synthase. Subsequent to the work in this chapter, Gnanasekaran *et al.* (2015) also utilized the *Agrobacterium*-mediated transient expression in *N. benthamiana* to heterologously express *P. abies* ISO (which shares 97% percent identity to PsISO in this work) and the same *P. sitchensis* CYP720B4 used in this work and showed that coinfiltration of plastid-targeted ISO and ER-targeted CYP720B4, together with DXS and GGDPS resulted in the depletion of isopimaradiene and concomitant production of isopimaric acid as high as 3000 ng g^{-1} FW (Gnanasekaran *et al.*, 2015). Conservative estimation of the levels of isopimaradiene observed in Gnanasekaran *et al.* (2015) and this work showed at least six-fold higher isopimaradiene accumulated than that of the highest levels of abietadiene accumulated in AgAS-YFP line 8-2 (Figure 3.4C). Perhaps most importantly, the generation of isopimaric acid by Gnanasekaran *et al.* (2015) makes clear that ER-localized *P. sitchensis* CYP720B4 is able to access its substrate isopimaradiene produced by plastid localized *P. abies* ISO when transiently expressed in *N. benthamiana*, an angiosperm.

While the work by Gnanasekaran *et al.* (2015) provided evidence for accessibility of plastid-synthesized diterpene olefin by ER-localized cytochrome P450 in transient, high-level expression, the approach has some potential disadvantages that could impact and confound the interpretation of the results. These include very short-term heterologous protein production (3-5 days) and heterogeneity in expression levels that could lead to some cells accumulating high amount of proteins that could overwhelm the protein transport systems and result in protein mislocalization (Wroblewski *et al.*, 2005; Kapila *et al.*, 1997). Gnanasekaran *et al.* (2015) attempted to retarget CYP720B4 to the plastid, and reported two distinct protein bands in their immunoblotting analysis, indicating that half of the protein was successfully imported into the plastids and half remained unprocessed in the cytosol or ER. For these reasons, I have opted to rely on generating stable Arabidopsis transgenic lines that would consistently express plastid-

localized bifunctional AgAS and PsISO and ER-localized CYP720B4. One can also select stable transgenic lines with differing levels of protein expression (low, medium and high) to assess dosage effects (as was done in Chapter 2 of this dissertation) and thorough confocal imaging of YFP-tagged proteins in ensuring the correct localization in these selected transgenic lines. The overall plan was to create a pool of non-polar diterpene olefins in plastids of AgAS-YFP and PsISO-YFP transgenic lines and then cross these selected expression lines with CYP720B4-YFP transgenic lines to reconstruct the simple two-step pathway and assess the efficacy and efficiency of *in vivo* transfer of the diterpene olefins to the ER, as measured by the production of corresponding DRAs in the crossed lines.

A prerequisite for these experiments is accumulation of high enough levels of the olefin substrate such that if modest levels of hydroxylation occur (e.g., 10%), it could be quantified. As described, expression of bifunctional diterpene synthases, AgAS-YFP and PsISO-YFP in the plastids generated their corresponding diterpene olefins, albeit at quite low levels. In AgAS-YFP, the levels of abietadiene produced in independent transgenic lines range from 110 to 540 ng g^{-1} FW, while in PsISO-YFP transgenic lines, the levels of isopimaradiene were even lower, which ranged from 14 to 62 ng g^{-1} FW, which stand 100-1000 fold lower than when CPS is expressed alone [up to 52000 ng g^{-1} FW in the plastid:CPS-YFP transgenic lines (Figure 2.6B)] or in combination with KS [up to 69000 ng g^{-1} FW (Fleet *et al.*, 2003)]. Conversion of modest percentages (e.g. 10%) of the abietadiene or isopimaradiene produced would generate the corresponding DRA product levels at near or below the limits of detection thus making assessment of the accumulation of specific metabolites in different compartments extremely challenging. For these and other reasons we elected to abandon this research effort.

There are several potential reasons for the low levels of diterpene olefins produced by AgAS-YFP and PsISO-YFP expressed in Arabidopsis and despite high levels of transcripts (Figure 3.5A), only low levels of the bifunctional diTPS proteins being produced. Transcript analysis showed about half of the transcripts contained cryptic splice sites (not detectable by software) that would produce truncated YFP proteins, and hence potentially undetectable with polyclonal GFP antibody that cross-reacts with YFP. However, half of the transcripts would produce full-length YFP-tagged proteins, which in turn should be detectable by both confocal imaging and immunoblotting analysis. Despite this, no YFP signal detected in both confocal imaging and immunoblotting analysis in stable Arabidopsis transgenic lines. A likely explanation is that the remaining full-length transcript may be cosuppressed, either post-transcriptionally or during translation despite the constructs being codon-optimized to limit such cosuppression. On the other hand, the positive YFP signal detected by confocal imaging (Figure 3.2) and at least 10-fold higher levels of abietadiene produced (Figure 3.4A) in the transient expression assay indicated that the levels of protein produced in transient expression were most likely much higher than in the stable transgenic lines. However, this true level of protein expression in the transgenic lines is difficult to address given that about 50% of the transcripts resulted in a truncated YFP tag and native antibodies that would cross-react with the AgAS and PsISO proteins are not available.

Another factor potentially contributing to the lower levels of diterpene olefins produced could be the inability of the bifunctional diTPSs to form a productive complex with the existing GGDP producing machinery in the plastid and hence potentially being susceptible to proteolysis. The whole plant phenotypes of AgAS-YFP and PsISO-YFP were indistinguishable from WT, which indicated that in these transgenic lines, GGDP was not being rerouted from either

chlorophyll, carotenoids, or GA production to produce diterpene olefins, as was observed in metabolic engineering study involving phytoene synthase (Fray *et al.*, 1995) which also utilizes the precursor GGDP. In my work, I speculate that heterologous expression of gymnosperm AgAs or PsISO in Arabidopsis may not allow for efficient protein-protein interaction between these bifunctional diTPS and the main Arabidopsis plastid-localized GGDPS, GGPS11, thus limiting the amount of GGDP accessible by either AgAS or PsISO. In Arabidopsis, the expression of GGDPS11 is highly correlated with the plastidic-localized chlorophyll, carotenoids, and GA biosynthetic enzymes (Meier *et al.*, 2011). In tomato and pepper chromoplasts, GGDPS was purified in a high-molecular-weight complex with isopentenyl diphosphate isomerase and phytoene synthase, suggesting that this could be part of a metabolon for synthesizing carotenoids (Camara, 1993; Fraser *et al.*, 2000; Nisar *et al.*, 2015). Metabolon formation allows for concerted channeling of pathway intermediates to consecutive enzymes in a biosynthetic pathway (Moller, 2010). In gymnosperms, one to four GGDP synthase is present in a species (Coman *et al.*, 2014), suggesting that some of the GGDP synthases could be involved exclusively for specialized metabolism. For example, in *Picea abies*, IDS5 (a GGDP synthase) is thought to be involved in oleoresin production while IDS6 is involved in the synthesis of chlorophylls and GAs (Schmidt and Gershenzon, 2007). In Chapter 2 of this dissertation, overexpression of CPS in the plastid led to levels of *ent*-kaurene 100-1000 fold higher than AgAS-YFP or PsISO-YFP, which suggests that there is a large flux of GGDP channeled into and capable of supporting the production of GAs. However, in this work, it is clear that AgAS-YFP and PsISO-YFP were not able to access this pool of GGDP available in the plastid.

Another likely explanation for the low levels of diterpene olefins produced in the transgenic lines could be due to the YFP tag impacting the protein activities of AgAS and PsISO.

However, this again seems unlikely as in Chapter 2 of this dissertation, copalyl diphosphate synthase, which is a monofunctional diTPS in the GA pathway was still active with its C-termini tagged with a fluorescent protein, as evident by the high amount of *ent*-kaurene produced in the transgenic lines. In addition, the addition of FLAG tag at the C-termini of *P. abies* ISO also did not appear to impact protein function in transient assays (Gnanasekaran *et al.*, 2015).

Finally, Table 3 shows that about half of the mRNA processing products for AgAS-YFP and PsISO-YFP were improperly spliced within the YFP coding region, resulting in truncated proteins that could be inactive or form inactive complexes. Bifunctional diTPSs contain both class I and II active sites on a single polypeptide chain. The class I active site is located in the γ domain at the C-terminus while the class II active site is contained in the αβ domains at the N-terminus (Zi *et al.*, 2014). The crystal structure for AgAS has been elucidated and showed that the C-terminus end projects out into the solution (Zhou *et al.*, 2012). The most common alternatively spliced variant (deletion of A3) would result in the in-frame deletion of amino acids 127 to 154 of the YFP protein (Figure 3.7, Table 3), thus resulting in misfolding of YFP tag. It is unlikely that any of the alternatively spliced variants will result in the misfolding of the AgAS and PsISO protein that would render the protein inactive. Another possibility is that the truncated proteins may form an inactive complex with other proteins in the MEP pathway as discussed above. It is unlikely, however, that diTPS form inactive complex with itself, as various active diTPS including CPS and AgAS have been purified as monomeric proteins in size exclusion chromatography (Duncan and West, 1981; Vogel *et al.*, 1996).

In this work, as the starting amount of diterpene olefins were already low, I decided not to cross the AgAS-YFP and PsISO-YFP lines into the CYP720B4-YFP background and characterize the resulting low levels of products and instead concentrate on the work reported in

Chapter 2. In addition, due to the low levels of YFP-tagged protein present, I would not be able to reliably demonstrate the correct localization of the YFP-tagged AgAS, PsISO and CYP720B4 by either confocal imaging or immunoblotting analysis as was done with the localization of CPS-YFP protein in Chapter 2 of this dissertation, thus making it difficult to conclude that any formation of DRAs is exclusively due to access of plastid-synthesized diterpene olefins by ER-localized CYP720B4.

Materials and methods:

Vector construction and generation of transgenic lines

Constructs encoding plastid-targeted bifunctional diTPS (AgAS and PsISO) and ER-targeted cytochrome P450 (CYP720B4) were codon-optimized and synthesized by GenScript (Piscataway, New Jersey). All constructs were flanked with a SpeI site at the 5'end and a BamHI site at the 3'end. Fragments containing the coding sequences were then ligated into a pUC57 entry vector between the cauliflower mosaic virus (CaMV) 35S promoter and an in frame, C-terminus yellow fluorescent protein (YFP) tag which was codon-optimized. The constructs for bifunctional diterpene synthases, AgAS-YFP and PsISO-YFP were subcloned into the kanamycin selectable pART27 plant expression vector while CYP720B4-YFP were subcloned into pMLBART plant expression vector with a glufosinate resistance marker. All constructs were transformed into the WT *Landsberg* background by the floral dipping method (Clough and Bent, 1998). For the AgAS-YFP and PsISO-YFP constructs, T_1 seedlings were selected on Murashige-Skoog plates containing 50 μg/μL kanamycin. For the CYP720B4-YFP construct, T_1 resistant seedlings were selected on soil by spraying with 300 μM glufosinate-ammonium (Ignite®, Bayer). Single insert homozygous transgenic lines were identified by resistance segregation ratios and progeny used for subsequent experiments.

Immunoblotting analysis

Total protein concentrations were determined using the RC DC™ protein assay (Bio-Rad). 20 μg of proteins were dissolved in 6x Laemmli buffer, separated on 8% SDS-PAGE gel and transferred to polyvinylidene fluoride (PVDF) membranes (GE healthcare). Membranes were blocked overnight with 5% BSA at 4°C, probed with an anti-GFP primary antibody (Life Technologies) at 1:2500 dilution and incubated for 1 hour at room temperature. Membranes were

subsequently washed four times with 1xTBST and incubated with horseradish peroxidase-conjugated goat anti-rabbit (Bio-Rad) diluted 1:10000 for 1 hour at room temperature. Membranes were washed and detected using the Clarity Western ECL Blotting substrates (Bio-Rad) and imaged on ChemiDoc (Bio-Rad).

Transient expression in tobacco and confocal imaging

This was performed as outlined in Chapter 2.

Metabolite analysis

Extraction of ~0.5 g Arabidopsis rosette leaves and transiently-expressed tobacco leaves followed the extraction outlined in *ent*-kaurene analysis described in Chapter 2. The GC-MS method was slightly different. The initial oven temperature was 40°C, held for 2 min, then raised by 40°C/min to 210°C, 4°C/min to 233°C, 50°C/min to 340°C and held at 340°C for 4.36 min. The mass spectrometer was set to scan for masses 30-400. QuanLynx was used for data analysis. Quantification was done using the mass-to-charge ratio (m/z) of 229 for abietadiene and isopimaradiene and 284 for methyl heptadecanoate.

Transcript analysis

Total mRNA was extracted from two-week-old T2 seedlings grown on MS plates using the RNeasy Plant Mini Kit (Qiagen) followed by DNase I on-column treatment (Qiagen). cDNA synthesis was performed from 1 μg of total mRNA using SuperScript II reverse transcriptase (Invitrogen) and oligo dT-18. Primers used for full-length transcript analysis were: Forward 5'-CCAACAACAACAAACAACAAACA-3'; Reverse 5'-AGGCGTCTCGCATATCTCAT-3'. Primers used to amplify 3' of the transgene to 3' of YFP were: AgAS-YFP, Forward 5'-GAGGCTGAATGGTCTGAGGC-3'; Reverse 5'-AGGCGTCTCGCATATCTCAT-3'; PsISO-YFP, Forward 5'-AGAGGACAGGGTGAAGTTGC-3'; Reverse 5'-

AGGCGTCTCGCATATCTCAT-3'; CYP720B4-YFP, Forward 5'-TGGCTTTCACTCAGTGTGTT-3'; Reverse 5'-AGGCGTCTCGCATATCTCAT-3'. Amplified PCR products were run on 1% agarose gel with ethidium bromide and visualized under UV illumination. Amplified PCR products were also subcloned into a pGEM-T Easy vector for isolation of single clones. Single bacterial colonies with PCR fragments were identified by colony PCR and twelve positive colonies for each fragment were grown overnight in Luria-Bertani liquid media at 37°C in the presence of carbenicillin (100 μg/mL). Plasmids were isolated and sequenced by Sanger sequencing at MSU Genomics Research Technology Support Facility. Nucleotide sequences were aligned using SnapGene Viewer.

REFERENCES

REFERENCES

Bannan, M.W. (1936) Vertical rein ducts in the secondary wood of the Abietineae. *New Phytol.*, **35**, 11–46.

Brückner, K. and Tissier, A. (2013) High-level diterpene production by transient expression in Nicotiana benthamiana. *Plant Methods*, **9**, 46.

Camara, B. (1993) Plant phytoene synthase complex: Component enzymes, immunology, and biogenesis. *Methods Enzymol.*, **214**, 352–365.

Chaturvedi, R., Venables, B., Petros, R. a, Nalam, V., Li, M., Wang, X., Takemoto, L.J. and Shah, J. (2012) An abietane diterpenoid is a potent activator of systemic acquired resistance. *Plant J.*, **71**, 161–172.

Chen, F., Tholl, D., Bohlmann, J. and Pichersky, E. (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.*, **66**, 212–229.

Christianson, D.W. (2006) Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev.*, **106**, 3412–42.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.*, **16**, 735–43.

Coman, D., Altenhoff, A., Zoller, S., Gruissem, W. and Vranová, E. (2014) Distinct evolutionary strategies in the GGPPS family from plants. *Front. Plant Sci.*, **5**, 230.

Duncan, J.D. and West, C.A. (1981) Properties of kaurene synthetase from Marah macrocarpus endosperm: evidence for the participation of separate but i Enzymes. *Plant Physiol.*, **68**, 1128–34.

Fleet, C.M., Yamaguchi, S., Hanada, A., Kawaide, H., David, C.J., Kamiya, Y. and Sun, T. (2003) Overexpression of AtCPS and AtKS in Arabidopsis Confers Increased ent -Kaurene Production But No Increase in Bioactive Gibberellins 1. *Plant Physiol.*, **132**, 830–839.

Fraser, P.D., Schuch, W. and Bramley, P.M. (2000) Phytoene synthase from tomato (Lycopersicon esculentum) chloroplasts-partial purification and biochemical properties. *Planta*, **211**, 361–369.

Fray, R.G., Wallace, A., Fraser, P.D., Valero, D., Hedden, P., Bramley, P.M. and Grierson, D. (1995) Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. *Plant J.*, **8**, 693–701.

Geisler, K., Jensen, N.B., Yuen, M.M.S., Madilao, L. and Bohlmann, J. (2016) Modularity of Conifer Diterpene Resin Acid Biosynthesis: P450 Enzymes of Different CYP720B Clades Use Alternative Substrates and Converge on the Same Products. *Plant Physiol.*, **171**, 152–164.

Gnanasekaran, T., Vavitsas, K., Andersen-Ranberg, J., Nielsen, A.Z., Olsen, C.E., Hamberger, B. and Jensen, P.E. (2015) Heterologous expression of the isopimaric acid pathway in Nicotiana benthamiana and the effect of N-terminal modifications of the involved cytochrome P450 enzyme. *J. Biol. Eng.*, **9**, 24.

Hamberger, B., Ohnishi, T., Hamberger, B., Séguin, A. and Bohlmann, J. (2011) Evolution of Diterpene Metabolism: Sitka Spruce CYP720B4 Catalyzes Multiple Oxidations in Resin Acid Biosynthesis of Conifer Defense against Insects. *Plant Physiol.*, **157**, 1677–1695.

Kapila, J., Rycke, R. De, Montagu, M. Van and Angenon, G. (1997) An Agrobacterium-mediated transient gene expression system for intact leaves. *Plant Sci.*, **122**, 101–108.

Keeling, C.I. and Bohlmann, J. (2006) Diterpene resin acids in conifers. *Phytochemistry*, **67**, 2415–23.

Keeling, C.I., Madilao, L.L., Zerbe, P., Dullat, H.K. and Bohlmann, J. (2011) The primary diterpene synthase products of Picea abies levopimaradiene/abietadiene synthase (PaLAS) are epimers of a thermally unstable diterpenol. *J. Biol. Chem.*, **286**, 21145–53.

Keeling, C.I., Weisshaar, S., Ralph, S.G., Jancsik, S., Hamberger, B., Dullat, H.K. and Bohlmann, J. (2011) Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (Picea spp.). *BMC Plant Biol.*, **11**, 43.

Martin, D., Tholl, D., Gershenzon, J. and Bohlmann, J. (2002) Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiol.*, **129**, 1003–18.

Martin, D.M., Fäldt, J. and Bohlmann, J. (2004) Functional Characterization of Nine Norway SpruceTPS Genes and Evolution of GymnospermTerpene Synthases of the TPS-d Subfamily. *Plant Physiol.*, **135**, 1908–1927.

Mehrshahi, P., Stefano, G., Andaloro, J.M., Brandizzi, F., Froehlich, J.E. and DellaPenna, D. (2013) Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 12126–31.

Meier, S., Tzfadia, O., Vallabhaneni, R., Gehring, C. and Wurtzel, E.T. (2011) A transcriptional analysis of carotenoid, chlorophyll and plastidial isoprenoid biosynthesis genes during development and osmotic stress responses in Arabidopsis thaliana. *BMC Syst. Biol.*, **5**, 77.

Moller, B.L. (2010) Dynamic Metabolons. *Science (80-.).*, **330**, 1328–1329.

Nisar, N., Li, L., Lu, S., Khin, N.C. and Pogson, B.J. (2015) Carotenoid metabolism in plants. *Mol. Plant*, **8**, 68–82.

Palmer, J.D., Soltis, D.E. and Chase, M.W. (2004) The plant tree of life: An overview and some points of view. *Am. J. Bot.*, **91**, 1437–1445.

Peters, R.J., Flory, J.E., Jetter, R., Ravn, M.M., Lee, H.J., Coates, R.M. and Croteau, R.B. (2000) Abietadiene synthase from grand fir (Abies grandis): Characterization and mechanism of action of the "pseudomature" recombinant enzyme. *Biochemistry*, **39**, 15592–15602.

Robert, S., Khalf, M., Goulet, M.-C., D'Aoust, M.-A., Sainsbury, F. and Michaud, D. (2013) Protection of recombinant mammalian antibodies from development-dependent proteolysis in leaves of Nicotiana benthamiana. *PLoS One*, **8**, e70203.

Schmidt, A. and Gershenzon, J. (2007) Cloning and characterization of isoprenyl diphosphate synthases with farnesyl diphosphate and geranylgeranyl diphosphate synthase activity from Norway spruce (Picea abies) and their relation to induced oleoresin formation. *Phytochemistry*, **68**, 2649–2659.

Sparkes, I.A., Runions, J., Kearns, A. and Hawes, C. (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.*, **1**, 2019–2025.

Tholl, D. and Lee, S. (2011) Terpene Specialized Metabolism in Arabidopsis thaliana. *Arab. B.*, **9**, e0143.

Troitsky, a. V., Melekhovets, Y.F., Rakhimova, G.M., Bobrova, V.K., Valiejo-Roman, K.M. and Antonov, a. S. (1991) Angiosperm origin and early stages of seed plant evolution deduced from rRNA sequence comparisons. *J. Mol. Evol.*, **32**, 253–261.

Vogel, B.S., Wildung, M.R., Vogel, G. and Croteau, R. (1996) Abietadiene Synthase from Grand Fir (Abies grandis). *J. Biol. Chem.*, **271**, 23262–23268.

Wroblewski, T., Tomczak, A. and Michelmore, R. (2005) Optimization of Agrobacterium-mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. *Plant Biotechnol. J.*, **3**, 259–73.

Zhou, K., Gao, Y., Hoy, J.A., Mann, F.M., Honzatko, R.B. and Peters, R.J. (2012) Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from Abies grandis. *J. Biol. Chem.*, **287**, 6840–6850.

Zi, J., Mafu, S. and Peters, R.J. (2014) To Gibberellins and Beyond! Surveying the Evolution of (Di)Terpenoid Metabolism. *Annu. Rev. Plant Biol.*, **65**, 259–286.

CHAPTER 4: Conclusions and Future Perspectives

Summary and Conclusions:

It has been increasingly clear that understanding how pathway intermediates are transported between different subcellular compartments in a biosynthetic pathway is not only important in gaining a comprehensive overview of plant metabolism, but also in aiding metabolic engineering efforts to increase flux towards the accumulation of high-value compounds. Plastids are central metabolic hubs within a plant cell, synthesizing a diverse array of compounds and exchanging pathway intermediates with other subcellular compartments including the endoplasmic reticulum (ER). With the characterization and identification of various plastid envelope transporter proteins, some transport processes, primarily the transport of polar metabolites, metals, ions and fatty acids/lipids are much better understood. However, how other non-polar compounds such as diterpenes, are transported from the plastid is still unknown. The goal of this dissertation is to shed more insight into how diterpene olefins are transported between the plastid and ER by utilizing a functional tool termed "transorganellar complementation" to express and complement two diterpene pathways, the gibberellin (GA) and diterpene resin acid (DRA) pathway.

In Chapter 2 of this dissertation, which described the transorganellar complementation of the GA pathway, the endogenous plastid-localized diterpene synthases, CPS and KS were retargeted into the ER lumen to create a pool of *ent*-kaurene in that compartment. This was done to assess whether this ER-synthesized *ent*-kaurene is accessible by the subsequent enzyme in the GA pathway, the OEM-localized KO. Since the mutants defective in the CPS and KS activities have dramatic dwarf phenotypes, complementation of the CPS and KS activities were presumed to be easily identified by the physiological appearance of the transgenic lines generated. This was indeed the case, as the *ga1-6* and *ga2-1* mutants complemented by the native plastid-targeted CPS-YFP and KS-mCherry protein, respectively, were phenotypically indistinguishable

from the WT (Figure 2.2, A and B). When both CPS and KS proteins were retargeted into the ER in the *ga1-6 ga2-1* double mutant background, the same WT appearance was observed, indicating that *ent*-kaurene synthesized in the ER was accessible by OEM-localized KO (Figure 2.2C).

Single targeting of CPS and KS into their corresponding mutant background was done as negative controls, as complementation in these cases would require movement of the polar intermediate, *ent*-CDP across the double plastid envelope membrane and the ER membrane, a transport activity that is not known in plants. This was indeed the case for ER-targeted KS-mCherry, as the transgenic lines remained dwarfs (Figure 2.2B). This indicated that *ent*-CDP is not transported from the plastid to the ER. Surprisingly, this was not the case with ER-targeted CPS-YFP, which complemented its *ga1-6* mutation and appeared phenotypically indistinguishable to WT. This indicated that *ent*-CDP generated in the ER were accessible by endogenous plastid-localized KS. These combined data point towards a mechanism for mediating the unidirectional transport of polar *ent*-CDP from the ER back into the plastid, which is likely mediated by transporters. Two independent techniques, namely confocal imaging and subcellular fractionation followed by immunoblotting, were used to confirm and corroborate the correct localization of ER-targeted CPS-YFP, which indicated that the physiological complementation of ER:CPS was not due to mislocalization of CPS protein (Figure 2.4 and 2.5).

Quantification of the GA metabolites in the CPS transgenic lines showed that in the plastid:CPS-YFP lines, the levels of *ent*-kaurene, GA_{12}, and GA_{24} correlated with the levels of CPS-YFP protein, consistent with the findings by Fleet *et al.* (2003). One of the plastid:CPS-YFP line, P3-CPS accumulated up to 200-fold higher levels of *ent*-kaurene, and up to 20-fold higher levels of GA_{12} and GA_{24} (Figure 2.6, B, C, and D). Conversely, overexpression of CPS

protein in the ER (E2-CPS and E3-CPS transgenic lines) only allowed for the accumulation of WT levels of GA metabolites (Figure 2.6, B, C, and D). The levels of GA metabolites quantified in the ER:CPS-YFP lines were thought to be limited by the availability of CPS substrate, GGDP. Therefore, ER-targeted GGDPS11-CFP was overexpressed in the E2-CPS and E3-CPS backgrounds in an attempt to increase the GGDP availability in that compartment and hence GA synthesis. However, the levels of *ent*-kaurene in the ER-targeted CPS:GGDPS double transgenic lines did not show any significant increases compared to the ER:CPS background (Figure 2.8), which indicated that the levels of upstream precursors such as IDP and DMADP, rather than GGDP were limiting.

In conclusion, the work in Chapter 2 showed that ER-synthesized *ent*-CDP is accessible by endogenous plastid-localized KS, pointing to a unidirectional transport mechanism for *ent*-CDP. In addition, OEM-localized KO could access ER-synthesized *ent*-kaurene, as evidenced by the WT phenotypes of the ER-CPS::ER-KS double transgenic lines.

In Chapter 3 of this dissertation, the *in vivo* transport of novel diterpene olefins of the DRA pathway in Arabidopsis was assessed. This was done by expressing conifer-specific enzymes in their native subcellular compartments. The bifunctional diterpene synthases, abietadiene synthase (AgAS) and isopimaradiene synthase (PsISO) were expressed in plastids, while the cytochrome P450, CYP720B4 was expressed in the ER to investigate whether an existing transport system in Arabidopsis can transport these novel diterpene olefins to the ER, thus generating the corresponding DRAs. In this work, metabolite analysis of pooled T2 plant materials from AgAs-YFP and PsISO-YFP showed very low levels of the diterpene olefin products were produced at levels 100- to 1000-fold lower than plastid CPS.

All three transgenic lines (AgAS-YFP, PsISO-YFP, and CYP720B4-YFP) showed no YFP signal detected in both confocal microscopy and immunoblotting analysis. In addition, the low levels of diterpene olefins detected in AgAS-YFP and PsISO-YFP indicated that there could be problems with the transgene expression. Transcript analysis showed that while full-length transcripts were expressed in the transgenic lines (Figure 3.5), 50% of all transcripts showed deletions from alternative splicing, which resulted in truncated YFP protein and partially explained the lack of YFP signal detected in confocal imaging and immunoblotting analysis. However, the remaining 50% of the transcripts were full-length and did not show alternative splicing, and should produce functional YFP protein. The most likely explanations for the undetectable level of YFP-fusion protein include cosuppression of the remaining full-length transcripts and inability of AgAS-YFP and PsISO-YFP produced to interact productively with the GGDP synthase in Arabidopsis. Although this work was unable to show the production of DRAs in the stable transgenic lines, a publication subsequent to this work was able to show production of isopimaric acid when ISO and CYP720B4 were transiently expressed in their native compartments in *N. benthamiana* (Gnanasekaran *et al.*, 2015). This indicated that *N. benthamiana*, which is an angiosperm, was able to transport non-endogenous diterpene olefin from the plastid to ER.

This dissertation was part of a larger effort by the laboratory to understand the operating principles at the plastid:ER interface by functionally determining the types of non-polar compounds that can be accessed across the plastid:ER interface *in vivo,* the directionality of this complementation and the classes of enzymes and pathways capable of transorganellar complementation. Through this work, ER-synthesized *ent*-kaurene was shown to be accessible by OEM-localized KO, indicating that the GA pathway can be complemented from a different

organelle. Complementation of the dwarf mutant phenotypes, in addition to the WT levels of GA intermediates quantified in the ER:CPS-YFP transgenic lines showed that both CPS and KS proteins are active in the ER lumen, indicating that soluble, monomeric diterpene synthases can catalyze their enzymatic reactions in a non-native organelle. However, more optimization is required in expressing conifer-specific bifunctional diterpene synthases such as AgAS and PsISO in Arabidopsis as they are shown to only produce low levels of their corresponding diterpene olefin products, even when they are expressed in their native organelle. Any future works that aim to express other diterpene synthases in the ER, however, need to address the limited availability of the GGDP precursor in that compartment.

Future perspectives:

Analysis of KS transgenic lines

In Chapter 2, we have shown the ER:CPS-YFP transgenic lines accumulated GA metabolites similar to WT levels. As CPS has been shown to be the "gatekeeper" in the synthesis of GAs (Silverstone *et al.*, 1997; Fleet *et al.*, 2003), we hypothesized that the ER-targeted CPS and KS in the *ga1ga2* mutant background would have levels of GA metabolites similar to the ER:CPS-YFP transgenic lines. As the ER-targeted KS-mCherry in the *ga2-1* mutant background failed to complement the dwarf mutant phenotype, we hypothesized that these transgenic lines would have levels of GA metabolites similar to the *ga2-1* mutant. Although physiological complementation gives a good indication to the levels of GA metabolites, quantification of the levels of GA metabolites such as *ent*-kaurene, $GA_{12,}$ and GA_{24} needs to be performed to support the phenotypes observed in these transgenic lines.

In addition, a second technique is needed to confirm and corroborate the subcellular localization of ER-targeted KS. The subcellular fractionation of whole Arabidopsis seedlings and protease protection assay of crude microsome fraction as performed in Chapter 2 of this dissertation will support the confocal imaging data that ER:KS-mCherry is ER-localized.

Coexpression analysis of metabolic pathways as a potential tool in identifying unknown proteins involved in intracellular polar diterpene intermediate transport.

In Chapter 2, *ent*-CDP, which is a polar compound, was shown to be transported in a unidirectional fashion from the ER into the plastid. One strategy could be used in potentially identifying these transporters in the plastid envelope membranes and ER membrane is the coexpression analysis of genes within a metabolic pathway. This strategy was successfully

employed in the identification of the plastidial glycolate glycerate translocator 1 (PLGG1) and tonoplast membrane-localized strictosidine exporter (Pick *et al.,* 2013; Payne *et al.,* 2017). Coexpression analysis showed that PLGG1 were coexpressed with all genes involved in the photorespiration pathway, and characterization of the *plgg1-1* mutant showed the PLGG1 transporter protein functions as a plastid-localized glycolate/glycerate transporter (Pick *et al.*, 2013). In an effort to identify candidate transporters, self-organizing maps of RNA-Seq datasets from *Catharanthus roseus* was used (Payne *et al.*, 2017). The strictosidine transporter was found to be coexpressed with the early monoterpene indole alkaloid pathway genes and further biochemical characterization showed the strictosidine transporter functions as a tonoplast-localized strictosidine exporter to the cytosol (Payne *et al.*, 2017). These two transporters described here are intracellular polar metabolite transporters, and this strategy may be useful in identifying candidate transporters for transporting *ent*-CDP or other prenyl diphosphates of different chain lengths.

REFERENCES

REFERENCES

Fleet, C.M., Yamaguchi, S., Hanada, A., Kawaide, H., David, C.J., Kamiya, Y. and Sun, T. (2003) Overexpression of AtCPS and AtKS in Arabidopsis Confers Increased ent -Kaurene Production But No Increase in Bioactive Gibberellins 1. *Plant Physiol.*, **132**, 830–839.

Gnanasekaran, T., Vavitsas, K., Andersen-Ranberg, J., Nielsen, A.Z., Olsen, C.E., Hamberger, B. and Jensen, P.E. (2015) Heterologous expression of the isopimaric acid pathway in Nicotiana benthamiana and the effect of N-terminal modifications of the involved cytochrome P450 enzyme. *J. Biol. Eng.*, **9**, 24.

Payne, R.M.E., Xu, D., Foureau, E., et al. (2017) An NPF transporter exports a central monoterpene indole alkaloid intermediate from the vacuole. *Nat. Plants*, **3**, 1–9.

Pick, T.R., Bräutigam, A., Schulz, M.A., Obata, T., Fernie, A.R. and Weber, A.P.M. (2013) PLGG1, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a unique class of metabolite transporters. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 3185–90.

Silverstone, A.L., Chang, C., Krol, E. and Sun, T.P. (1997) Developmental regulation of the gibberellin biosynthetic gene GA1 in Arabidopsis thaliana. *Plant J.*, **12**, 9–19.