2017

TESIS DOCTORAL

Biosynthesis of Phenylalanine in Plants

JORGE EL-AZAZ CIUDAD

Biosynthesis of Phenylalanine in Plants

TESIS DOCTORAL
JORGE EL-AZAZ CIUDAD

PROGRAMA DE DOCTORADO
BIOLOGÍA CELULAR Y MOLECULAR
Director: Francisco M. Cánovas Ramos
Co-director: Fernando N. de la Torre Fazio

Facultad de Ciencias
Universidad de Málaga
2017

AUTOR: Jorge El-Azaz Ciudad

http://orcid.org/0000-0001-7452-1018

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga

DEPARTAMENTO DE BIOLOGÍA MOLECULAR Y BIOQUÍMICA

FACULTAD DE CIENCIAS

TESIS DOCTORAL

Biosynthesis of Phenylalanine in Plants

Jorge El-Azaz Ciudad

2017

DEPARTAMENTO DE BIOLOGÍA MOLECULAR Y BIOQUÍMICA

FACULTAD DE CIENCIAS

El Prof. Francisco M. Cánovas Ramos, Catedrático del Departamento Biología Molecular y Bioquímica de la Facultad de Ciencias de la Universidad de Málaga, y el Dr. Fernando N. de la Torre Fazio, investigador de la Universidad de Málaga, CERTIFICAN:

Que D. Jorge El-Azaz Ciudad ha realizado bajo nuestra dirección, en el Departamento de Biología Molecular y Bioquímica de la Facultad de Ciencias de la Universidad de Málaga, el trabajo de investigación recogido en esta memoria de tesis doctoral, con el título original "Biosynthesis of Phenylalanine in Plants".

Tras la revisión de la presente memoria, se ha estimado oportuna su presentación ante la Comisión de Evaluación correspondiente, por lo que autorizamos su exposición y defensa para optar al grado de Doctor en Biología.

Y para que así conste, firmamos el presente certificado en Málaga, a 4 de JULIO de 2017:

Prof. Dr. Francisco M. Cánovas Ramos

Dr. Fernando N. de la Torre Fazio

Agradecimientos

Esta tesis doctoral ha sido posible gracias a la dirección del profesor Francisco M. Cánovas, que lidera desde hace más de treinta años el grupo de investigación en metabolismo del nitrógeno en plantas en la Universidad de Málaga, quien me brindó la oportunidad de incorporarme a su grupo en septiembre de 2011 como estudiante de máster y empezar a trabajar en lo que se ha convertido en esta tesis doctoral. Por otro lado, Fernando de la Torre, investigador post-doctoral en este mismo grupo, ha desempeñado como co-director una labor indispensable para que este trabajo se haya podido concluir satisfactoriamente, y es a él a quien siempre deberé la base de mi formación en el laboratorio en muchas jornadas de trabajo codo con codo.

Este texto refleja el resultado de los esfuerzos de ambos, director y co-director, por hacer avanzar mi formación y ampliar, en un sentido más genérico, los horizontes del grupo de inves tigación con la incorporación de nuevas líneas de trabajo.

Igualmente fundamental ha resultado el poder contar con los medios materiales necesarios para la realización de los experimentos que aquí se recogen, gracias al respaldo económico de los proyectos concedidos por el Ministerio de Economía, Industria y Competitividad (BIO2015-69285-R) y la Junta Andalucía (BIO2012-0474), y así mismo, a la beca de Formación de Profesorado Universitario de la que he sido beneficiario estos años.

Quisiera también dedicar unas palabras a todas las personas de nuestro laboratorio, compañeros y amigos, que han estado presentes durante estos años de mi formación y me han prestado su ayuda, colaboración y apoyo tanto en los ámbitos profesional como personal: Belén, Blanca, Canales, Cañas, Concha, Marina, Mayte y Vanessa. ¡Gracias a todos vosotros!

Table of Contents

I. SUMMARY
RESUMEN EN ESPAÑOL

1. Introducción

El aminoácido fenilalanina (Phe) desempeña en las plantas terrestres una función esencial, al actuar como precursor tanto de las proteínas como de la síntesis de fenilpropanoides, una amplia familia de metabolitos secundarios que cumplen funciones muy diversas y cuya aparición y diversificación está interrelacionada con la propia evolución de las plantas terrestres. La importancia de la síntesis de Phe es tal que se estima que más del 30% del CO_2 fijado por las plantas en la fotosíntesis es finalmente derivado hacia la síntesis de este aminoácido, y de ahí hacia la biosíntesis de fenilpropanoides (Boerjan *et ál*., 2003), más particularmente ligninas, uno de los componentes fundamentales de las paredes celulares secundarias de las plantas.

La biosíntesis de los aminoácidos aromáticos se produce en las plantas a partir del corismato, el producto final de la ruta metabólica conocida como ruta del siquimato (**Figura 1**). Esta ruta, que se desarrolla en el interior de los cloroplastos, parte de los metabolitos fosfoenol-piruvato (PEP) y D-eritrosa-4-fosfato (E4P). Ambos metabolitos son intermediarios, de forma respectiva, de la glucólisis y de la ruta de las pentosas fosfato; de esta forma, la ruta del siquimato conecta el metabolismo central del carbono con la biosíntesis de aminoácidos aromáticos y el metabolismo secundario que depende de estos aminoácidos. El corismato producido en la ruta del siquimato actua como precursor de los tres aminoácidos aromáticos: triptófano (Trp), tirosina (Tyr) y Phe. El Trp se produce a partir de seis reacciones enzimáticas adicionales en una ruta metabólica propia, mientras que en la biosíntesis de Tyr y Phe intervienen tres reacciones adicionales de la ruta del arogenato o la del fenilpiruvato/4-hidroxifenilpiruvato.

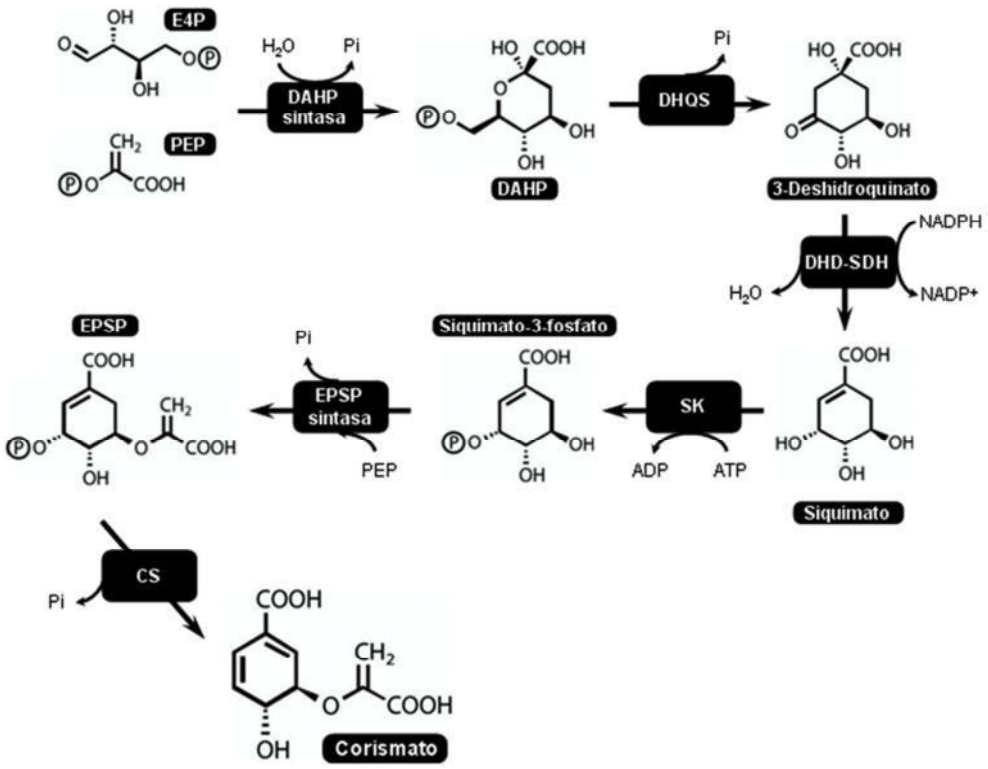


Figura 1. **Esquema de la ruta del siquimato**. La biosíntesis de corismato se produce a partir de una D-eritrosa-4-fosfato (E4P) y dos moléculas de fosfo*enol*-piruvato (PEP). A partir del corismato se sintetiza, en la ruta del prefenato, los aminoácidos fenilalanina y tirosina, mientras que el triptófano es sintetizado en una ruta propia independiente. DAHP, 3-Deoxy-D-arabino-heptulosonato-7-fosfato; DHQS, deshidroquinato sintasa; DHD-SDH, deshidroquinato deshidratasa-siquimato deshidrogenasa; SK, siquimato kinasa; EPSP, enol-piruvil-siquimato-3-fosfato; CS, corismato sintasa. Adaptado a partir de Maeda y Dudareva, (2012).

Los aminoácidos Tyr y Phe son sintetizados en una ruta metabólica propia diferenciada que es conocida como ruta del prefenato (**Figura 2**), la cual parte del corismato, que como ya se ha dicho es el producto final de la ruta del siquimato. En el primer paso de esta ruta, el corismato es transformado en prefenato por la enzima corismato mutasa (CM). El prefenato puede ser entonces transformado en Phe a través de dos sub-rutas alternativas de la ruta del prefenato:

- Ruta del arogenato. Considerada actualmente como la cuantitativamente más importante para las plantas (Maeda *et ál*., 2010), en ella el prefenato es primero transformado en arogenato a través de la enzima prefenato aminotransferasa (PAT). A continuación, el arogenato es transformado en Phe mediante la actividad arogenato deshidratasa.

- Ruta del fenilpiruvato. Es la ruta de biosíntesis de Phe característica de la mayoría de microorganismos, así como de hongos. En esta ruta, el prefenato es descarboxilado a fenilpiruvato por la enzima prefenato deshidratasa (PDT). Con posterioridad, este último compuesto sufre una transaminación que lo transforma en Phe, la cual está catalizada por transaminasas de aminoácidos aromáticos.

De forma análoga, la Tyr puede ser sintetizada a través de dos vías alternativas que parten del prefenato (**Figura 2**): su conversión en arogenato por la enzima PAT, y posterior descarboxilación a Tyr mediante la enzima arogenato deshidrogenasa (ADH); o bien, de manera alternativa, mediante la descarboxilación del prefenato a 4-hidroxifenilpiruvato mediante la enzima prefenato deshidrogenasa (PDH) y después la transaminación de este compuesto a Tyr mediante una aminotransferasa de aminoácidos aromáticos (AAA).

Como se ha mencionado en el párrafo anterior, se considera que en las plantas la Phe es sintetizada principalmente a partir de arogenato, síntesis que es dependiente de la actividad enzimática ADT. La familia de enzimas ADT presenta un elevado número de isoformas en todas las plantas, lo que podría estar relacionado con la regulación de la síntesis de Phe, y más particularmente la regulación del uso que se le va a dar este aminoácido (síntesis de proteínas frente a síntesis de fenilpropanoides). De esta forma, en *Arabidopsis thaliana* se han descrito por ejemplo un total de seis miembros para esta familia (Cho *et ál*., 2007). Aunque la ruta del arogenato es la principal ruta de biosíntesis de Phe, recientemente se ha propuesto que la ruta del fenilpiruvato podría funcionar como una ruta alternativa para la biosíntesis de Phe en plantas (Yoo *et ál*., 2013). La ruta del fenilpiruvato requiere la existencia de enzimas con actividad PDT en plantas. Diferentes trabajos apuntan a que dicha actividad enzimática es realizada por las mismas proteínas que las ADT. Así, se ha descrito que en *Arabidopsis thaliana* dos de las seis enzimas ADT son bifuncionales, presentando tanto actividad ADT como PDT *in vitro* (Cho *et ál*., 2007), lo que posibilitaría la síntesis de Phe no dependiente de arogenato. Estas enzimas bifuncionales ADT/PDT podrían hacer posible la síntesis de Phe a través de dos rutas alternativas pero no necesariamente excluyentes, lo que podría ser un factor adicional en la regulación de la síntesis de este aminoácido y su utilización final.

Uno de los dos posibles destinos metabólicos de la Phe es, aparte de la síntesis de proteínas, la biosíntesis de fenilpropanoides. Estos compuestos forman un amplísimo y diverso grupo de metabolitos formados a partir de la polimerización de monómeros compuestos por un anillo fenilo con una cadena lateral de 3C de longitud. Aunque sean englobados como "metabolitos secundarios", son fundamentales para el normal desarrollo y

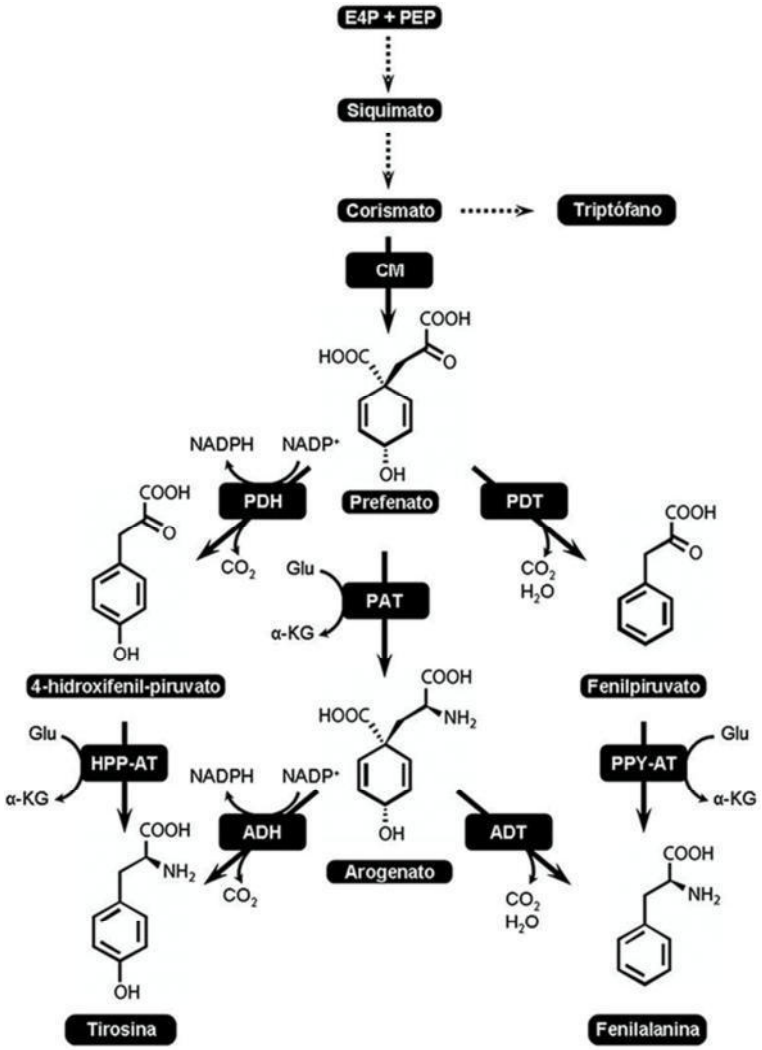


Figura 2. **Esquema de la ruta de biosíntesis de los aminoácidos triptófano, tirosina y fenilalanina en plantas**. Los pasos en los que están implicadas múltiples reacciones enzimáticas se han indicado mediante líneas discontinuas. PEP, fosfoenolpiruvato; E4P, eritrosa-4-fosfato; α-KG, α-cetoglutarato; Glu, glutamato; CM, corismato mutasa; PDH, prefenato deshidrogenasa; PDT, prefenato deshidratasa; PAT prefenato-arogenato aminotransferasa; HPP-AT, hidroxifenilpiruvato aminotransferasa; PPY-AT, fenilpiruvato aminotransferasa; ADH, arogenato deshidrogenasa; ADT, arogenato deshidratasa. Adaptado a partir de Maeda *et ál*., 2011.

supervivencia de las plantas (Fraser y Chapple, 2011). Los fenilpropanoides derivan en su inmensa mayoría del aminoácido Phe, aunque en algunas plantas también de la Tyr (Heldt y Heldt, 2011).

El primer paso en la ruta de biosíntesis de fenilpropanoides está catalizado por la enzima fenilalanina amonio-liasa (PAL), que genera ácido trans-cinámico y amonio. La enzima PAL es considerada como una de las más estudiadas a nivel de regulación en el metabolismo secundario de las plantas (Heldt y Heldt, 2011). Posteriormente el ácido cinámico será sustrato de una serie de reacciones adicionales que dan lugar a varios derivados de los que parten las rutas específicas de biosíntesis de los diferentes grupos de fenilpropanoides. La síntesis de los fenilpropanoides se basa en la modificación química y polimerización de estos monómeros básicos, y ocasionalmente también en la incorporación de otras moléculas, como malonil-CoA.

El tipo de fenilpropanoide más abundante en términos cuantitativos en las plantas vasculares es el de las ligninas, un componente fundamental de la pared celular secundaria de las células vegetales, a las que confiere resistencia mecánica y vuelve impermeables, características que tienen una gran importancia funcional en los tejidos vasculares de las plantas. Se estima que a nivel de la biosfera las ligninas suponen en torno a un 30% del carbono total en forma orgánica, lo que las convierte en el segundo biopolímero por importancia cuantitativa, únicamente por detrás de la celulosa (Boerjan *et ál*., 2003). La lignina es un polímero ramificado altamente completo que es formado a partir de una serie de fenilpropanoides que actúan como monómeros. La ruta específica de biosíntesis de estos monómeros parte del ácido p-cumárico y el cumaroil-CoA, que en una compleja sucesión de reacciones enzimáticas darán lugar

a los alcoholes cumarilo, coniferilo y sinapilo. Estos alcoholes son generalmente agrupados bajo la denominación común de monolignoles (Heldt y Heldt, 2011). La polimerización de las ligninas tiene lugar en el exterior de la célula vegetal a partir de la oxidación de monolignoles mediante actividades tipo peroxidasa y lacasa, que generan radicales fenólicos que reaccionan entre si en rondas sucesivas de polimerización, pudiendo de esta forma dar lugar a una gran diversidad de estructuras diferentes. En las gimnospermas, las ligninas están formadas a partir de los monolignoles coniferilo y cumarilo, mientras que en las angiospermas incorporan también el monómero sinapilo, inexistente en las ligninas de gimnospermas (Lewis y Yamamoto, 1990).

Aparte de su importancia fisiológica en las plantas, la importancia industrial y biotecnológica de las ligninas es indiscutible. Estos compuestos han recibido particular atención en el último par de décadas al estar estrechamente relacionados con la producción de bio-combustibles de segunda generación, es decir, aquellos que son producidos a partir de la celulosa de los tejidos vegetales (Mottiar *et ál*., 2016). El aprovechamiento de esta fuente de energía no es actualmente viable a gran escala como consecuencia de que, en las paredes celulares secundarias, la celulosa se presenta químicamente vinculada con ligninas, haciendo de la extracción de celulosa un proceso costoso y altamente dañino para el medio ambiente. El caso particular de los bio-combustibles de segunda generación ejemplifica el interés que, también en investigación aplicada, puede tener profundizar en el estudio de la síntesis de aminoácidos aromáticos y sus derivados.

2. Objetivos

El objetivo general de esta tesis es obtener un mejor conocimiento la biosíntesis de fenilalanina en las plantas y su regulación en relación a la biosíntesis de metabolitos secundarios. Los objetivos específicos son:

1. **Determinar si la fenilalanina puede ser sintetizada en las plantas a través de la ruta del fenilpiruvato**. Hasta la fecha, hay pocas evidencias fisiológicas al respecto de la participación de la ruta del fenilpiruvato en la biosíntesis de fenilalanina en plantas. La posibilidad de que este aminoácido sea sintetizado a través de dos rutas alternativas podría proveer a las plantas de un nivel adicional de regulación, lo que podría desempeñar un papel importante en la disponibilidad del fenilalanina para su uso en el metabolismo secundario.

2. **Identificar las bases moleculares subyacentes al hecho de que ciertas arogenato deshidratasas presenten también actividad prefenato deshidratasa**. Con anterioridad, varios autores han demostrado que algunas arogenato deshidratasas de plantas también pueden actuar como prefenato deshidratasas *in vitro*. Sin embargo, actualmente se desconocen las bases mecanísticas o estructurales detrás de ello.

3. **Arrojar luz sobre la regulación de la biosíntesis de fenilalanina, más particularmente en coníferas, empleando *P. pinaster* como especie modelo**. Las coníferas presentan un metabolismo secundario especialmente complejo, y poseen pues un particular interés en el estudio de la regulación de estas rutas metabólicas. Como plantas

leñosas, las coníferas derivan además grandes cantidades de carbono orgánico hacia la síntesis de fenilpropanoides, principalmente ligninas, un componente esencial de los tejidos vasculares. Las arogenato deshidratasas son, por su posición en el metabolismo biosintético de la fenilalanina y la gran variedad de isoformas encontradas en las plantas, buenas candidatas para actuar como enzimas reguladoras en este proceso fisiológico.

3. Resultados y discusion

Al inicio de esta tesis doctoral, existía una cantidad de información realmente limitada sobre la ruta del prefenato y los mecanismos a través de los cuales las plantas regulan la biosíntesis de Tyr y Phe. Este hecho contrastaba especialmente con la ruta de biosíntesis de Trp, a cuyo estudio se había dedicado mayor atención en los últimos años, particularmente por la relación existente entre la biosíntesis de Trp y las auxinas (Maeda y Dudareva, 2012). Durante esta tesis, se ha generado un conjunto importante de resultados, tanto en *Pinus pinaster* como en las plantas modelo *Arabidopsis thaliana* y *Nicotiana benthamiana*, que consideramos han supuesto una contribución importante al campo estudiado. Testigo de ello son las publicaciones científicas, en revistas internacionales de alto impacto en el campo de estudio, que han surgido del trabajo de investigación básica realizado durante esta tesis doctoral. En consecuencia, se ha optado por presentar esta tesis doctoral de acuerdo al modelo de compendio de publicaciones, incluyéndose un total de tres artículos científicos ya publicados y otros tres en formato de borrador, con la intención de que en breve sean sometidos a revisión en revistas internacionales del área de biología molecular y bioquímica de plantas. A continuación se presentarán y discutirán, por orden de inclusión en la tesis, los resultados más importantes que se desprenden de este conjunto de artículos científicos.

En el artículo número 1, publicado originalmente con el título *Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism*, se han descrito los efectos fisiológicos y bioquímicos provocados por el silenciamiento de los genes *NbPAT* y *NbAsp5*, empleándose para ello una estrategia basada en VIGS (acrónimo del

inglés *Virus Induced Gene Silencing*) en hojas de *Nicotiana benthamiana*. Los plastos de las células vegetales poseen dos enzimas con actividad aspartato aminotransferasa (AAT): una de tipo eucariota, Asp5, y otra de tipo procariota, la cual también presenta actividad PAT. En este artículo, se han identificado la familia al completo de enzimas con actividad AAT de *N. benthamiana*, y se han aislado y clonado sendos genes codificantes de enzimas AAT de localización plastidial: *NbAsp5* y *NbPAT*. Mediante una aproximación basada en VIGS, se han generado plantas de *N. benthamiana* silenciadas para los genes *NbAsp5* y/o *NbPAT*, las cuales han sido analizadas fenotípica y bioquímicamente con el fin de determinar la participación de sendas enzimas en la biosíntesis plastidial de aminoácidos. Las plantas silenciadas para *NbAsp5* no mostraron cambios en su fenotipo, presentando niveles similares de aspartato y glutamato libre a los encontrados en las plantas control. No obstante, se encontró una disminución significativa en los niveles del aminoácido asparaguina, y un fuerte incremento en los niveles de lisina. En contraposición, la supresión de la expresión de *NbPAT* ocasionó una severa reducción en el crecimiento de las plantas, las cuales mostraron evidentes signos de clorosis. Estas últimas plantas también mostraron unos niveles de asparagina extremadamente bajos y, más especialmente, fuertes alteraciones en el metabolismo de la Phe y en la cantidad de ligninina depositada en los tallos, que se vio drásticamente reducida. Por otro lado, el silenciamiento génico de *NbPAT* también ocasionó en estas plantas la sobreexpresión de *NbPDT1* y *NbPDT2*, genes putativamente codificantes de enzimas con actividad PDT, así como de *NbPPAT*, que codifica una aminotransferasa de aminoácidos aromáticos. La sobreexpresión de *NbPDT1*, *NbPDT2* y *NbPPAT* en respuesta al silenciamiento de *NbPAT* sugiere que, en respuesta al bloqueo de la ruta del arogenato, se está produciendo una re-programación transcripcional de la biosíntesis de Phe

que permitiría sintetizar este aminoácido a través de la ruta del fenilpiruvato. En conclusión, los resultados obtenidos indican que la enzima PAT presenta una función que es al menos parcialmente redundante con Asp5 en relación a la biosíntesis de aspartato y sus aminoácidos derivados pero, por otro lado, que desempeña una función esencial en la biosíntesis de Phe, y por ende también en relación al metabolismo de los fenilpropanoides.

En el artículo número 2, publicado bajo el título *Identification of a small protein domain present in all plant lineages that confers high prephenate dehydratase activity*, se ha llevado a la cabo la caracterización molecular de la familia de enzimas con actividad ADT/PDT de *P. pinaster*. Como se ha mencionado en la introducción, la biosíntesis de Phe en las plantas puede producirse a través de dos vías alternativas de la ruta del prefenato. Una de estas vías, que es considerada como la cuantitativamente más importante, es dependiente de arogenato. En ella, el prefenato es transformado en arogenato mediante la enzima PAT. Posteriormente, el arogenato es descarboxilado por la enzima ADT, produciendo Phe. En la otra vía para sintetizar Phe, que es conocida como ruta del fenilpiruvato, el prefenato es primero descarboxilado a fenilpiruvato mediante la enzima PDT. El fenilpiruvato producido es entonces sustrato de una reacción de transaminación catalizada por transaminasas de aminoácidos aromáticos (AAA), en la cual se produce la Phe. La vía del fenilpiruvato ha sido considerada como característica de microorganismos y hongos, pero sin embargo en los últimos años varias publicaciones han aportado resultados que sugieren que esta vía de síntesis de Phe podría ser funcional en las plantas. Así, varias enzimas de plantas con actividad ADT poseen también actividad PDT *in vitro* (Cho *et ál*., 2007). Así mismo, la supresión de la actividad PAT induce la expresión de genes con actividad PDT y AAA que

podrían estar involucrados en la ruta del fenilpiruvato, como se ha mostrado en el primer artículo de esta tesis, observación que también ha sido realizada por otros autores (Yoo *et ál*., 2013). Así, la identificación de genes que pudiesen estar involucrados en esta hipotética ruta alternativa para sintetizar Phe es pues un motivo de interés. Con objeto de identificar genes codificantes de enzimas con actividad PDT en *P. pinaster*, se han identificado y clonado un total de 9 genes pertenecientes a la familia de enzimas ADT. Mediante análisis filogenético y complementación funcional en levaduras, hemos identificado la existencia de al menos 2 genes de *P. pinaster* codificantes de proteínas bifuncionales con actividad ADT/PDT. El análisis comparativo de estas secuencias, conjuntamente con experimentos de mutagénesis dirigida, ha permitido identificar un dominio de 22 aminoácidos en la región C-terminal que es esencial para una actividad PDT eficiente. La supresión de este motivo lleva a las enzimas con actividad PDT a una drástica reducción en su eficiencia catalítica y a un incremento de en torno a 10 veces en la K_m aparente por el sustrato de la reacción (prefenato). Más particularmente, dentro de este dominio, se ha identificado el residuo Ala^{314} como esencial para una actividad PDT eficiente. Hemos denominado PAC (*PDT Activity Conferring*) a este dominio. Los análisis filogenéticos realizados han mostrado que pueden encontrarse enzimas ADT con el dominio PAC en todos los linajes de plantas terrestres, además de en algas verdes, algas rojas y glaucofitas, los tres linajes evolutivos que surgieron de la adquisición de los cloroplastos. Estos resultados sugieren que la actividad PDT, y por tanto la capacidad de sintetizar Phe a través de fenilpiruvato, se ha conservado durante la evolución de las plantas y sus ancestros. Se sugiere también que la posibilidad de sintetizar Phe a través de dos rutas diferentes podría haber desempeñado una importante función en la evolución y regulación del metabolismo secundario en plantas.

En el artículo número 3, con el título *New insights into the metabolic regulation of phenylalanine biosynthesis in plants*, se propone una hipótesis explicativa para los resultados observados en el artículo precedente. De acuerdo con los resultados que se han publicado en el artículo 2, la presencia del dominio PAC confiere a las enzimas bifuncionales ADT/PDT una afinidad aparente por el prefenato muy superior a las proteínas en que este dominio no está presente, así como una mayor constante de especificidad (K_{cat}/K_m). Las proteínas de la familia ADT poseen dos dominios funcionales claramente diferenciados. Mientras que la región amino-terminal está constituida por el dominio catalítico de la enzima, la región carboxilo-terminal tiene funciones reguladoras mediante un mecanismo alostérico de respuesta a metabolitos. Este dominio regulador carboxilo-terminal es denominado ACT, y se encuentra también en otras enzimas relacionadas con la biosíntesis de aminoácidos. Curiosamente, el dominio PAC, que como se ha visto es crítico para la actividad PDT, se encuentra ubicado en el dominio regulador ACT, y no en la región catalítica. El dominio ACT ha sido descrito previamente en relación a un mecanismo alostérico de retroalimentación negativa en respuesta al producto de la reacción que catalizan, la Phe. De esta forma, este aminoácido es reconocido en un sitio de unión específico en el dominio ACT y provoca un cambio de conformación en la estructura terciaria de la proteína, haciendo que el dominio catalítico pase de una configuración activa o abierta, a una inactiva o cerrada. Mediante el modelado tridimensional *in silico* de PpADT-G, usando la estructura cristalina de la proteína ADT/PDT de *C. tepidum* como modelo, se ha determinado que el dominio PAC de PpADT-G forma parte del bolsillo donde la Phe es reconocida en el dominio ACT de la proteína. Así mismo, el residuo Ala^{314}, que se ha demostrado como crítico para la actividad PDT, se encuentra ubicado entre dos motivos altamente conservados (GVLF y ESRP) que

están involucrados en el reconocimiento directo de la Phe en este bolsillo. Estas observaciones sugieren que la relación entre el dominio PAC y la actividad PDT estaría relacionada con un cambio en la conformación terciaria de la proteína, que afectaría al dominio catalítico. Dado el hecho de que las enzimas PDT de plantas estudiadas presentan actividad *in vitro* con la simple presencia del sustrato de la reacción, el prefenato, planteamos la hipótesis de que el prefenato, además de ser sustrato de la reacción, podría ser reconocido en el dominio ACT y actuar como un activador alostérico de la propia reacción de la que es sustrato. De acuerdo con esta hipótesis, el dominio PAC sería esencial para el reconocimiento alostérico del prefenato, determinando con ello la eficiencia de la reacción PDT. La confirmación de esta hipótesis que en un futuro se lleven a término experimentos de cristalización y determinación de la estructura tridimensional de las enzimas ADT/PDT de plantas.

En el artículo número 4, titulado *Identification and transcriptional regulation of arogenate dehydratase genes linked to lignin biosynthesis in maritime pine*, se ha descrito la relación existente entre la expresión de genes codificantes para enzimas PDT de *P. pinaster* y la formación de madera de compresión en esta especie. La madera de compresión es un tipo particular de tejido vascular de coníferas que se forma en los tallos o ramas que han perdido la verticalidad, más concretamente en el lado que ha quedado orientado hacia el suelo, ayudándole a recuperar la verticalidad perdida. Este tipo especializado de madera está enriquecido en ligninas (Villalobos *et ál*., 2012). Las ligninas, como fenilpropanoides, son metabolitos que derivan de la Phe, por lo que durante la formación de la madera de compresión hay elevada demanda de este aminoácido. Partiendo de una librería substractiva de cDNA de madera de compresión de ejemplares adultos de *P. pinaster* (cortesía del Prof. Francisco J. Ruiz

Cantón, Universidad de Málaga) se han determinado mediante qPCR los niveles de expresión de la familia *ADT* de *P. pinaster* en este tejido, encontrándose una activación generalizada de la transcripción de genes *ADT*. En un trabajo previo publicado por nuestro grupo de investigación (Craven-Bartle *et ál*., 2013) se demostró que el factor de transcripción PpMYB8 actúa como un activador general de la biosíntesis de Phe y ligninas en *P. pinaster*. El análisis *in silico* de las regiones promotoras de los nueve genes *ADT* incluidos en el estudio mostró la presencia de cajas de unión AC para PpMYB8, sugiriendo que los genes *ADT* podrían estar controlados por este factor de transcripción. Para probar esta hipótesis, se silenció la expresión de este factor de transcripción en plantas de *P. pinaster*. En las plantas silenciadas se encontró que la expresión de *PpADT-A*, *PpADT-D* y *PpADT-I* disminuye significativamente, lo que sugiere que estos genes de la familia *ADT* están controlados por PpMYB8. La interacción de PpMYB8 con las regiones reguladoras de *PpADT-A* y *PpADT-D* se analizó *in vitro* mediante la técnica de retraso de banda en gel (EMSA), demostrándose la existencia de una interacción física entre el factor de transcripción y las regiones analizadas. Adicionalmente, para el caso de *PpADT-A* se confirmó la interacción directa *in vivo* entre PpMYB8 y la región reguladora de este gen mediante un ensayo de actividad β-galactosidasa en *S. cerevisiae* W303. En conjunto, estos resultados ponen de manifiesto el papel de PpMYB8 como regulador de la biosíntesis de Phe en un contexto fisiológico en el que la biosíntesis de ligninas está fuertemente activada.

En el artículo número 5, *Biosynthesis and metabolic fate of phenylalanine in conifers*, se ha revisado el estado actual de la cuestión en lo que respecta al metabolismo de la Phe y su regulación en coníferas. Este grupo de plantas presenta un metabolismo secundario especialmente

diverso (Warren *et ál*., 2015). Al mismo tiempo, al tratarse de plantas leñosas, acumulan grandes cantidades de fenilpropanoides en forma de ligninas durante ciclos de vida muy largos. Ello, sumado a su indudable importancia forestal, convierte a las coníferas en un modelo interesante para el estudio del metabolismo secundario y su regulación. Como ya se ha mencionado a lo largo de este resumen, la Phe desempeña un papel crítico en las plantas como precursor de los fenilpropanoides, conectado el metabolismo primario con el secundario y siendo esencial en el desarrollo y crecimiento de las plantas, así como en numerosos mecanismos de respuesta a múltiples estreses bióticos y abióticos. Como punto de partida de la ruta de biosíntesis de fenilpropanoides, grandes cantidades de Phe deben ser sintetizadas en las plantas en momentos concretos de su desarrollo, en un proceso que debe estar finamente regulado. Al mismo tiempo, se repasa la importancia que el reciclaje del amonio (NH_4^+) tiene en la biosíntesis de fenilpropanoides durante el desarrollo de los tejidos vasculares: la primera enzima de la ruta de biosíntesis de los fenilpropanoides, la fenilalanina amonio-liasa, libera el grupo NH_4^+ de la Phe, que debe ser re-asimilado en nuevos aminoácidos para evitar la pérdida masiva de N en el proceso. En esta revisión se discuten también las novedades en el campo de la investigación de la función de los fenilpropanoides en los mecanismos de respuesta a estrés biótico y abiótico en coníferas, así como las iniciativas de manipulación genética de la ruta que se han llevado a término en los últimos años.

Finalmente, en el artículo número 6, titulado *A member of the arogenate deshydratase gene family (ADT2) is essential for female gametophyte development in* Arabidopsis**,** se presenta el trabajo de caracterización y complementación funcional del mutante *adt2* de *Arabidopsis thaliana*. Publicaciones previas (Corea *et ál*., 2012a; 2012b) han demostrado que,

en *A. thaliana*, los genes *ADT1*, *ADT3*, *ADT4*, *ADT5* y *ADT6* presentan funciones parcialmente solapantes en el crecimiento y desarrollo de los tejidos vasculares. Así, los citados autores generaron mutantes homocigotos para estos genes individualmente, y mutante múltiples, generados mediante el cruzamiento de los primeros. Las plantas mutantes en los genes *ADT* presentaban menor acumulación de ligninas en los tallos, fenotipo que se acentuaba en los mutantes múltiples. Sin embargo, los autores no pudieron encontrar un mutante de inserción homocigoto para el gen *ADT2*, sugiriendo que este gen podría desempeñar una función esencial en el desarrollo de la planta. En el presente estudio, hemos analizado dos líneas mutantes independientes para *ADT2*: SALKseq_044042 (*adt2-1*) y SALKseq_081342 (*adt2-2*). Del análisis de estos mutantes se ha deducido que la mutación *adt2* es letal en homocigosis: en un total de 213 plantas genotipadas para la línea mutante *adt2-1* y 279 para *adt2-2*, no se detectó ningún homocigoto, confirmando las observaciones de Corea *et ál*., (2012a). La proporción entre plantas con genotipo silvestre y mutantes heterocigotos se ajustó a una segregación distorsionada 1:1, característica de genes esenciales para el desarrollo del gametofito. Los mutantes heterocigotos *adt2* presentan silicuas más cortas de lo habitual, con una dotación de semillas aproximadamente reducida a la mitad, lo que se relaciona con un fenotipo de semi-esterilidad femenina. La expresión transgénica de *ADT2* sobre el fondo mutante heterocigoto rescató el fenotipo mutante en la línea *adt2-1*, confirmado que la mutación *adt2* es la causante del fenotipo observado. La expresión heteróloga de *PHA2*, que codifica una PDT de levaduras sin actividad ADT (Bross *et ál*., 2011), y de *ADT3*, una ADT estricta de Arabidopsis, ambos bajo el control del promotor de *ADT2*, rescató también el fenotipo mutante. Estos resultados marcan una clara diferenciación entre *ADT2* y los cinco genes restantes *ADT* de *A. thaliana*, al ser *ADT2* el único de ellos para el que

hasta el momento se ha demostrado una función esencial en el ciclo de vida de la planta. Al mismo tiempo, se ha demostrado que *PHA2* puede rescatar este fenotipo, mostrando que la ruta del fenilpiruvato es operativa en el órgano estudiado. El hecho de que *ADT3* también pueda revertir el fenotipo *adt2* evidencia que la Phe puede ser sintetizada en este tejido también por la vía del arogenato, sugiriendo que la enzima ADT2 pueda estar actuando de manera bifuncional y que sendas rutas de biosíntesis coexistan simultáneamente. En conclusión, nuestros resultados subrayan la importancia del metabolismo de la Phe para el desarrollo embrionario, ya sea para la biosíntesis de proteínas o para la de fenilpropanoides, aportándose nuevas evidencias al respecto de la coexistencia de dos rutas alternativas y no excluyentes para la biosíntesis de Phe.

4. Conclusiones

1. La enzima bifuncional aspartato aminotransferasa / prefenato aminotransferasa es esencial para el crecimiento y desarrollo de las plantas, desempeñando en los plastos una función esencial en la homeostasis de la biosíntesis de aminoácidos.

2. Se ha identificado un dominio proteico de 22 aminoácidos de longitud, al que se ha denominado PAC, que confiere alta actividad prefenato deshidratasa a las enzimas de la familia arogenato deshidratasa. Este dominio se encuentra tanto en plantas terrestres como en las algas de las que son descendientes, sugiriendo que dicho dominio, y por tanto la actividad prefenato deshidratasa, se han conservado durante la evolución de este linaje. Por otro lado, la posición del dominio PAC en la región reguladora ACT de estas proteínas sugiere la existencia de un mecanismo de regulación alostérico que determina la actividad prefenato deshidratasa.

3. El silenciamiento de la actividad prefenato aminotransferasa en las hojas de *Nicotiana benthamiana* activa la expresión de genes codificantes de proteínas con actividad prefenato deshidratasa y transaminasas de aminoácidos aromáticos, sugiriendo que la fenilalanina puede ser sintetizada a través de la ruta del fenilpiruvato cuando la ruta del arogenato es bloqueada.

4. La mutación en el gen *ADT2* provoca en *Arabidopsis thaliana* un fenotipo de semi-esterilidad femenina y constituye el único caso de gen *ADT* para el que hasta el momento se haya descrito una función esencial en el ciclo de vida de las plantas.

5. La prefenato deshidratasa de *Saccharomyces cerevisiae* codificada en el gen *PHA2* puede rescatar el fenotipo de semi-esterilidad femenina presentado por los mutantes *adt2* de *Arabidopsis thaliana*, demostrando que la biosíntesis de fenilalanina puede darse a partir de fenilpiruvato en el órgano estudiado. Al mismo tiempo, este resultado sugiere que ADT2, una enzima bifuncional arogenato deshidratasa / prefenato deshidratasa, podría estar actuando como prefenato deshidratasa durante el desarrollo del gametofito femenino de *Arabidopsis thaliana*, sin que ello probablemente excluya su funcionamiento simultáneo como arogenato dehidratasa.

6. Dos genes de la familia de las arogenato deshidratasas de *Pinus pinaster*, *PpADT-A* y *PpADT-D*, han sido relacionados con la formación de madera de compresión y la lignificación de tejidos vasculares, bajo el control transcripcional de PpMYB8, un factor de transcripción que regula múltiples genes de la ruta de biosíntesis de ligninas. Estos resultados sugieren que sendas isoformas, PpADT-A y PpADT-D, podrían estar especializadas en la biosíntesis de la fenilalanina que posteriormente se dirige hacia la producción de ligninas.

5. Referencias

Boerjan W, Ralph J y Baucher M (2003) *Lignin biosynthesis*. Annu Rev Plant Biol, 54:519-546.

Bross CD, Corea OR, Kaldis A, Menassa R, Bernards MA y Kohalmi SE (2011) *Complementation of the* pha2 *yeast mutant suggests functional differences for arogenate dehydratases from* Arabidopsis thaliana. Plant Physiol Biochem, 49(8):882-890.

Cho MH, Corea OR, Yang H, Bedgar DL, Laskar DD, Anterola AM, Moog-Anterola FA, Hood RL, Kohalmi SE, Bernards MA, Kang C, Davin LB y Lewis NG (2007) *Phe biosynthesis in* Arabidopsis thaliana*. Identification and characterization of arogenate dehydratases*. J Biol Chem, 282(42):30827-30835

Corea OR, Ki C, Cardenas CL, Kim SJ, Brewer SE, Patten AM, Davin LB y Lewis NG (2012a) *Arogenate dehydratase isoenzymes profoundly and differentially modulate carbon flux into lignins*. J Biol Chem, 287(14):11446-11459.

Corea OR, Bedgar DL, Davin LB y Lewis NG (2012b) *The arogenate dehydratase gene family: towards understanding differential regulation of carbon flux through phenylalanine into primary versus secondary metabolic pathways*. Phytochemistry, 82:22-37.

Craven-Bartle B, Pascual MB, Cánovas FM y Ávila C (2013) *A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine*. Plant J, 74: 755–766

Heldt HW y Heldt F (2005) *Phenylpropanoids comprise a multitude of plant secondary metabolites and cell wall components*. En: Heldt HW y Heldt F, *Plant Biochemistry, Third Edition* (435-454). Elsevier Inc.

Lewis NG y Yamamoto E (1990) *Lignin: occurrence, biogenesis and biodegradation*. Annu Rev Plant Physiol Plant Mol Biol, 41:455-496

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, Cooper BR, Rhodes D, Pichersky E y Dudareva N (2010) *RNAi suppression of Arogenate Dehydratase 1 reveals that Phe is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22(3):832-849.

Maeda H, Yoo H y Dudareva N (2011) *Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate*. Nat Chem Biol, 7(1):19-21.

Maeda H y Dudareva N (2012) *The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants*. Annu Rev Plant Biol, 63:73-105.

Mottiar Y, Vanholme R, Boerjan W, Ralph J y Mansfield SD (2016) *Designer lignins: harnessing the plasticity of lignification*. Curr Opin Biotechnol, 14(37):190-200.

Villalobos DP, Díaz-Moreno SM, Said el-SS, Cañas RA, Osuna D, Van Kerckhoven SH, Bautista R, Claros MG, Cánovas FM y Cantón FR (2012) *Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-*

adenosylmethionine, lignin and lignan related genes. BMC Plant Biol, doi: 10.1186/1471-2229-12-100.

Warren RL, Keeling CI, Yuen MM, *et ál*. (2015) *Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism*. Plant J, 83:189- 212.

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, Jannasch AS, Gonda I, Lewinsohn E, Rhodes D y Dudareva N (2013) *An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase*. Nat Commun, 2833. doi:10.1038/ncomms3833.

II. INTRODUCTION

1. Phenylpropanoids as a key innovation for plant terrestrialization

Land colonization by plants in the Devonian was probably one of the most important evolutionary event in life history on Earth, and a key step towards the development of the terrestrial ecosystems as we know them today. Coming from a marine environment, first proto-terrestrial plants found the necessity to deal with the requirements derived from their new habitat. Ancestral land plants had to face multiple new stresses that were absent at their original aquatic environment: desiccation, ultraviolet radiation, terrestrial microbial communities and forces of *gravity.* In this evolutionary transition from aquatic to terrestrial habitats, plants acquired a whole set of key adaptations, promoting the emergence of specialized metabolic pathways.

The phenolic metabolism in plants is one of the major innovations that occurred at early stages of adaptation to the land environment. All these compounds are derived from the carbon skeleton of the amino acid phenylalanine (Phe) (Fraser and Chapple, 2011), and in some plants also from tyrosine (Tyr) (Heldt and Piechulla, 2011). The acquisition of the ability to deaminate the amino acid Phe and hydroxylate its aromatic ring probably led to the accumulation of the first phenylpropanoids, simple compounds that were able to absorb UV light, making possible the survival of archaic land plants (Weng and Chapple, 2010). Although these innovative compounds facilitated the initial transit from water onto land, for tens of millions of years plants remained small because of a lack of mechanical reinforcement to support vertical growth (Bateman et al., 1998).

The development of the ability to deposit the phenolic polymer lignin in the cell walls made possible the evolution of specialized tissues to transport

water from soil to aerial organs of the plant and, at the same time, to provide structural support to withdraw the forces of gravity. This was one of the most striking innovations during land colonization by plants, and made possible the occurrence of larger and more complex body plans: the tracheophytes, also called as “vascular plants”. This was achieved through the development of progressively more sophisticated vascular cells, the xylem vessel elements, with thickened cell walls and highly regulated developmental programs, which include programmed cell death at its final stages. In the vascular bundles of higher plants, xylem cells are also accompanied by fiber cells, sclerenchyma cells with a thickened secondary cell wall. The development of lignin metabolism is considered a key factor for the secondary cell walls formation in this emerging tissues (Weng and Chapple, 2010).

Specialized tissues for water transport and mechanical sustentation are found at the most primitive terrestrial plant groups, particularly in mosses, a group of bryophytes considered to be the ancestors of vascular plants (Kenrick, 2000). In these primitive land plants, there is a water conducting tissue composed by hydroid and stereid cells. Hydroid cells could be considered as an analog structure to xylem vessels from higher plants, having several characteristics in common, such as an elongated cell shape or the absence of cellular content. Despite these similitudes, hydroid cells from bryophytes do not have a lignified cell wall, and lack the characteristic pits that are found in the cell walls of xylem vessels. In fact, bryophytes do not synthesize lignin, although they accumulate abundant soluble phenylpropanoids, as monolignols and flavones (Umezawa, 2003). This suggest that phenylpropanoid biosynthesis evolved before the rise of the first tracheophytes and the lignin polymers, probably as UV-protectant adaptation. The remainder component of the bryophytes bundles, the

stereid cells, are thickened wall cells that are thought to work as supporting elements and could be compared to the fibers cells from vascular plants. Nonetheless, the evolutionary relationship between hydroids-stereids and vessels-fibers is far to be clear. Besides the strong differences at the cell wall structure level previously mentioned, xylem vessels and fibers cells are developed exclusively in the sporophyte generation of vascular plants, whereas hydroids and stereids are formed mainly in the gametophytic generation of bryophytes. It makes even more difficult to establish an evolutionary link, suggesting the possibility of a convergent evolution phenomena that could have conduced to the generation of analogous structures with similar form and function.

In recent times, evidences have been reported about the existence of a conserved regulatory network for water-conducting tissues formation in mosses and vascular plants (Xu et al., 2014). Moreover, it have been shown the existence of a conserved genetic basis controlling bundles development among both plant groups, which seems to indicate a common origin for both water-conducting and supporting elements in land plants. All these events highlight to what extent phenylpropanoid metabolism is tightly linked to occurrence of first terrestrial plants.

2. Phenylpropanoids have multiple roles in plants

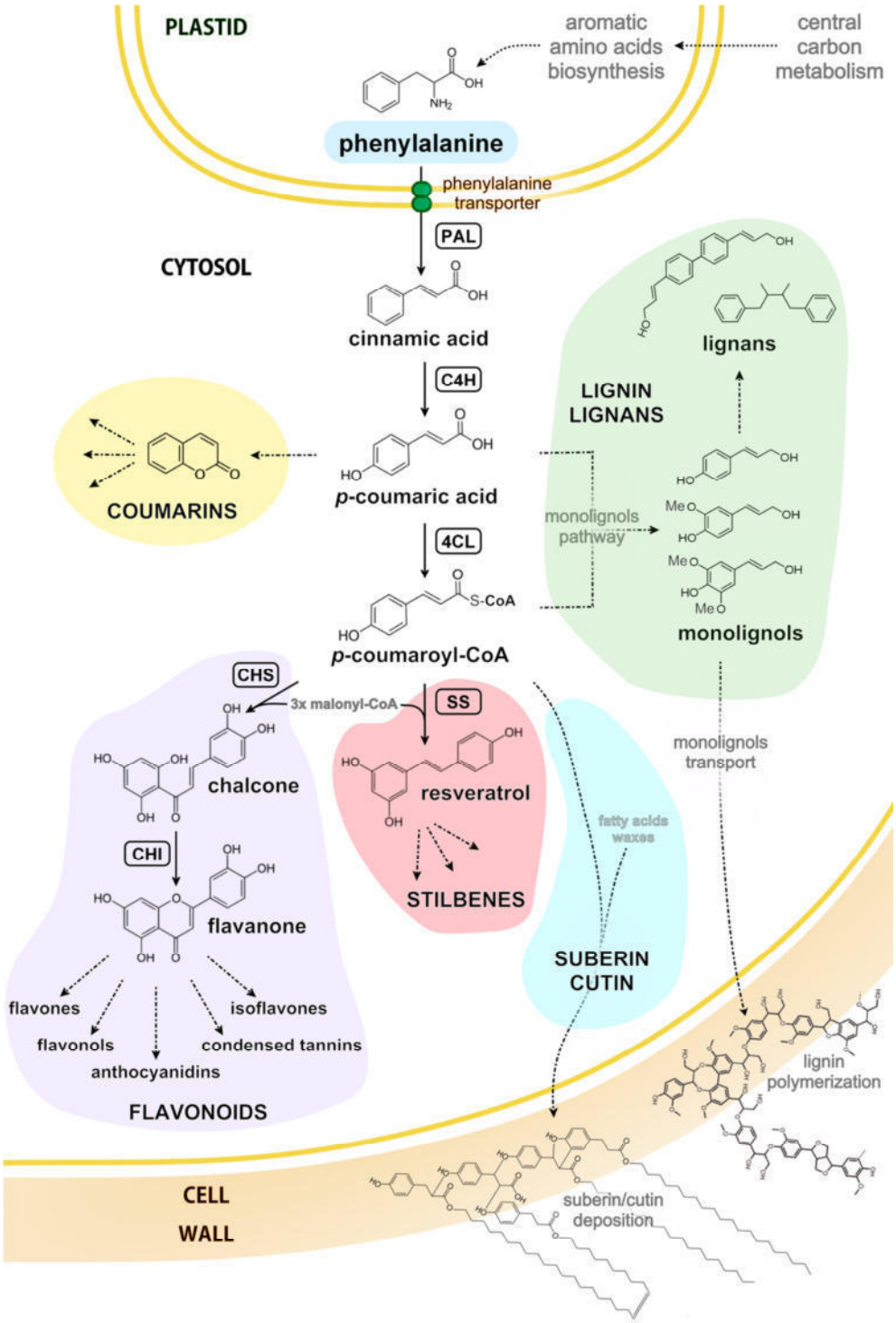
Phenolic compounds are an abundant and diverse group of metabolites characteristic of plants (Tohge et al., 2013). The phenylpropanoid pathway is required for the biosynthesis of lignin, and serves as starting point for the production of many other derivative compounds such as suberins, flavonoids, isoflavonoids, stilbenes or coumarins. Despite being usually classified as "secondary metabolites", multiple and essential roles have been described for phenylpropanoids in plants, including their involvement in response both to biotic and abiotic stresses. These rather structurally divergent compounds have important functions as antibiotics, natural pesticides, signal substances for the establishment of symbiosis with rhizobia, attractants for pollinators, protective agents against ultraviolet radiation, insulating materials to make cell walls impermeable to gases and water, and structural material to assist plant stability (**Table II.1**). As there is an extensive bibliography concerning the metabolism of plant phenolic compounds, to comprehensively review here our current knowledge about this topic would take much more pages than desirable. A concise overview about the metabolism of phenylpropanoids will be provided in this section.

Table II.1. **Some functions of phenylpropanoids** (from Heldt and Piechulla, 2011).

Lignin	Secondary cell walls constituent
Lignan	Antibiotics, toxins against browsing animals
Flavonoids	Antibiotics, signal for interaction with symbionts, flower pigments, light protection substances
Coumarins	Antibiotics, toxins against browsing animals
Suberin and cutin	Formation of impermeable layers
Stilbenes	Antibiotics, especially fungicides
Tannin	Fungicides and protective substances against herbivores

In short, phenylpropanoids are the result of the combination and modification of a relatively limited set of core structures. These component blocks are derived from the shikimate pathway, which is used by bacteria, fungi, algae and plants to synthetize 3 aromatic amino acids: tryptophan (Trp), Phe and Tyr (Herrmann, 1995). As mentioned before, Phe is used by the majority of the plants as the entry point to the phenylpropanoid metabolism after being transported from the plastids into the cytosol, where the phenylpropanoids pathway takes place (**Figure II.1**). The phenylalanine export from plastids into cytosol has remained uncharacterized till recent times, when the first phenylalanine exporter of plants has been identified in *Petunia hybrida* (Widhalm et al., 2015). This transporter is able to export the three aromatic amino acids, and its downregulation reduces the metabolic flux through the aromatic amino acids biosynthetic pathway, suggesting that the exporting at the plastid membrane contributes to regulate the biosynthesis of aromatic amino acids inside the plastid (Widhalm et al., 2015).

Figure II.1 (next page). An overview of plants phenylpropanoid metabolism. Phe, synthesized inside plant plastids, is exported to the cytosol and converted into cinnamic acid by the enzyme phenylalanine ammonia lyase (PAL). In subsequent steps, cinnamate 4-hydroxylase (C4H) and 4-coumarate-CoA ligase (4CL) render the precursors for the various phenylpropanoid biosynthetic branches. In an analogous way, grasses are able to synthesize *p*-coumaric acid from Tyr by the enzyme tyrosine ammonia liase (TAL; not shown). *p*-coumaric acid is the precursor for the biosynthesis of coumarins derivatives. Flavonoids (purple) and stilbenes (magenta) incorporate, either by chalcone synthase (CHS) and chalcone isomerase (CHI) or stilbene synthase (SS), an additional aromatic ring built from three molecules of malonyl-CoA. Suberin and cutin (turquoise) form esters with long-chain fatty acids and waxes, conferring gas and water impermeable properties. Subsequent rounds of hydroxylations and methylations over the aromatic ring will provide monolignols, the basic monomers for lignans and lignin biosynthesis (dark green).


PLASTID
aromatic amino acids biosynthesis
central carbon metabolism
phenylalanine
phenylalanine transporter
PAL
CYTOSOL
cinnamic acid
C4H
lignans
LIGNIN LIGNANS
COUMARINS
p-coumaric acid
monolignols pathway
4CL
S-CoA
p-coumaroyl-CoA
monolignols
CHS
3x malonyl-CoA
SS
monolignols transport
chalcone
resveratrol
fatty acids waxes
CHI
STILBENES
flavanone
SUBERIN CUTIN
flavones
isoflavones
flavonols
condensed tannins
anthocyanidins
FLAVONOIDS
lignin polymerization
CELL WALL
suberin/cutin deposition

The enzyme phenylalanine ammonia lyase (PAL) catalyzes the first step in the phenylpropanoids biosynthesis: the deamination of Phe. It produces the formation of a carbon-carbon double bond during the deamination, rendering *trans*-cinnamic acid. The phenyl ring will be hydroxylated in subsequent reactions via cytochrome P_{450}-dependant monooxygenase enzymes (**Figure II.1**). In some grasses, PAL enzymes can also work as tyrosine ammonia lyase (TAL) with similar efficiency, yielding *p*-cumaric acid in an analogous way (Barros et al., 2016). Subsequent rounds of hydroxylations and methylations on the aromatic ring will render a set of simple phenols that act as precursors of more complex molecules. The combination of these simple phenols, either amongst themselves and/or with other compounds from a different metabolic origin, will provide thousands of different molecules with the most diverse functions (**Figure II.1**).

For its biological and economic significance, monolignol metabolism and lignin are between the most well-studied phenolic compounds. For an in-depth reading about lignin metabolism, the review published by Boerjan et al. in 2003 could be suggested as a main reference in the field, as well as Mottiar et al. 2016 for a better understanding about current research in lignin biotechnology and its interesting future prospects. In addition to their physiological role, lignins have a great interest from an economic point of view, as far as they block access to cell wall sugars and strongly limit biomass harnessing for livestock feed, paper manufacturing, and second generation biofuel production (Chapple et al., 2007). Lignin metabolism will be described in further detail in the subsequent section.

Flavonoids, with around 8000 estimated metabolites, are by far the largest class of plant phenols and are present in different amounts, according to

the plant species, organs, developmental stages and growth conditions, performing a wide range of functions (Petrussa et al., 2013; Tohge et al., 2013). They have an in common three phenolic ring chemical structure (C6-C3-C6) that comes from the condensation of three molecules of malonyl-CoA and one of *p*-coumaryl-CoA by the enzyme chalcone synthase. Chalcone is the precursor for the synthesis of all flavonoids. The flavonoid biosynthetic pathway has been characterized in *Arabidopsis thaliana* and *Zea mays* to a large extent, but also in *Vitis vinifera* (Petrussa et al., 2013). Stilbenes, another important group of phenylpropanoids, are also synthetized from the same substrates as chalcone, but the basic structure resulting from the stilbene synthase activity differs from those formed by chalcone synthase.

A particular attention has been paid to an increasing number of flavonoids and stilbenes with health beneficial properties. It is possible to find a large body of scientific publications reporting evidences about such effects. In this way, certain isoflavonoids, flavonoids, stilbenes, and lignans, collectively termed as phytoestrogens, can act with estrogenic or antiestrogenic activity toward mammals, with multiple applications for human health (Dixon, 2004). Resveratrol, a phytoalexin belonging to the stilbene group synthetized by various plants in response to external stressors, has been suggested to have multiple effects in the prevention or improvement of glucose metabolism disorders, inflammation, cancer and certain liver diseases, when supplied to laboratory animals (Poulsen et al., 2013).

3. Lignins are the most abundant phenylpropanoid in plants

Phenylpropanoids can represent, in terms of carbon metabolic fluxes in the plant, up to 30% of photosynthetically fixed carbon, with lignin as the second biopolymer on the biosphere just beside cellulose (Boerjan et al., 2003; Vogt, 2010). For the particular case of lignin, it is estimated that could represent from 18% to 35% of the plant biomass by weigh (Sarkanem and Ludwig, 1971). Thus, with the occurrence of the phenylpropanoid pathway and lignin metabolism during plant evolution, aromatic amino acid biosynthesis had to be substantially improved in terms of efficiency, in order to supply the emerging massive demand of carbon skeletons required for this novel metabolic route.

Lignin is a polymer formed by a complex, ramified network of phenolic monomers, and its presence is the most distinguishing feature of plant secondary cell walls (Carpita et al., 2015). Three simple phenylpropanoids, the hydroxycinnamyl alcohols or monolignols, account for the formation of most of the lignins: p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. The biosynthesis of monolignols is well documented and appears to occur in the cytosol. It encomprises successive hydroxylation and O-methylation of the aromatic ring and conversion of the side chain carboxyl to an alcohol function (Dixon et al., 2001) (**Figure II.2**). The three monolignols only differ in the degree of the phenolic ring methoxylation (Mottiar et al., 2016). Once incorporated into the lignin polymer, they produce *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) moieties.

Pathways leading to monolignols formation are currently viewed as a metabolic grind, through which the successive reactions may occur at different levels and in a different sequence. These pathways occurs in the

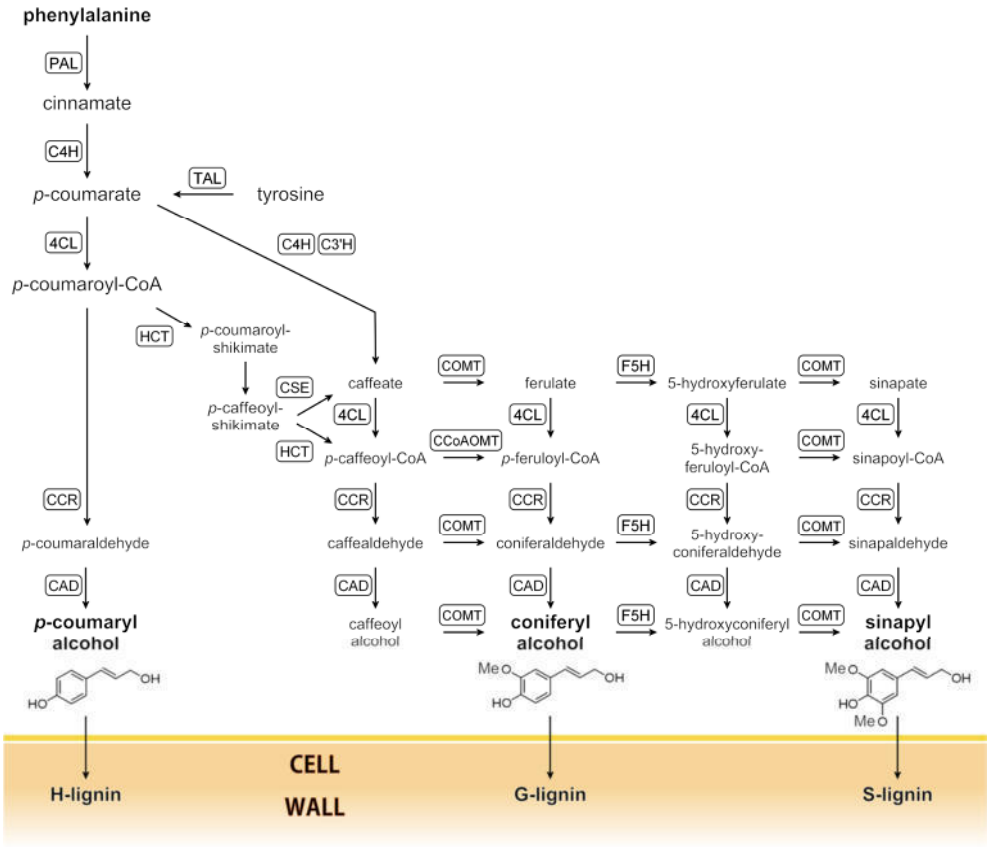


Figure II.2. Monolignol biosynthesis in plants. The monolignols, *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol, are produced in the general phenylpropanoid and the monolignol pathway in the cytosol. Then, monolignols are exported outside the cell, where the lignin polymerization takes place. Some of the enzymes involved in the monolignol pathway are thought to be linked to the endoplasmic reticulum membrane. Not all the depicted enzymatic reactions are present in all plants: TAL seems to be distinctive from grasses, gymnosperms lack F5H, CSE activity has only been demonstrated in *Arabidopsis* and the direct route from *p*-coumarate to caffeate has only been demonstrated in poplar. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate-CoA ligase; TAL, tyrosine ammonia lyase; HCT, shikimate/quinate *p*-hydroxycinnamoyltransferase; C3'H, *p*-coumaroyl-shikimate/quinate-3-hydroxylase; CSE, caffeoyl shikimate esterase; CCR, cinnamoyl-CoA reductase; CCoAOMT, caffeoyl-CoA *O*-methyltransferase; F5H, ferulate-5-hydroxylase; COMT, caffeic acid/5-hydroxyconiferaldehyde *O*-methyltransferase; CAD, cinnamyl alcohol dehydrogenase.

cytosol and, after their biosynthesis, monolignols may be exported out of the cell and reach the site of polymerization in the cell walls. Monolignols can also be glycosylated in reactions associated with the endoplasmic reticulum and Golgi apparatus, changing their stability and solubility. Glycosilation seems to participate in detoxification processes that involve targeting to subcellular compartments (Dima et al., 2015; Le Roy et al., 2016). The relationship between monolignols glycosylation/de-glycosilation and lignification is, on the other hand, still not fully clarified.

Once in the cell wall, laccase and peroxidase enzymatic activities generate monolignol radicals that will be incorporated into the growing lignin polymer. The combinatorial coupling of radicals results in highly variable racemic polymers with different physicochemical features (Ralph et al., 2004). The rate of the different lignin moieties changes between plant species and tissues: whereas lignins from gymnosperms and related species are rich in G units and contain low amounts of H units, dicot lignins are mainly composed of G and S units (Wang et al., 2013).

4. The aromatic amino acids are formed from the shikimate pathway

Aromatic amino acids are used in plants for protein biosynthesis and also as precursors of numerous natural products. Aromatic amino acids are derived from the shikimate pathway, which does not exist in animals. It makes these amino acids essential components of the diet of humans. For the same reason, the enzymes of this pathway have been largely targeted for the development of herbicides. For instance, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase is the target for the broad-spectrum herbicide glyphosate (a structural analogue of phospho*enol*pyruvate) that competitively inhibits the enzyme by occupying the phospho*enol*pyruvate binding-site (Schönbrunn et al., 2001).

The shikimate pathway (**Figure II.3**) comprises a sequence of 7 metabolic steps, than begins with the condensation of phospho*enol*pyruvate and D-erythrose-4-phosphate, and ends with the synthesis of chorismate, the final product of the pathway (Herrmann and Weaver, 1999). All the shikimate pathway enzymes have been characterized in both plants and microbes, and their corresponding coding genes have been identified (Maeda and Dudareva, 2012). Chorismate is a key precursor in plants of at least four different metabolic pathways leading to the formation of Trp, Phe/Tyr, salicylate/phylloquinone, and folate. Aside from aromatic amino acids, the plant hormone salicylic acid or vitamins B_9 and K_1 are examples of chorismate-derived metabolites (Coruzzi et al., 2015). Hence, chorismate represents a branch point for two subsequent aromatic amino acid biosynthetic pathways:

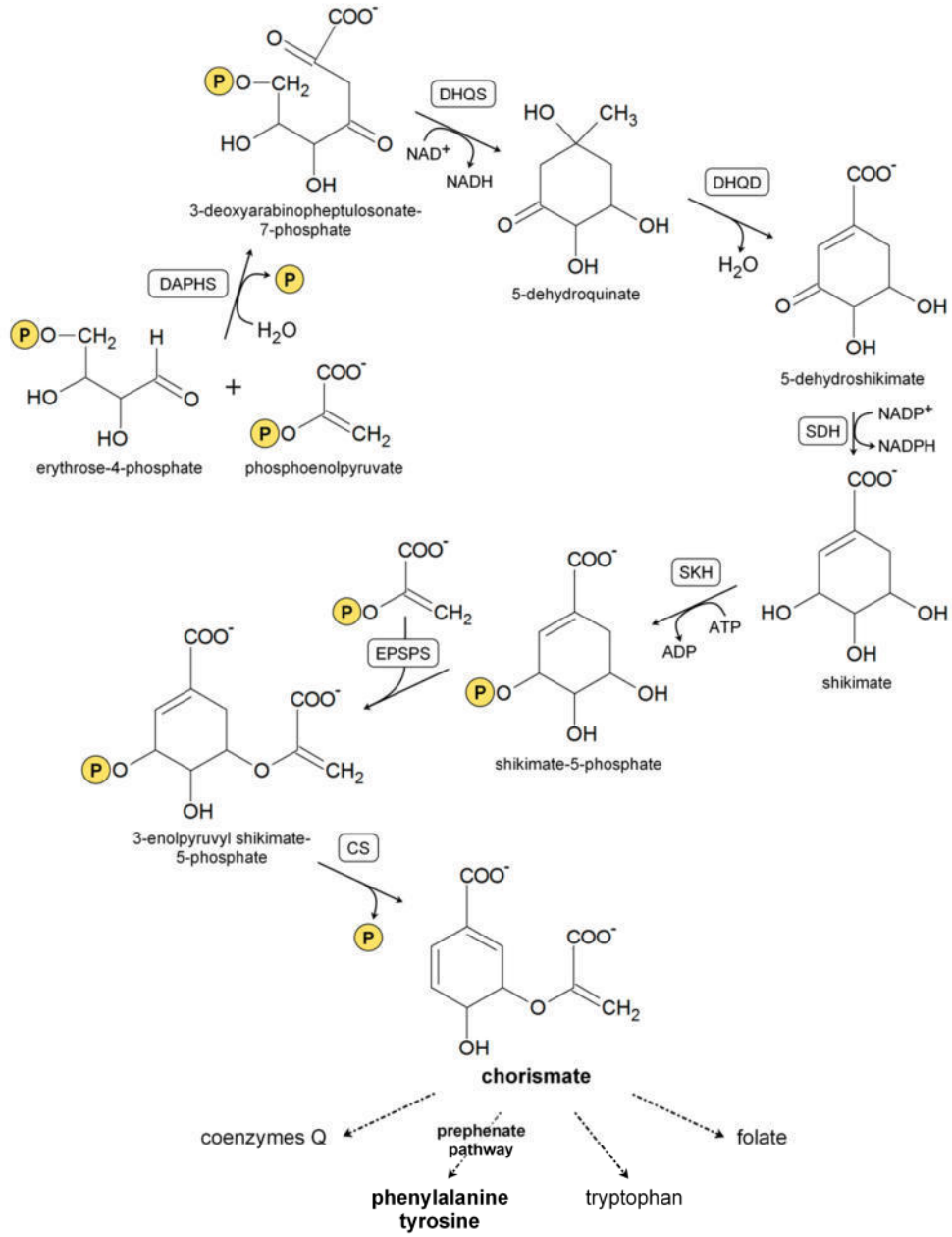
1- Tryptophan pathway, which starts with the enzyme anthranilate synthase and involves 6 enzymatic reactions.

2- Prephenate pathway. Prephenate is produced by a rearrangement of the side chain to the 1'-position of the ring. Subsequent transamination reactions and dehydratation or $NAD(P)^+$-dependent oxidation of the ring, accompanied by decarboxylation, results in the formation of Phe and Tyr.

While Trp pathway will be concisely described later in this section, prephenate pathway for the biosynthesis of Tyr and Phe will be dealt with in depth in section 5.

The shikimate pathway acts as the link between plant central carbon metabolism and the synthesis of the secondary metabolites derived from Trp, Phe and Tyr (Hermann and Weaver, 1999). Yet, the shikimate pathway may not be considered as a pathway restricted to the generation of amino acids for protein biosynthesis, because it also provides the precursors required for a large variety of other compounds (see section I.2 and I.3). Regardless of the importance of aromatic amino acids and their derivatives for plants, our current knowledge thereon has been mainly inferred from microbial organism, resulting in a poor understanding about aromatic amino acid biosynthesis regulation in plants (Tzin and Galili, 2010).

Figure II.3 (next page). The shikimate pathway. In plants, the shikimate pathway occurs inside the plastids. One erythrose-4-phosphate and two phospho*enol*pyruvate molecules are used for biosynthesis of chorismate via seven enzymatic reactions. Chorismate, the final product of the shikimate pathway, is the precursor for the biosynthesis of tryptophan, Tyr and Phe, in addition to other important compounds. Enzyme abbreviations: DAPHS, 3-deoxyarabinoheptulosonate-7-phosphate synthase; DHQS, 5-dehydroquinate synthase; DHQD, 3-dehydroquinate dehydratase; SDH, shikimate dehydrogenase; SK, shikimate kinase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; CS, chorismate synthase.



Anthranilate synthase is the enzyme responsible for the initiation of the tryptophan pathway, which converts chorismate into this amino acid through five additional enzymatic reactions: phosphoribosylanthranilate transferase, phosphoribosylanthranilate isomerase, indole-3-glycerol phosphate synthase and tryptophan synthase (this last enzyme catalyzes both last steps of the pathway). All six tryptophan pathway enzymes have been shown or predicted to have plastidial localization. Indole-3-acetic acid (IAA), the main auxin synthetized by plants, can be produced *de novo* from a tryptophan-dependent pathway that has been demonstrated to be essential in different developmental processes, as embryogenesis, seedling growth or flower development (Zhao, 2012). In addition to serve as a source for auxins biosynthesis, Trp is also used as precursor for many other secondary metabolites, included indole alkaloids, phytoalexins, cyclic hydroxamic acids indole glucosinolates and acridone alkaloids (Coruzzi et al., 2015).

5. Phe and Tyr are synthesized from chorismate in the prephenate pathway

In contrast with tryptophan pathway, Phe and Tyr biosynthesis have received much less attention since the development of the molecular biology of plants. Despite the extreme significance of both amino acids, most especially Phe as a precursor for phenylpropanoids, many genes encoding for critical enzymes of this pathway have remained unidentified for years. Yet, in the last decade there has been an increasing interest for the study of Phe and Tyr pathways, and only since recent times we know the coding genes for all the involved enzymatic activities (Cho et al., 2007; Maeda et al. 2011).

The committed step in the prephenate pathway (**Figure II.4**) is the conversion of chorismate to prephenate by the enzyme chorismate mutase (CM). CMs enzymes largely diverge between the different organisms that produce aromatic amino acids. Despite monofunctional CM are more frequent in nature and characteristic of plants, there are a few examples of bifunctional CM enzymes. *Escherichia coli* CMs are also able to work as prephenate dehydratase (CM/PDT, encoded by *PheA* gene; Tzin et al., 2009) or prephenate dehydrogenase (CM/PDH, encoded by *TyrA* gene; Lütke-Eversloh and Stephanopoulos, 2005). Plants always contain at least two different CM isoforms (3 in some plants, included *Arabidopsis*), denominated as CM1 and CM2. Whereas CM1 is a plastidial isoform with feedback-inhibition in response to Tyr and Phe and is activated by Trp, CM2 is cytosolic and has been shown to be usually insensitive to allosteric regulation by aromatic amino acids (Colquhoun et al., 2010). As no putative source of cytosolic chorismate has been currently described, and all the

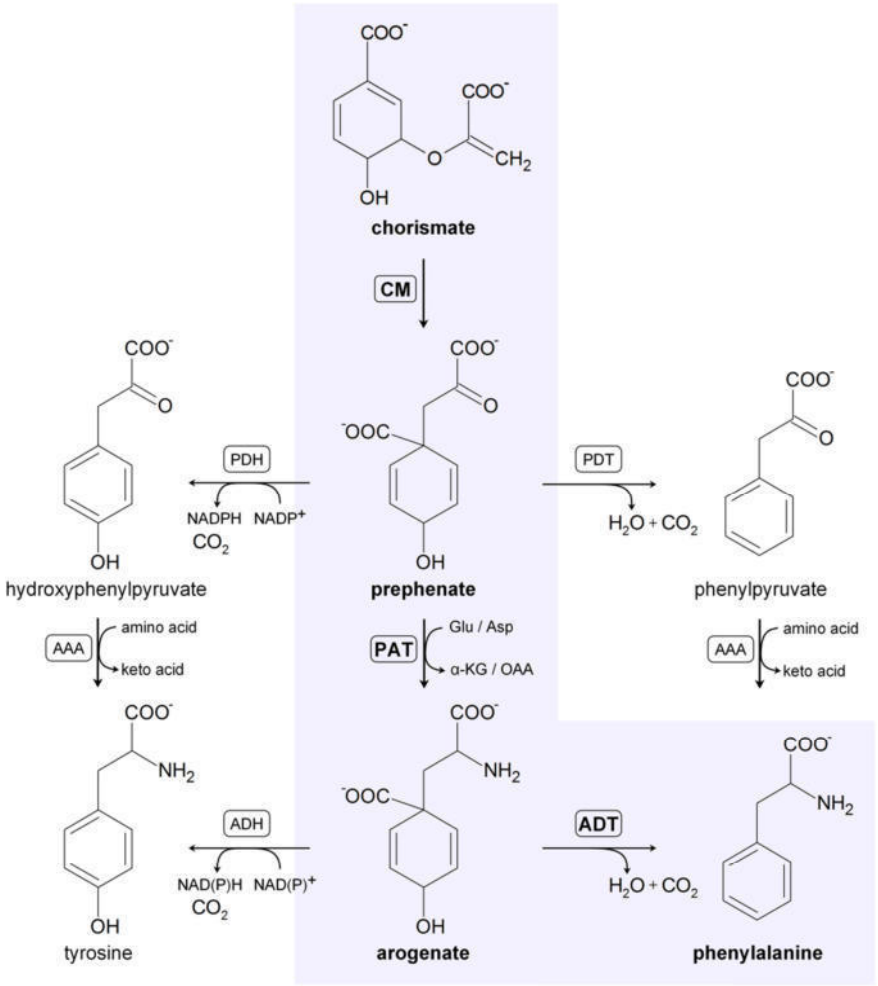


Figure II.4. The prephenate pathway in plants. Chorismate from prephenate pathway is converted into prephenate by the enzyme chorismate mutase (CM). Either Phe and Tyr can be synthesized by two alternative pathways which branch from prephenate, although the arogenate branch (remarked in pale blue) is considered to be the major source of Phe in plants. On the other hand, prephenate dehydrogenase activity (PDH) in Tyr biosynthesis has only been demonstrated in legumes. Other enzyme abbreviations: PDT, prephenate dehydratase; PAT, prephenate amino transferase; AAA, aromatic amino acid amino transferase; ADH, arogenate dehydrogenase; ADT, arogenate dehydratase.

remaining prephenate pathway enzymes seems to be localized inside the plastids, physiological function of the cytosolic CM2 remains unclear.

The conversion of prephenate to Phe or Tyr may occur via two alternative pathways: the arogenate pathway and the phenylpyruvate/4-hydroxyphenylpyruvate pathway. In the arogenate pathway, the keto group from prephenate is transaminated to L-arogenate, also termed as pre-tyrosine (more particularly when consulting older references) by the enzyme prephenate-arogenate aminotransferase (PAT). This is a reversible reaction that uses PLP as cofactor, and PAT enzymes are only able to use L-glutamate or L-aspartate as amino donors (Bonner and Jensen, 1985). Kinetic characterization of PAT enzymes from plants showed that they have roughly 10-fold higher affinity toward prephenate than toward arogenate, suggesting that the reaction net flux occurs mainly in the forward direction (Maeda and Dudareva, 2012). Interestingly, the coding gene for this enzyme was the last Phe and Tyr biosynthesis gene in being identified (Graindorge et al., 2010; Maeda et al., 2011). It also functions as an aspartate aminotransferase putatively involved in the plastidial biosynthesis of other amino acids (de la Torre et al., 2006; 2009).

After this shared step, L-arogenate can either be dehydrated/decarboxylated to Phe by the enzyme arogenate dehydratase (ADT), or dehydrogenated/decarboxylated to Tyr by arogenate dehydrogenase (ADH). The arogenate pathway is considered to be the main route for Phe biosynthesis in plants (Maeda et al., 2010). As reported by the cited publication, ADT1 suppression via RNA interference in *Petunia hybrida* petals significantly reduced ADT activity, levels of Phe, and derived phenylpropanoid/benzenoid volatiles.

In the phenylpyruvate/4-hydroxyphenylpiruvate pathways, reactions occur in reverse order than in the arogenate pathway: first prephenate is dehydrated/decarboxylated to phenylpyruvate by prephenate dehydratase (PDT) or dehydrogenated/decarboxylated to form 4-hydroxyphenylpyruvate by prephenate dehydrogenase (PDH). Recently, it has been reported the existence of a PDH enzyme from legumes that seems to be localized outside the plastids and insensitive to Tyr inhibition (Schenck et al., 2015). The corresponding products from PDT or PDH reaction endure transamination to Phe or Tyr, respectively, by aromatic amino acid aminotransferases (AAAs). In most microorganisms studied to date, included *Escherichia coli* and *Saccharomyces cerevisiae* (baker's yeast), Phe and Tyr biosynthesis occurs primarily by this pathway (Zamir et al., 1987).

6. Can plants ADTs also work as PDTs?

PDT and ADT activities convert prephenate into phenylpyruvate and arogenate into Phe, respectively, and can be housed in the same proteins in plants. The kinetic characterization of *Arabidopsis* ADT family (Cho et al., 2007), which consist of a total number of 6 isoforms, showed that some ADTs, more particularly ADT1, ADT2 and ADT6, are also able to display PDT activity although with lower catalytic efficiency. The yeast *S. cerevisiae* only synthesizes Phe through the phenylpyruvate pathway by the single copy PDT gene *PHA2*, and its mutation causes Phe auxotrophy. Bifunctional ADT/PDT proteins from plants can rescue the Phe auxotrophic phenotype caused by the deletion of the endogenous PDT activity in the yeast mutant *pha2*, reinforcing the idea that PDT activity in plants could work *in vivo* (Bross *et al*., 2011).

Although arogenate pathway is thought to be quantitatively more important for Phe biosynthesis than phenylpyruvate pathway, PDT activity has been detected in plant tissues (Warpeha et al., 2006; Maeda et al., 2010; Yoo et al., 2013). Additionally, RNAi interference of PAT RNAi in petunia petals showed in some cases less than 20% of reduction in Phe levels (Maeda et al., 2011), suggesting that the phenylpyruvate pathway may be functional, although its particular role remains unclear. In the same line, Yoo *et al*. (2013) reported a cytosolic phenylpyruvate-aminotransferase from *P. hybrida* able to catalyze efficiently the interconversion of phenylpyruvate and Tyr to Phe and 4-hydroxyphenylpyruvate, respectively. Through feeding experiments with labelled phenylpyruvate as well as silencing and overexpression analysis, these authors proposed the existence of a functional phenylpyruvate pathway that has enhanced flux when blocking the arogenate pathway. Putting it all together, there are strong evidences

indicating the existence of a functional phenylpyruvate pathway in plants. However, none of these studies have reported evidences about a specific role for the phenylpyruvate pathway in plant development.

7. Phe biosynthesis is regulated at transcriptional and post-transcriptional levels

Unlike microorganisms and algae, that only need to maintain a basal level of aromatic amino acids to satisfy the demand for protein biosynthesis, plants have to synthesize a large amount of secondary metabolites from these amino acids. Ultimately, all this carbon flux goes through shikimate pathway and the two branches derived from chorismate. For this reason, aromatic amino acid biosynthesis must be regarded as one of the major routes in plant metabolism, and its regulation should be coordinated with the biosynthesis of downstreamed secondary metabolites, whose demand can dramatically change in response to environmental and developmental conditions (Maeda and Dudareva, 2012). The regulation of metabolic fluxes to aromatic amino acid biosynthesis in plants occurs both at transcriptional and post-transcriptional levels.

There is limited information in plants about how aromatic amino acids levels affect the expression of the shikimate pathway genes. Previous publications reported that reduced Phe levels in petunia flowers of *ADT1*-RNAi lines increased the expression of the shikimate pathway genes (Maeda et al., 2010). Although the underlying molecular mechanisms are currently unknown, it seems that reduced levels of Phe may act as an inductor for the expression of the shikimate pathway genes. As reviewed by Tzin and Galili (2010) and Maeda and Dudareva (2012), a number of publications have reported the implication of MYB and NAC transcription factors in the transcriptional regulation of the shikimate pathway. Petunia ODORANT1, a R2R3-MYB transcription factor, has been linked to Phe biosynthesis and volatiles production in petunia flowers (Dal Cin et al., 2011). More recently, it has been reported that R2R3-MYB transcription

factors from *Picea glauca* have a putative role in co-regulation of the shikimate pathway and phenylpropanoid biosynthesis (Bomal et al., 2014). In the last year, transcriptome analysis of protoxylem vessels elements have indicated that cell differentiation involves the active up-regulation of genes encoding the enzymes catalyzing phospho*enol*pyruvate biosynthesis from oxaloacetate and shikimate pathway enzymes for Phe biosynthesis (Ohtani et al., 2016).

Post-translational regulation is also important in the Phe and Tyr pathways. CM, the first specific enzyme of this pathway, is feedback inhibited by Phe and Tyr and induced by Trp (Eberhard et al., 1996). ADTs, as branch point enzymes, are also an important control step in the metabolic flux of the pathway. In tobacco, spinach, and *Sorghum bicolor*, ADT activity has been shown to be positively regulated by Tyr and negatively regulated by Phe (Jung et al., 1986; Siehl and Conn, 1988), although this aspect has not been studied in *Arabidopsis* (Cho et al., 2007). In addition, ADTs from rice are negatively regulated by Phe as well (Yamada et al., 2008).

Plants ADTs consist of a catalytic domain and a C-terminal allosteric regulatory domain called ACT-domain (from **a**spartokinase, **c**horismate mutase and **T**yrA) also present in a variety of other enzymatic proteins (Chipman and Shaanan, 2001). The ACT domain is involved in an allosteric feedback-inhibition: it has been reported that point mutations in this domain can disrupt the enzyme allosteric regulation in response to Phe, resulting in an over accumulation of this amino acid in the mutants (Yamada et al., 2008).

8. References

Bateman RM, Crane PR, DiMichele WA, Kenrick PR, Rowe NP, Speck T and Stein WE (1998) *Early evolution of land plants: phylogeny, physiology, and ecology of the primary terrestrial radiation*. Annu Rev Ecol Sys, 29:263-292.

Barros J, Serrani-Yarce JC, Chen F, Baxter D, Venables BJ and Dixon RA (2016) *Role of bifunctional ammonia-lyase in grass cell wall biosynthesis*. Nat Plants, 16050:1-9.

Boerjan W, Ralph J and Baucher M (2003) *Lignin biosynthesis*. Annu Rev Plant Biol, 54:519-546.

Bomal C, Duval I, Giguère I, Fortin É, Caron S, et al. (2014) *Opposite action of R2R3-MYBs from different subgroups on key genes of the shikimate and monolignol pathways in spruce*. J Exp Bot, 65(2):495-508.

Bonner CA and Jensen RA (1985) *Novel features of prephenate aminotransferase from cell cultures of* Nicotiana silvestris. *Arch Biochem Biophys,* 238:237-246.

Bross CD, Corea OR, Kaldis A, Menassa R, Bernards MA and Kohalmi SE (2011) *Complementation of the* pha2 *yeast mutant suggests functional differences for arogenate dehydratases from* Arabidopsis thaliana. Plant Physiol Biochem, 49(8):882-890.

Carpita NC, Ralph J and McCann MC. *The Cell Wall*, pp. 45-110. In: Buchanan BB, Gruissem W and Jones RL (2015) *Biochemistry & Molecular Biology of Plants, Second Edition*. John Wiley & Sons, Ltd. Oxford, UK.

Chapple C, Ladisch M and Meilan R (2007) *Loosening lignin's grip on biofuel production*. Nat Biotechnol, 25:746-748

Chipman DM and Shaanan B (2001) *The ACT domain family*. Curr Opin Struct Biol, 11(6):694-700.

Cho MH, Corea OR, Yang H, Bedgar DL, Laskar DD, et al. (2007) *Phe biosynthesis in* Arabidopsis thaliana*. Identification and characterization of arogenate dehydratases*. J Biol Chem, 282(42):30827-30835

Colquhoun TA, Schimmel BCJ, Kim JY, Reinhardt D, Cline K and Clark DG (2010) *A petunia chorismate mutase specialized for the production of floral volatiles*. Plant J, 61(1):145-155.

Coruzzi G, Last R, Dudareva N and Amrhein N. *Amino Acids*, pp. 289-336. In: Buchanan BB, Gruissem W and Jones RL (2015) *Biochemistry & Molecular Biology of Plants*, *Second Edition*. John Wiley & Sons, Ltd. Oxford, UK.

Dal Cin V, Tieman DM, Tohge T, McQuinn R, de Vos RC, et al. (2011) *Identification of genes in the Phe metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit*. Plant Cell, 23(7):2738-2753.

de la Torre F, De Santis L, Suárez MF, Crespillo R and Cánovas FM (2006) *Identification and functional analysis of a prokaryotic-type aspartate aminotransferase: implications for plant amino acid metabolism*. Plant J, 46(3):414-425.

de la Torre F, Moya-García AA, Suárez MF, Rodríguez-Caso C, Cañas RA, Sánchez-Jiménez F and Cánovas FM (2009) *Molecular modeling and site-directed mutagenesis reveal essential residues for catalysis in a prokaryote-type aspartate aminotransferase*. Plant Physiol, 149(4):1648-1660.

Dima O, Morreel K, Vanholme B, Kim H, Ralph J and Boerjan W (2015) Small *Glycosylated Lignin Oligomers Are Stored in Arabidopsis Leaf Vacuoles*. Plant Cell, 27(3):695-710.

Dixon RA, Chen F, Guo D and Parvathi K. (2001) *The biosynthesis of monolignols: a "metabolic grid", or independent pathways to guaiacyl and syringyl units?* Phytochemistry, 57(7):1069-1084.

Dixon RA (2004) *Phytoestrogens*. Annu Rev Plant Biol, 55:225-261

Eberhard J, Ehrler TT, Epple P, Felix G, Raesecke HR, Amrhein N and Schmid J (1996) *Cytosolic and plastidic chorismate mutase isozymes from* Arabidopsis thaliana*: molecular characterization and enzymatic properties*. Plant J, 10:815-821.

Fraser CM and Chapple C (2011) *The Phenylpropanoid Pathway in* Arabidopsis. The Arabidopsis Book, 9: e0152

Graindorge M, Giustini C, Jacomin AC, Kraut A, Curien G and Matringe M (2010) *Identification of a plant gene encoding glutamate/aspartate-prephenate aminotransferase: the last homeless enzyme of aromatic amino acids biosynthesis*. FEBS Lett, 584(20):4357-4360

Heldt HW and Piechulla B. *Phenylpropanoids comprise a multitude of plant secondary metabolites and cell wall components,* pp. 431-449. In: Heldt HW and Piechulla B (2011) *Plant Biochemistry, Fourth Edition*. Academic Press, Elsevier Publishing. London, UK

Herrmann K (1995) *The shikimate pathway as an entry to aromatic secondary metabolism*. Plant Physiol, 107:7-12

Herrmann KM and Weaver LM (1999) *The Shikimate Pathway*. Annu Rev Plant Physiol Plant Mol Biol, 50:473-503.

Jung E, Zamir LO and Jensen RA (1986) *Chloroplasts of higher plants synthesize L-Phe via L-arogenate*. Proc Natl Acad Sci USA, 83:7231-7235.

Kenrick P (2000) *The relationships of vascular plants*. Philos Trans R Soc Lond B Biol Sci, 355(1398):847–855.

Le Roy J, Huss B, Creach A, Hawkins S and Neutelings G (2016) *Glycosylation Is a Major Regulator of Phenylpropanoid Availability and Biological Activity in Plants*. Front Plant Sci, 7:735.

Lütke-Eversloh T and Stephanopoulos G (2005) *Feedback inhibition of chorismate mutase/prephenate dehydrogenase (*TyrA*) of* Escherichia coli*: generation and*

characterization of tyrosine-insensitive mutants. Appl Environ Microbiol, 71(11):7224-7228.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, et al. (2010) *RNAi suppression of Arogenate Dehydratase 1 reveals that Phe is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22(3):832-849.

Maeda H, Yoo H, Dudareva N (2011) *Prephenate aminotransferase directs plant Phe biosynthesis via arogenate*. Nat Chem Biol, 7(1):19-21.

Maeda H and Dudareva N (2012) *The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants.* Annu Rev Plant Biol, 63:73-105.

Mottiar Y, Vanholme R, Boerjan W, Ralph J and Mansfield SD (2016) *Designer lignins: harnessing the plasticity of lignification*. Curr Opin Biotechnol, 14(37):190-200

Ohtani M, Morisaki K, Sawada Y, Sano R, Uy AL, et al. (2016) *Primary Metabolism during Biosynthesis of Secondary Wall Polymers of Protoxylem Vessel Elements*. Plant Physiol, 172(3):1612-1624.

Petrussa E, Braidot E, Zancani M, Peresson C, Bertolini A, Patui S and Vianello A (2013) *Plant Flavonoids – Biosynthesis, Transport and Involvement in Stress Responses*. Int J Mol Sci, 14(7):14950-14973.

Poulsen MM, Jørgensen JO, Jessen N, Richelsen B and Pedersen SB (2013) *Resveratrol in metabolic health: an overview of the current evidence and perspectives*. Ann NY Acad Sci, 1290:74-82.

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, et al. (2004) *Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids*. Phytochem Rev, 3:29-60

Sarkanen KV and Ludwig CH (1971) *Lignins: Occurrence, Formation, Structure and Reactions*. Wiley-Interscience

Schenck CA, Chen S, Siehl DL and Maeda HA (2015) *Non-plastidic, tyrosine-insensitive prephenate dehydrogenases from legumes*. Nat Chem Biol, 11(1):52-57

Schönbrunn E, Eschenburg S, Shuttleworth WA, Schloss JV, Amrhein N, et al. (2001) *Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail*. Proc Natl Acad Sci USA, 98:1376-1380

Siehl DL and Conn EE (1988) *Kinetic and regulatory properties of arogenate dehydratase in seedlings of* Sorghum bicolor (L.) Moench Arch Biochem Biophys, 260:822-829

Tohge T, Watanabe M, Hoefgen R and Fernie AR (2013) *The evolution of phenylpropanoid metabolism in the green lineage*. Crit Rev Biochem Mol Biol, 48(2):123-152

Tzin V, Malitsky S, Aharoni A and Galili G (2009) *Expression of a bacterial bi-functional chorismate mutase/prephenate dehydratase modulates primary and secondary metabolism associated with aromatic amino acids in* Arabidopsis. Plant J, 60(1):156-167.

Tzin V and Galili G (2010) *The Biosynthetic Pathways for Shikimate and Aromatic Amino Acid Biosynthesis in* Arabidopsis thaliana. The Arabidopsis Book, e0132

Umezawa T (2003) *Diversity in lignan biosynthesis*. Phytochem Rev, 2:371-390

Vogt T (2010) *Phenylpropanoid biosynthesis*. Mol Plant, 3(1):2-20

Wang Y, Chantreau M, Sibout R and Hawkins S (2013) *Plant cell wall lignification and monolignol metabolism*. Front Plant Sci, 4:220

Warpeha KM, Lateef SS, Lapik Y, Anderson M, Lee B-S and Kaufman LS (2006) *G-protein-coupled receptor 1, G-protein Gα-subunit 1, and prephenate dehydratase 1 are required for blue light-induced production of Phe in etiolated* Arabidopsis. Plant Physiol, 140:844-855

Weng JK and Chapple C (2010) *The origin and evolution of lignin biosynthesis*. New Phytol, 187(2):273-285.

Widhalm JR, Gutensohn M, Yoo H, Adebesin F, Qian Y, et al. (2015) *Identification of a plastidial phenylalanine exporter that influences flux distribution through the phenylalanine biosynthetic network*. Nat Commun, 6:8142.

Xu B, Ohtani M, Yamaguchi M, Toyooka K, Wakazaki M, et al. (2014) *Contribution of NAC Transcription Factors to Plant Adaptation to Land*. Science, 343(6178):1505-1508

Yamada T, Matsuda F, Kasai K, Fukuoka S, Kitamura K, Tozawa Y, Miyagawa H and Wakasa K (2008) *Mutation of a Rice Gene Encoding a Phe Biosynthetic Enzyme Results in Accumulation of Phe and Tryptophan*. Plant Cell, 20(5):1316-1329.

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, et al. (2013) *An alternative pathway contributes to Phe biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase*. Nat Commun, 4:2833.

Zamir LO, Tiberio R, Fiske M, Berry A and Jensen RA (1985) *Enzymatic and nonenzymatic dehydration reactions of L-arogenate*. Biochemistry, 24:1607-1612.

Zhao Y (2012). *Auxin Biosynthesis: A Simple Two-Step Pathway Converts Tryptophan to Indole-3-Acetic Acid in Plants*. Mol Plant, 5(2):334-338.

III. OBJECTIVES

The general aim of this thesis manuscript is to better understand the biosynthesis of Phe in plants, and how the availability of this amino acid affects secondary metabolism, growth and development. The specific objectives are:

1. **To determine whether phenylalanine can be synthesized *in planta* using phenylpyruvate instead of arogenate**. So far, there is limited evidence about the participation of the "phenylpyruvate branch" in the biosynthesis of phenylalanine in plants. The possibility of synthesizing phenylalanine through two different branches of the prephenate pathway could be an additional regulatory level, immediately before the commitment step in the biosynthesis of phenylpropanoids.

2. **To identify the molecular basis determining why some arogenate dehydratases can also exhibit prephenate dehydratase activity**. Previous reports have shown that certain arogenate dehydratases from plants exhibit prephenate dehydratase activity *in vitro*. Nevertheless, the structural basis of this feature is unknown.

3. **To shed light on the biosynthesis of phenylalanine in conifers and its regulation, using *Pinus pinaster* as model organism**. Conifers exhibit a highly complex secondary metabolism derived from aromatic amino acids, and therefore these trees are interesting models for studying the regulation of these pathways. Woody plants such as conifers direct large amounts of organic carbon into the biosynthesis of phenylpropanoids, mostly lignin, during wood formation. Given the pivotal branch point position of arogenate dehydratases between phenylalanine and tyrosine biosynthesis, and the large number of isoforms found in plants, we hypothesize that these enzymes are a good candidates as potential regulators of the phenylalanine pathway.

IV. RESULTS

Paper 1

Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism

Fernando de la Torre, Jorge El-Azaz, Concepción Ávila, and Francisco M. Cánovas

Plant Physiology, January 2014. 164: 92-104.

Summary

Chloroplasts and plastids of nonphotosynthetic plant cells contain two aspartate (Asp) aminotransferases: a eukaryotic type (Asp5) and a prokaryotic-type bifunctional enzyme displaying Asp and prephenate aminotransferase activities (PAT). We have identified the entire Asp aminotransferase gene family in *Nicotiana benthamiana* and isolated and cloned the genes encoding the isoenzymes with plastidic localization: *NbAsp5* and *NbPAT*. Using a virus-induced gene silencing approach, we obtained *N. benthamiana* plants silenced for *NbAsp5* and/or *NbPAT*. Phenotypic and metabolic analyses were conducted in silenced plants to investigate the specific roles of these enzymes in the biosynthesis of essential amino acids within the plastid. The *NbAsp5* silenced plants had no changes in phenotype, exhibiting similar levels of free Asp and glutamate as control plants, but contained diminished levels of asparagine and much higher levels of lysine. In contrast, the suppression of *NbPAT* led to a severe reduction in growth and strong chlorosis symptoms. *NbPAT* silenced plants exhibited extremely reduced levels of asparagine and were greatly affected in their phenylalanine metabolism and lignin deposition.

Furthermore, *NbPAT* suppression triggered a transcriptional reprogramming in plastid nitrogen metabolism. Taken together, our results indicate that *NbPAT* has an overlapping role with *NbAsp5* in the biosynthesis of Asp and a key role in the production of phenylalanine for the biosynthesis of phenylpropanoids. The analysis of *NbAsp5/NbPAT* cosilenced plants highlights the central role of both plastidic aminotransferases in nitrogen metabolism; however, only *NbPAT* is essential for plant growth and development.

INTRODUCTION

Inorganic nitrogen is primarily assimilated into the amino acids Gln, Glu, Asn, and Asp; these four amino acids serve as nitrogen metabolic precursors and nitrogentransport compounds in most crops and higher plants (Buchanan et al., 2000). Glu is synthesized in the plastids and occupies a central position in plant amino acid metabolism, regulation, and signaling (Forde and Lea, 2007). Glu is the nitrogen donor for the biosynthesis of essential amino acids through the Asp and aromatic amino acid pathways, which are entirely located in plastids (Lea and Azevedo, 2003; Rippert et al., 2009; **Figure 1**).

The enzyme aspartate aminotransferase (AAT; EC 2.6.1.1) catalyzes a reversible transamination between Glu and oxaloacetate to yield Asp and 2-oxoglutarate and plays a crucial role channeling nitrogen from Glu to Asp in all living organisms. In plants, AAT is present as several isoenzymes located in different subcellular compartments: the cytosol, mitochondria, peroxisomes, and plastids (Ireland and Joy, 1985). In *Arabidopsis* (*Arabidopsis thaliana*), a small gene family encodes AAT isoenzymes: *Asp2*

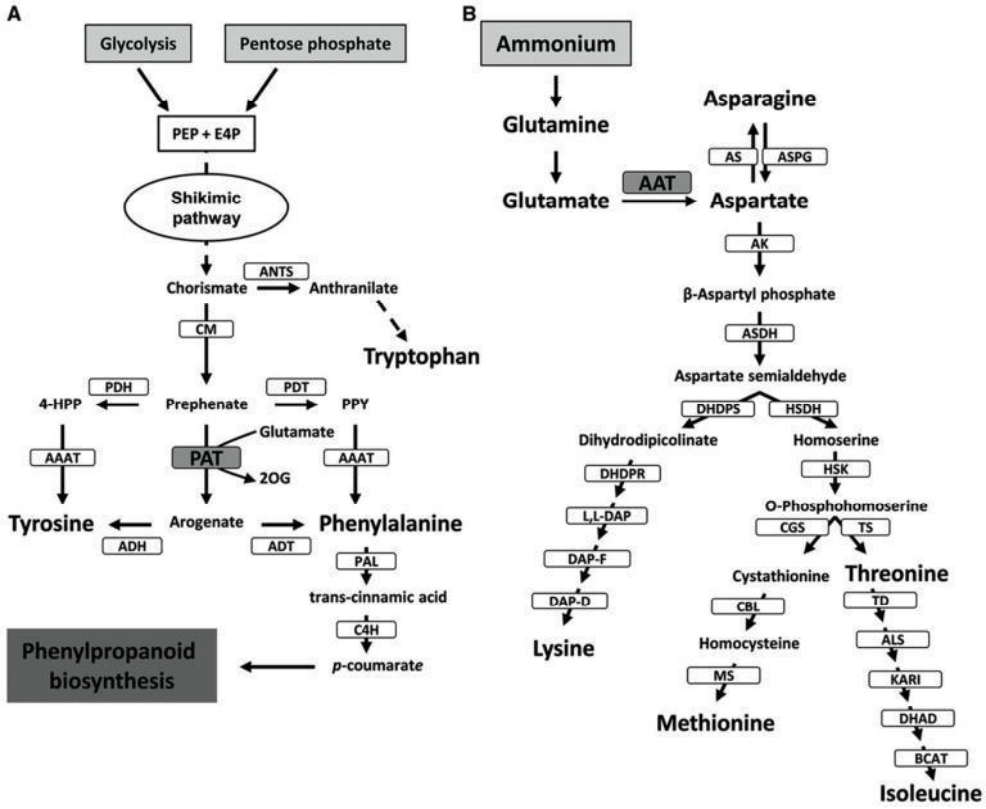


Figure 1. Asp and aromatic amino acid biosynthetic pathways in plants. **(A)** Postchorismate pathway leading to the biosynthesis of aromatic amino acids. CM, Chorismate mutase; E4P, D-erythrose 4-phosphate; 4-HPP, 4-hydroxyphenylpyruvate; PEP, phospho*enol*pyruvate; PPY, phenylpyruvate. **(B)** Asp-derived amino acid biosynthesis. ALS, Acetolactate synthase; AS, Asn synthetase; ASN, asparaginase; BCAT, branched-chain aminotransferase; CBL, cystathionine β-lyase; CGS, cystathionine γ-synthase; DAP-D, diaminopimelate decarboxylase; DAP-F, diaminopimelate epimerase; DHAD, dihydroxy acid dehydratase; DHDPR, dihydrodipicolinate reductase; DHDPS, dihydrodipicolinate synthase; HSDH, homoserine dehydrogenase; HSK, homoserine kinase; KARI, ketol-acid reductoisomerase; MS, Met synthase; TD, Thr deaminase; TS, Thr synthase

and *Asp4* (cytosolic), *Asp1* (mitochondrial), *Asp3* (peroxisomal), and *Asp5* (plastidial; Schultz and Coruzzi, 1995; Wilkie and Warren, 1998). The function of cytosolic and plastidic AAT has been investigated using

Arabidopsis mutants (Schultz et al., 1998; Miesak and Coruzzi, 2002). A specific nonredundant role has been proposed for *Asp2* related to the biosynthesis of a specific Asp pool during the light phase that is used to produce Asn during the dark phase (Miesak and Coruzzi, 2002). The Asp2 gene has also been proposed to interact with plant defense responses in [i]Arabidopsis[i] plants infected with the necrotrophic pathogen Botrytis cinerea (Brauc et al., 2011). *Asp5* mutants affected in the plastidic isoenzyme had no visible phenotype except that plants contained increased Gln levels either under light or dark growth conditions (Schultz et al., 1998; Miesak and Coruzzi, 2002). According to this, *Asp5* might have a role in shuttling reducing equivalents, as proposed for the peroxisomal and mitochondrial isoenzymes (Liepman and Olsen, 2004).

In addition, de la Torre et al. (2006) reported the existence of a novel plastid-located AAT, a prokaryotic-type AAT unrelated to other eukaryotic AATs from plants and animals but closely related to cyanobacterial enzymes. The kinetic parameters of this enzyme were determined, highlighting its high affinity for Glu and Asp compared with other plant AATs (de la Torre et al., 2006). Recently, Graindorge et al. (2010) and Maeda et al. (2011) independently discovered that plant prokaryotic-type AATs displayed not only AAT activity but also prephenate aminotransferase (PAT) activity. PAT activity catalyzes the reversible transamination between Glu/Asp and prephenate to yield 2-oxoglutarate/oxaloacetate and arogenate. This metabolic reaction represents a key step in the biosynthesis of aromatic amino acids in the plastids. The product of PAT, arogenate, is the immediate precursor for the biosynthesis of Phe and Tyr using arogenate dehydratase (ADT) and arogenate dehydrogenase (ADH), respectively (Tzin and Galili, 2010; Maeda and Dudareva, 2012; **Figure 1**). Recent genetic evidence has indicated that the arogenate pathway is the

predominant route for Phe biosynthesis in plants (Maeda et al., 2010; Maeda and Dudareva, 2012). In maritime pine (*Pinus pinaster*), a Myb transcription factor (Myb8) regulates genes involved in the Phe pathway (Craven-Bartle et al., 2013).

It is important to remark that in the presence of Glu as nitrogen donor, PAT exhibits similar values of the specificity constant for oxaloacetate and prephenate, indicating that the enzyme can operate both as a PAT and as a classical AAT (Graindorge et al., 2010). Consequently, PAT could be involved not only in the biosynthesis of aromatic amino acids but also in the biosynthesis of Asp-derived amino acids. Asp is a precursor in two main pathways: (1) the biosynthesis of Asn in the cytosol (Lea and Azevedo, 2003; Lea et al., 2007); and (2) the biosynthesis of Lys, Met, Thr, and Ile through the socalled Asp metabolic pathway in the plastids (Azevedo et al., 2006; Jander and Joshi, 2009; Galili, 2011; **Figure 1**).

In regard to PAT function, silencing of the gene encoding the enzyme was achieved in petals of petunia (*Petunia hybrida*) following an RNA interference (RNAi) strategy using a petal-specific promoter (Maeda et al., 2011). The targeted tissue was found to be impaired for Phe biosynthesis. However, the phenotypic and metabolic effects of *PAT* gene suppression in a whole plant were still unknown.

In this work, we report a new strategy to study the biochemical function of AAT and PAT activities in plastids, virus-induced gene silencing (VIGS), as a viable alternative experimental approach to obtain gene silencing in whole *Nicotiana benthamiana* plants. This strategy provided us with an efficient tool to study the in vivo role of the two plastidial aminotransferases encoded by *NbAsp5* and *NbPAT*, respectively. Molecular and metabolic

analysis of silenced plants revealed their contributions to different pathways of amino acid biosynthesis within the plastids.

RESULTS

Identification of the AAT Gene Family in *N. benthamiana*

A search in the EST databases of *N. benthamiana* (www. sgn.cornell.edu) allowed us to identify the complete AAT gene family. The *N. benthamiana* AAT gene family is similar to the well-characterized AAT gene family in *Arabidopsis* and includes five genes encoding eukaryotictype AATs (*NbAsp1-NbAsp5*), predicted to be targeted to different subcellular compartments (mitochondria, cytosol, plastids, and peroxisomes) and an extra gene encoding a prokaryotic-type bifunctional PAT/AAT enzyme targeted to plastids (*NbPAT*). Therefore, *N. benthamiana* contains two enzymes located within the plastids housing AAT activity, NbAsp5 and NbPAT. These AAT proteins showed a high degree of identity with the corresponding counterparts from *Arabidopsis* (Supplemental Table S1).

VIGS of *NbAsp5* and *NbPAT*

We developed a VIGS approach in order to study the phenotypic and metabolic impact associated with the loss of function of NbAsp5 and/or NbPAT proteins. The starting hypothesis was that these two enzymes are responsible for Asp and prephenate biosynthesis within the plastid. Using reverse transcription-PCR with primers based on *NbAsp5* and *NbPAT*, we obtained and sequenced two complementary DNA (cDNA) fragments 469 and 440 bp in length, respectively, corresponding to specific regions in the

59 open reading frame of both genes (Supplemental Figure S1). We named these fragments *Nb5'Asp5* and *Nb5'PAT*, respectively. These sequences had been previously selected as unique sequences, since in BLAST searches no other sequences were detected in the latest genome assembly of *N. benthamiana* that contained stretches longer than 21 to 24 nucleotides displaying 100% identity. This indicates the unlikelihood of cosilencing genes other than *Asp5* and *PAT* using *Nb5'Asp5* or *Nb5'PAT* in the VIGS assay. Fragments were inserted into pTRVGW, a Gateway-compatible tobacco rattle virus vector (Liu et al., 2002), courtesy of Savithramma P. Dinesh-Kumar and Dr. Olga del Pozo (pTRV-*Asp5* and pTRV-*PAT*). In order to substantiate the specificity of their silencing phenotypes, we generated additional nonoverlapping silencing constructs for *NbAsp5* (pTRV-*Nb3'Asp5*) and NbPAT (pTRV-*NbMPAT* and pTRV-*Nb3'PAT*; Supplemental Figure S1). Furthermore, *Nb5'Asp5* and *Nb5'PAT* fragments were PCR combined in a single DNA fragment, *NbAsp5/NbPAT*, and Gateway inserted into pTRVGW in a DNA construct for VIGS of both genes at the same time, pTRV-*Asp5/PAT* (Supplemental Figure S1). Thus, *N. benthamiana* plantlets were silenced for *NbAsp5*, *NbPAT*, or both using pTRV-*Asp5*, pTRV-*PAT*, or pTRV-*Asp5/PAT*, respectively (**Figure 2**). One month after initiation of VIGS, the morphology of the plants was observed and compared with the wild type and plants transfected with the empty vector, pTRV-EV. Silencing of the endogenous phytoene desaturase gene, *NbPDS*, was used as a control for the effectiveness of VIGS (Supplemental Figure S2). When the *NbAsp5* gene was targeted, no evident phenotype was detected in plants silenced using either pTRV-*Asp5* or pTRV-*Nb3'Asp5*. In contrast, plants silenced for *NbPAT* using pTRV-*PAT*, or

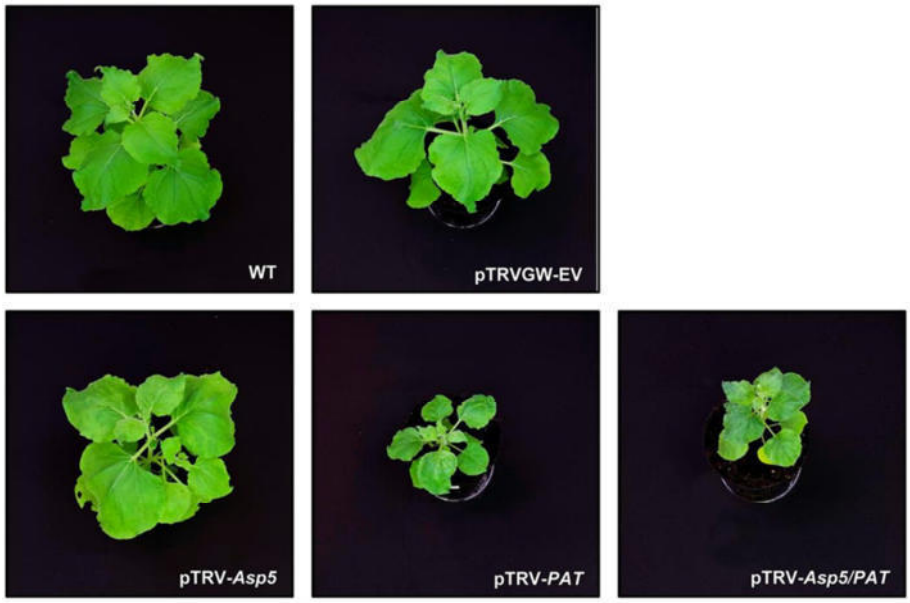


Figure 2. Phenotypes of *N. benthamiana* plants subjected to VIGS of *NbAsp5* and/or *NbPAT* genes. Photographs of wild-type (WT) and control (pTRVGW-EV) plants are shown in the top row. Phenotypes of 6-week-old plants silenced for *NbAsp5*, *NbPAT*, or *NbAsp-5/NbPAT* are shown in the bottom row. Similar phenotypes were observed in six different experiments, each with five biological replicates. All photographs were taken under the same magnification.

cosilenced for both *NbPAT* and *NbAsp5* using pTRV-*Asp5/PAT*, exhibited severe reduction in growth, strong symptoms of chlorosis, and altered fresh weight-to-dry weight ratio in comparison with pTRV-EV or wild-type plants (**Figure 2**; Supplemental Figures S3 and S4). Similar effects were observed in plants silenced using alternative constructs for VIGS of *NbPAT*: pTRV-*NbMPAT* and pTRV-*Nb3'PAT* (data not shown).

The degree of silencing observed was initially assessed by quantitative PCR (qPCR) determination of transcript abundance (**Figure 3**). The

transcript levels of both targeted genes, *NbAsp5* (**Figure 3A**) and *NbPAT* (**Figure 3B**), were reduced approximately 90% both in single- and double silenced *N. benthamiana* plants when compared with wild-type or pTRV-EV plants. **Figure 3** also shows that the silencing of a single plastidic aminotransferase, either *NbAsp5* or *NbPAT*, did not affect the transcript abundance of the other enzyme in the same plant. We also examined whether silencing of plastidic enzymes could alter the expression of other members of the gene family. As shown in Supplemental Figure S5, *NbAsp5* and/or *NbPAT* silencing did not affect the transcript abundance of *Asp1*, *Asp2*, *Asp3*, and *Asp4*. The specific silencing of *NbAsp5* and *NbPAT* was further analyzed by AAT activity determination on native gels. Our results clearly confirm effective suppression of the corresponding AAT activity bands in plants silenced for either *NbAsp5* or *NbPAT* (**Figure 3**). Moreover, when we conducted a similar experiment to test PAT activity on native gels, we observed complete suppression of the PAT activity band in plants silenced for *NbPAT* using both pTRV-*PAT* or the chimera containing both AAT sequences, pTRV-*Asp5/PAT* (**Figure 3B**). The absence of the NbPAT polypeptide in these samples was further confirmed by western-blot analysis (**Figure 3B**) using specific antibodies raised against the maritime pine protein (de la Torre et al., 2006).

Silencing of NbPAT Affects Lignin Deposition in Vascular Bundles

Recently, the arogenate pathway has been proposed to be the predominant route for Phe biosynthesis in plants (Maeda et al., 2011). Considering that the phenylpropane skeleton required for lignin biosynthesis is provided by the deamination of Phe, and the phenotype observed, we conducted an experiment in order to evaluate variations in lignin content in the *NbPAT* silenced plants (**Figure 4**). The determination of lignin content showed a

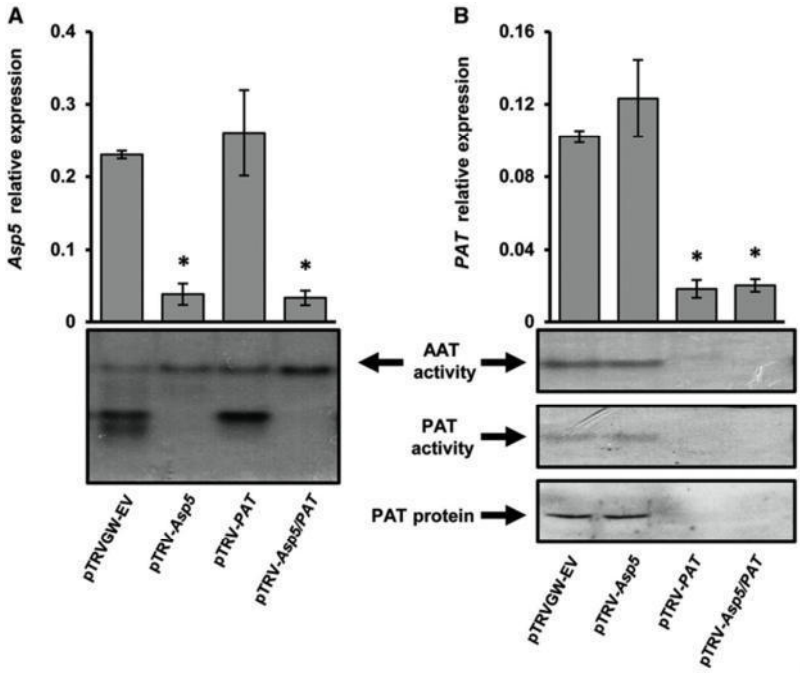


Figure 3. Molecular characterization of VIGS of *NbPAT* and/or *NbAsp5* genes in *N. benthamiana* leaves. **(A)** Relative *NbAsp5* gene expression in silenced plants compared with controls. At bottom, AATactivity is shown on native gels with protein extracts prepared from control and silenced plants. **(B)** Relative *NbPAT* gene expression in control and silenced plants. At bottom, AAT activity and PAT activity are shown in protein extracts prepared from control and silenced plants resolved by native gel electrophoresis; PAT protein levels were analyzed by western blotting using anti-PAT specific antibodies (de la Torre et al., 2006). The expression levels for all genes were normalized to that of *NbActin2*. Error bars represent SE. Asterisks indicate significant differences compared with control plants. AAT activity corresponding to NbASP5 and NbPAT on the gels was determined using extracts from *N. benthamiana* leaves (Fig. 2) according to a previously described protocol (de la Torre et al., 2007).

significant reduction compared with controls in leaves and stems of 6-week-old plants silenced for *NbPAT* (**Figure 4A**). A similar reduction was observed in plants silenced using pTRV-*Asp5/PAT* (**Figure 4A**). These results were further confirmed by *in situ* lignin detection in stem sections of plants silenced for *NbAsp5*, *NbPAT*, or both *NbPAT* and *NbAsp5*. The plants silenced for *NbPAT* or *NbPAT* and *NbAsp5* showed greatly reduced phloroglucinol staining in comparison with control plants or plants silenced only for *NbAsp5* (**Figure 4B**).

NbPAT silencing also affects the chlorophyll content

Concomitantly, and based on the observed chlorosis phenotype of plants silenced for *NbPAT* (**Figure 2**; Supplemental Figure S3), chlorophyll contents were also determined in the leaves of control and silenced plants. A significant reduction (30%-40%) in both chlorophyll *a* and *b* contents was observed in leaves of plants silenced for *NbPAT* and *Asp5/PAT* compared with control plants (**Figure 5**). No comparable alteration in chlorophyll a or b could be detected in plants silenced for *NbAsp5* (**Figure 5**).

NbAsp5 and *NbPAT* silencing alter carbohydrate levels

Since the silencing of PAT activity affected chlorophyll levels, we next examined the relative abundance of Glc, Fru, Suc, and the polysaccharides cellulose and starch. Major changes in free sugar levels and starch were observed in co-silenced plants; meanwhile, no significant changes were observed in cellulose content (Supplemental Figure S6).

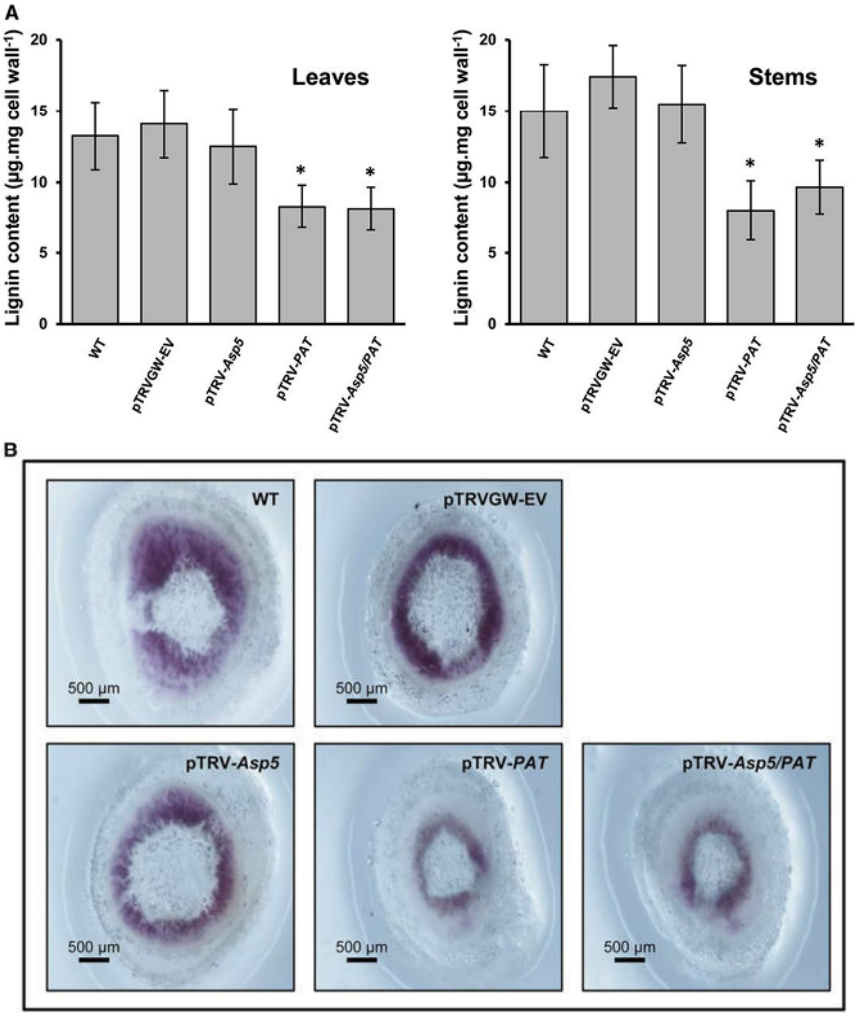


Figure 4. Lignin content and histological lignin staining in *N. benthamiana* plants. **(A)** Quantification of lignin in leaves (left) and stems (right) from *N. benthamiana* plants. WT, Wild type; pTRVGW-EV, control vector; pTRV-*Asp5*, silenced *NbAsp5*; pTRV-*PAT*, silenced *NbPAT*; pTRV-*Asp5/PAT*, cosilenced *NbAsp5/NbPAT*. Six independent biological replicates were measured. Average values and SD were calculated. Error bars represent SE, and asterisks indicate significant differences compared with control plants. **(B)** Histological detection of lignin in fresh-cut stem sections of *N. benthamiana* corresponding to control and silenced plants. Red-violet color shows the reaction of phloroglucinol-HCl with cinnamaldehyde end groups of lignin.

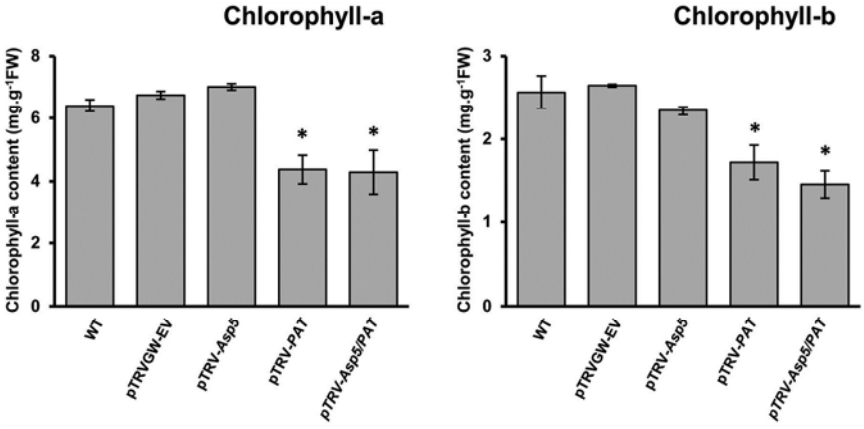


Figure 5. Chlorophyll content in leaves of *N. benthamiana* plants. Chlorophyll *a* and *b* contents in *N. benthamiana* leaves were determined as described in "Materials and Methods." For each measurement, the average value and SD were calculated using five independent biological replicates. Error bars represent SE. Asterisks indicate significant differences compared with control plants. FW, Fresh weight; WT, wild type.

Suppression of AAT activity in plastids of *N. benthamiana* results in altered amino acid profiles

To get further insights into the metabolic role of plastidic AAT enzymes, we next examined the profiles of a set of amino acids. We first analyzed whether the silencing of *NbPAT* and/or *NbAsp5* affected levels of key amino acids, Asp, Asn, Glu, and Gln (**Figure 6A**). No significant changes were observed in the levels of Asp or Glu in plants silenced for either *NbAsp5* or *NbPAT*. In contrast, plants silenced for both genes showed a dramatic reduction in the content of Asp and a concomitant increase in free Glu levels (approximately 5-fold; **Figure 6A**). In addition, plants silenced alternately for *NbAsp5* or *NbPAT* showed significant reductions in the levels of free Asn, close to 55% and 65%, respectively (**Figure 6A**). As observed

with Asp, the silencing of both genes simultaneously resulted in the additive reduction of the level of the amino acid of around 83% (**Figure 6**). When we measured free Gln levels, only a small reduction was detected in plants silenced for *NbAsp5* or *NbAsp5/NbPAT*. Absence of a major disturbance in Gln levels in plants silenced for *NbAsp5* and/or *NbPAT* is consistent with the observed corresponding levels of Glu that guarantee its biosynthesis through the action of Gln synthetase.

In order to investigate a putative role of *NbPAT* and/or *NbAsp5* in providing Asp for the biosynthetic pathways of Asp-derived amino acids in plastids, we also determined the levels of Lys, Thr, and Ile in the leaves of plants silenced for both aminotransferases (**Figure 6B**). Lys levels were dramatically increased in plants silenced for *NbAsp5* or *NbAsp5/NbPAT* when compared with control plants. However, Lys levels in plants silenced for *NbPAT* were close to those observed in control plants (**Figure 6B**). **Figure 6B** also shows that the levels of Thr and Ile in plants silenced alternately for *NbPAT* or *NbAsp5* were similar to those observed in control plants. Nevertheless, the levels of both amino acids clearly increased (approximately 2-fold for Thr and 3-fold for Ile) in plants cosilenced for both aminotransferases when compared with control plants. To further study the impact of plastidic AAT silencing, the expression levels of genes encoding three key enzymes in the Asp metabolic pathway were determined: Asp kinase (AK1), Asp semialdehyde dehydrogenase (ASDH1), and LL-diaminopimelate aminotransferase (LL-DAP; Azevedo et al., 2006; Hudson et al., 2006). The relative abundance of AK1 transcripts was similar in all samples examined; however, the expression of ASDH1 and LL-DAP significantly increased in *NbPAT* and *NbAsp5/NbPAT* silenced plants (**Figure 7A**).

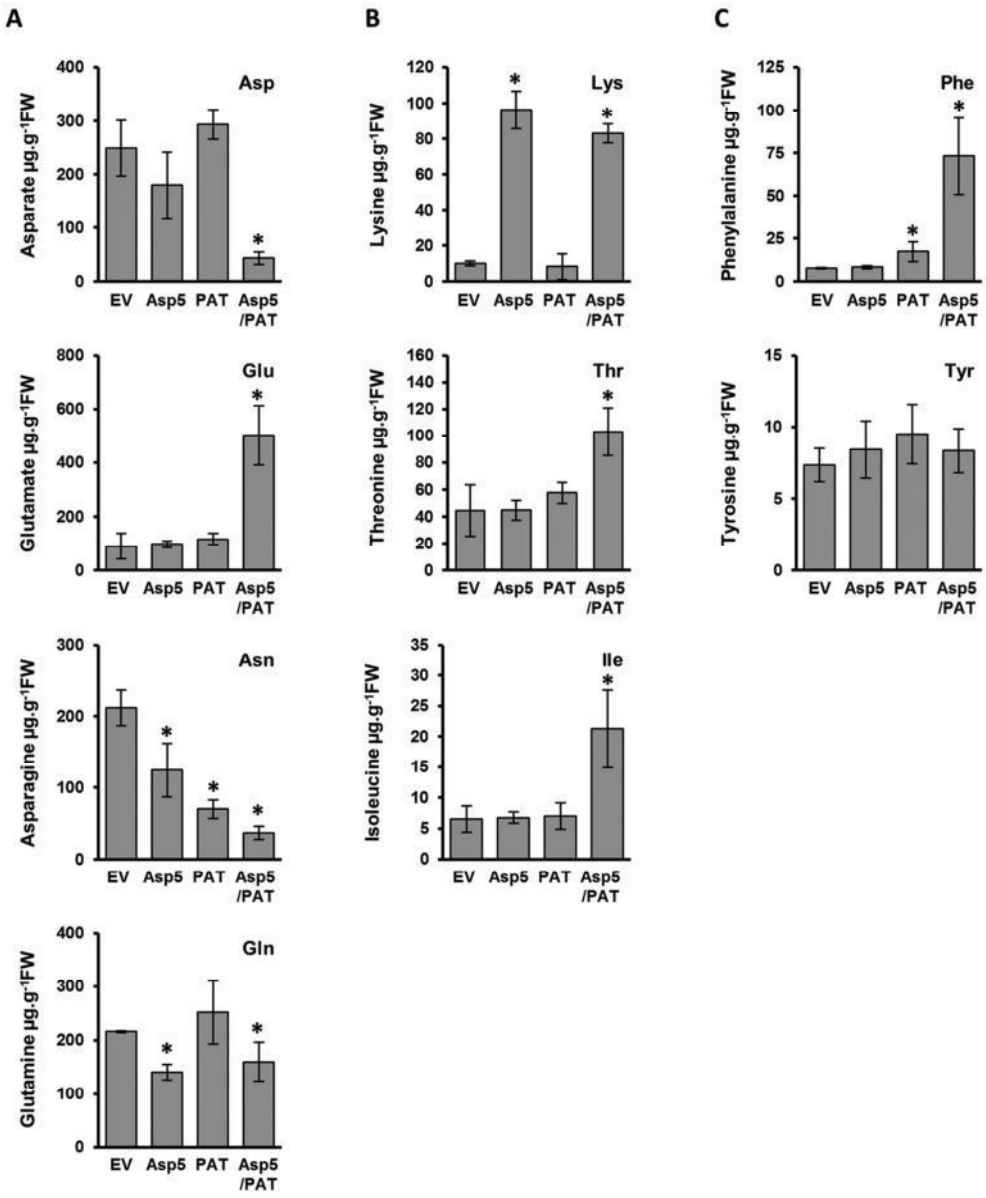


Figure 6. Amino acid content in plants silenced for *NbPAT* and/or *NbAsp5*. Amino acid profiles are shown in leaves from *N. benthamiana* silenced for *NbPAT*, *NbAsp5*, or both *NbPAT* and *NbAsp5* compared with empty-vector controls (EV). Amino acid content was determined using an HPLC method with a UV detector developed for the determination of individual amino acids. A, Asp, Glu, Asn, and Gln. B, Lys, Thr, and Ile. C, Phe and Tyr. Data represent means of five plants. Significant differences between the control and silenced plants based on Student's *t* test ($P \leq 0.05$) are indicated with asterisks. FW, Fresh weight.

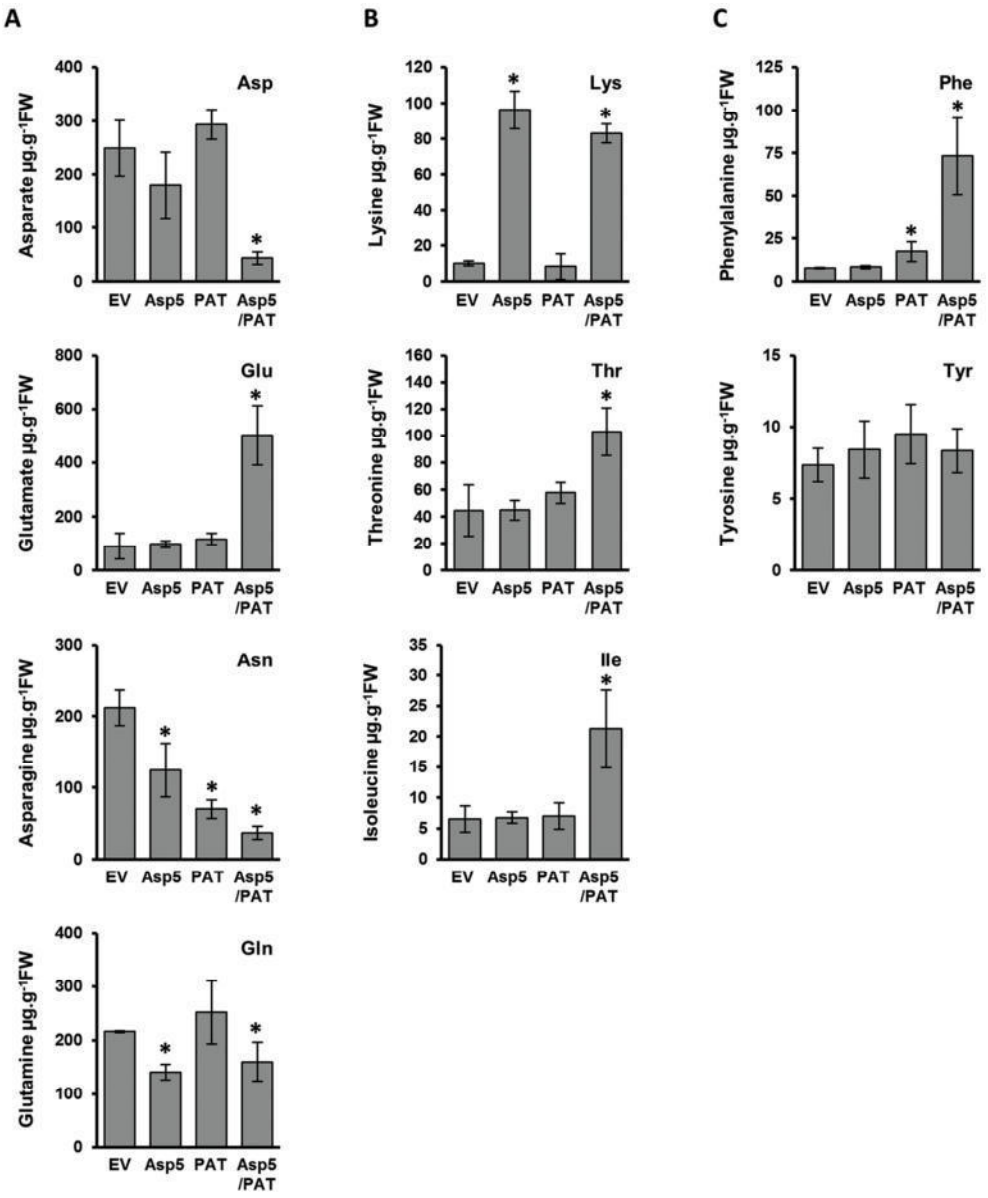


Figure 7. qPCR expression analysis in plants silenced for *NbPAT* and/or *NbAsp5*. Total RNA was extracted from *N. benthamiana* leaves as described. Transcript levels were determined by qPCR. **(A)** *NbAK1*, *NbASDH1*, and *NbLL-DAP*. **(B)** *NbPDH1* and *NbADH1*. C, *NbPDT1*, *NbPDT2*, and *NbAAAT*. **(D)** *NbPAL1*, *NbC4H*, and *NbANTS1*. The expression level for all genes was normalized to that of *NbActin2*. Values represent means of two assays of real-time qPCRanalysis, with four biological replicates in each. Error bars represent SE. Asterisks indicate significant differences compared with control plants.

Suppression of PAT activity in *N. benthamiana* leaves affects Phe metabolism

A database search analysis demonstrated the existence of a single locus coding for PAT in most plant genomes characterized (www.phytozome.net). Our VIGS-based approach revealed no detectable PAT activity in *N. benthamiana* leaves silenced for *NbPAT*, confirming that this enzyme is uniquely responsible for this activity in this tissue. In order to evaluate how the silencing of *Asp5* and *PAT* genes affects the profiles of aromatic amino acids, Phe and Tyr levels were compared in control and silenced plants (**Figure 6C**). No significant changes were observed in Phe levels in *NbAsp5* silenced plants (**Figure 6C**). In contrast, plants silenced for *NbPAT*, using either TRV-*PAT* or pTRV-*Asp5/PAT*, showed significant increases in the levels of Phe of 2- or 8-fold, respectively, compared with control plants. However, no significant alteration in Tyr levels was observed in plants silenced for *NbPAT* and/or *NbAsp5* (**Figure 6C**). This finding confirms the previous work of Maeda et al. (2011) reporting that the suppression of PAT in petals of petunia had no effect on the levels of free Tyr.

A putative alternate route for the biosynthesis of Phe could be the dehydration and decarboxylation of prephenate by arogenate/prephenate dehydratase (PDT) followed by transamination of phenylpyruvate to Phe using an aromatic-type aminotransferase. We have identified two full-length cDNAs encoding arogenate/PDTs expressed in *N. benthamiana* leaves, *NbPDT1* and *NbPDT2*, as well as a full-length sequence for *NbPPAT*, the homolog of the *Arabidopsis* gene *At5g36160*, which encodes the single aminotransferase described so far that is able to catalyze the transamination of phenylpyruvate to Phe (Prabhu and Hudson, 2010). We

conducted a set of qPCR measurements for transcripts encoded by *NbPDT1*, *NbPDT2*, and *NbPPAT* in plants silenced for *NbPAT* and/or *NbAsp5*. Our results showed a clear increase in the expression levels corresponding to these three genes in plants silenced for *NbPAT* using either pTRV-*PAT* or pTRV-*Asp5/PAT* (**Figure 7B**). No significant alteration in expression levels was detected for the same genes in plants silenced only for *NbAsp5* (**Figure 7B**). Subsequently, we also analyzed the expression levels of two genes encoding the first two enzymes downstream of PAT in channeling Phe toward secondary metabolites: Phe ammonia lyase (*NbPAL*) and cinnamate 4-hydroxylase (*NbC4H*). The results indicated that the expression of these genes was clearly reduced in plants silenced for *NbPAT* (**Figure 7C**), consistent with the observed reduction of lignin deposition in leaves and stems observed in plants silenced for *NbPAT* (**Figure 4**). We also analyzed the expression levels of *NbANTS1*, coding for ANTHRANILATE SYNTHASE1, the first enzyme in the biosynthesis of Trp from chorismate, which showed significant overexpression in plants silenced for *NbPAT* (**Figure 7C**).

Since Tyr levels were not altered significantly by silencing of *NbPAT*, we also analyzed the expression levels of *NbPDH1*, the gene encoding for PREPHENATE DEHYDROGENASE1, an enzyme that would be necessary in an alternate Tyr biosynthetic pathway using 4-hydroxyphenylpyruvate as intermediate. The qPCR analysis showed no significant alteration in *NbPDH1* expression levels (**Figure 7D**). Similarly, no alteration was detected in *NbADH1*, the gene encoding for AROGENATE DEHYDROGENASE1, the enzyme involved in the generation of Tyr from arogenate.

DISCUSSION

Plastids are subcellular organelles present in nearly all living plant cells and the exclusive site of many important biological processes, such as the major metabolic reactions involved in the biosynthesis of amino acids, including the Asp and aromatic amino acid pathways. In this paper, we analyze, to our knowledge for the first time, the effects on nitrogen metabolism following the suppression of plastidic AAT and PAT activities at the whole-plant level in *N. benthamiana* using VIGS of *NbPAT* and *NbAsp5* enzymes as an experimental approach. Our results clearly demonstrated that VIGS is an effective method to study the metabolic impact of Asp5 and PAT suppression *in planta* given the effective gene silencing observed, close to 90% (**Figure 3**).

Suppression of AAT activities in plastids of *N. benthamiana*

The analysis of the transcriptome of *N. benthamiana* allowed us to identify the complete AAT gene family, consisting of five genes encoding eukaryotic-type AAT (*NbAsp1-NbAsp5*) and a single gene encoding prokaryotictype AAT (*NbPAT*). The existence of two AATs within the plastids of *N. benthamiana* matches the previously described models in *Arabidopsis* (Schultz et al., 1998) and *Pinus pinaster* (de la Torre et al., 2007). Despite efforts made by several groups during the past two decades, the role of each member of the AAT gene family has not been sufficiently established. The presence of various isoenzymes located in different subcellular compartments is an argument against a complete functional redundancy and suggests the existence of functional specificities. In *Arabidopsis*, a role for *AtAsp2* was proposed in generating the bulk of Asp in the cytosol under illuminated conditions to be utilized

later for the synthesis of Asn in the dark (Miesak and Coruzzi, 2002). Moreover, in the same article, the characterization of transgenic lines in *Arabidopsis* deficient for the plastidic isoenzyme AtAsp5 was described, without identifying any phenotypic change. Similarly, we found no evident phenotypic changes when *NbAsp5* was suppressed in *N. benthamiana*, with no changes in growth or development, lignin deposition, or pigment contents (**Figures 2**, **4**, and **5**). In this work, we focus on the analysis of plastidial AAT activity in *N. benthamiana*, considering the contributions of both the NbAsp5 and NbPAT enzymes. Silencing of *NbPAT* resulted in strong growth reduction and leaf chlorosis (**Figure 2**; Supplemental Figure S3), consistent with a role of the enzyme in the biosynthesis of amino acids needed not only for protein synthesis but also to serve as precursors for a wide range of secondary metabolites. The phenotype observed in plants silenced for *NbPAT* was similar to that observed in plants silenced for both the *NbAsp5* and *NbPAT* genes (**Figure 2**), indicating a minor contribution, if any, of *NbAsp5* to the phenotype. In this respect, we observed that the levels of free Glu were not altered in plants individually silenced for *NbAsp5* or *NbPAT* but were greatly increased in plants silenced for both enzymes at the same time (**Figure 6**). Concomitantly, a decrease in free-Asp levels was observed in plants showing cosilencing of *NbAsp5* and *NbPAT*, indicating that the bulk of Asp within the leaf is only reduced when none of these plastidic AAT enzymes is active. These data strongly suggest the participation of NbPAT in Asp biosynthesis and further support that cytosolic AAT (NbAsp2) has a limited contribution to the observed levels of Asp. Interestingly, the individual silencing of either *NbAsp5* or *NbPAT* significantly reduced Asn levels even when Asp and Glu levels were not affected. It can be argued that NbAsp2 could be responsible for maintaining Asp levels in NbAsp5 silenced plants with a subsequent reduction in Asn biosynthesis. However, the strong reduction in Asn levels in *NbPAT*

silenced plants is not consistent with this possibility (**Figure 6**). Alternately, the above data may reflect a retention of Asp in the plastid to compensate a deficiency in its biosynthesis. Thus, our data suggest that when both AAT enzymes were blocked, Asn biosynthesis in the cytosol from Asp was greatly reduced. Gln levels did not show strong alterations in plants silenced for *NbAsp5* and/or *NbPAT*, which is consistent with a stable flux from available Glu to assimilate ammonium through Gln synthetase activity.

These findings strongly suggest that both enzymes play a partially redundant role in the plastids. Furthermore, these results clearly point out the physiological relevance of the AAT activity housed by *NbPAT* and could explain why the single suppression of the eukaryotic-type plastidial AAT enzyme in *N. benthamiana* (this work) or in *Arabidopsis* (Miesak and Coruzzi, 2002) produced no phenotypic effects or alterations in free Asp levels.

Lys levels clearly increased in plants silenced for NbAsp5 using pTRV-Asp5 or pTRV-*Asp5/PAT* but remained close to the control in plants silenced only for *NbPAT* (**Figure 6**). This profile was not observed for other amino acids of the Asp metabolic pathway, such as Thr and Ile. These results suggest an essential, and likely exclusive, role of the eukaryotic-type AAT (NbAsp5) in the homeostasis of Lys. Furthermore, our studies regarding the plastidial AAT contribution to the Asp metabolic pathway suggest that the observed phenotype in plants silenced for *NbPAT* could be explained at least in part by the suppression of its AAT activity and not only by the suppression of its PAT activity. The expression levels observed for key genes in the Asp metabolic pathway are consistent with this assumption (**Figure 7**). These results are compatible with a model previously proposed (de la Torre et al., 2009) according to which PAT

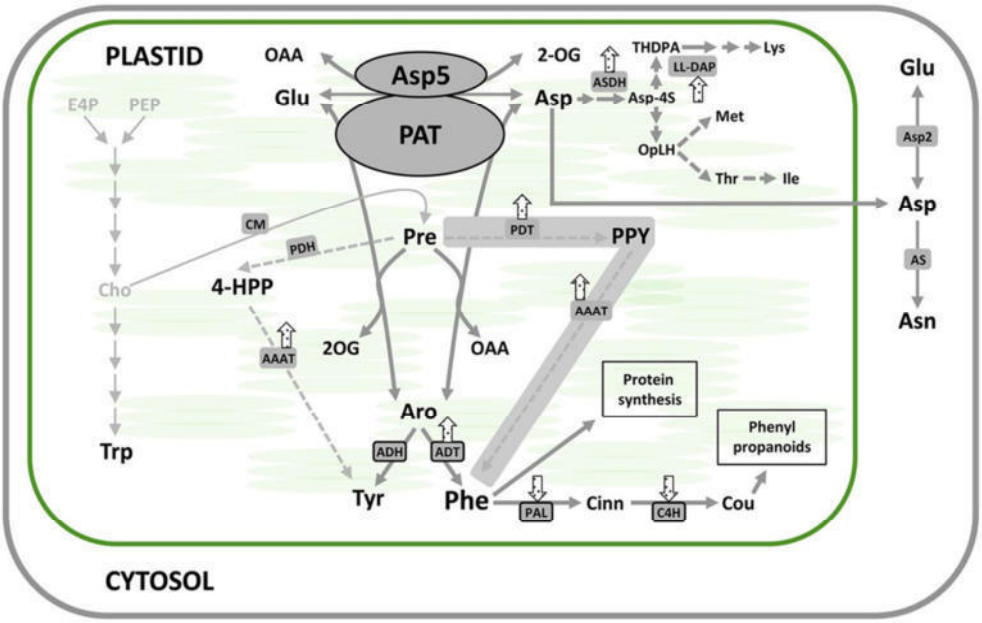


Figure 8. Schematic representation of major metabolic effects of plastidic AAT and PAT silencing. VIGS resulted in altered amino acid profiles and altered expression of genes encoding enzymes involved in amino acid biosynthesis. Arrows indicate transcriptional up- or down-regulation of the corresponding genes under *NbPAT* silencing in *N. benthamiana* leaves. The alternate pathway leading to the biosynthesis of Phe is shadowed in gray. AAAT, Aromatic amino acid aminotransferase; Aro, arogenate; AS, Asn synthetase; ASDH, Asp semialdehyde dehydrogenase; Asp-4S, L-Asp 4-semialdehyde; Cho, chorismate; Cinn, trans-cinnamate; CM, chorismate mutase; Cou, *p*-coumarate; E4P, D-erythrose 4-phosphate; 4-HPP, 4-hydroxyphenylpyruvate; OAA, oxaloacetate; 2-OG, 2-oxoglutarate; OpLH, *o*-phospho-L-homoserine; PEP, phospho*enol*pyruvate; PPY, phenylpyruvate; Pre, prephenate; THDPA, L-2,3,4,5-tetrahydrodipicolinate.

enzymes could also be involved in the flux toward the Asp-derived pathway (**Figure 8**). We cannot rule out that the metabolic effects generated by the suppression of PAT activity could be masking other effects corresponding to the suppression of the AAT activity housed by the same enzyme.

The observed changes in nitrogen metabolism and carbohydrate levels in *NbAsp5* and *NbPAT* cosilenced plants strongly suggest a deregulation of carbon/nitrogen balance triggered by suppression of the two plastidic enzymes.

Suppression of PAT activity in *N. benthamiana*

PATs are bifunctional enzymes competent to function not only as PATs but also as AATs in *Arabidopsis* (Graindorge et al., 2010; Maeda et al., 2011), petunia (Maeda et al., 2011), tomato (*Solanum lycopersicum*; Dal Cin et al., 2011), and *P. pinaster* (de la Torre et al., 2006). The analysis of recombinant PAT protein indicates that the enzyme exhibits a high affinity for substrates involved in both PAT and AAT activities. The functional analysis performed in *N. benthamiana* silenced plants demonstrates that the enzyme is also able to display both activities in planta (**Figure 3**). However, despite this knowledge of the biochemical characteristics and properties of the enzyme, its full metabolic role has not been fully elucidated. To a large extent, this is due to the lack of a system for efficient gene silencing in whole plants.

The difficulty in obtaining plants suppressed for the PAT enzyme was first supported by the work of Pagnussat et al. (2005), describing the identification of 130 transposon mutants of *Arabidopsis* with defects in female gametogenesis and embryo development. One of these lines, a mutant defective in the *At2g22250* locus encoding PAT, was affected in early embryo development. This phenotype was consistent with the suppression of an essential enzymatic activity required for the biosynthesis of amino acids (de la Torre et al., 2009). Several attempts were carried out in our laboratory to obtain transgenic *Arabidopsis* plants with suppressed

PAT gene expression using a targeted RNAi approach under the control of the constitutive 35S promoter, but all of them proved to be unsuccessful. PAT was locally silenced in petunia petals using an RNAi strategy, and the enzyme was strongly implicated in the biosynthesis of Phe (Maeda et al., 2011). However, analysis of PAT gene suppression in a whole plant is necessary to confirm the role of the enzyme without any masking effect provoked by the transfer of metabolites from other parts of the plant to offset the effect in a local tissue. In this work, the impact of silencing *NbPAT* was analyzed in a plantlet developmental stage to overcome its essential role during the embryonic phase.

Surprisingly, amino acid quantification registered a moderate increase in Phe levels in plants silenced for *NbPAT*. The increase in Phe was much more evident in plants silenced simultaneously for *NbPAT* and *NbAsp5*. Although silencing of *NbPAT* was not complete in our working model (close to 90%), these results appear inconsistent with a unique route for the synthesis of Phe *in planta*. Maeda et al. (2011), in petals of petunia silenced for PAT, observed similar levels of Phe to those measured in the same tissue in control plants. Both lines of evidence suggest the existence of an alternate source of Phe. An alternative pathway for the biosynthesis of Phe using phenylpyruvate as an intermediate has been proposed, a pathway known to exist in most microorganisms (**Figure 1**). In this paper, we present a set of results that together indicate the existence of transcriptional reprogramming in plants silenced for *NbPAT*, resulting, on the one hand, in the activation of genes encoding enzymes whose activities are required in an alternate pathway based on phenylpyruvate as an intermediate (*NbADT1*, *NbADT2*, and *NbAAAT*), and on the other hand, in reduced expression of genes encoding enzymes involved in the subsequent use of Phe toward compounds such as lignins or

phenylpropanoids, *NbPAL1* and *NbC4H* (**Figure 8**). *NbANTS1*, a gene encoding the first enzyme in the biosynthesis of Trp (Dubouzet et al., 2007), was also up-regulated in PAT suppressed plants, supporting a putative recycling of chorismate toward Trp in a physiological context with reduced demand for prephenate. These findings further confirm that Phe biosynthesis competes with Trp biosynthesis from their common precursor chorismate (Tzin et al., 2009). Overall, the reduction of PAT activity in these plants seems to be, at least partially, compensated by an alternate biosynthetic pathway and a reduced consumption toward secondary metabolism, probably ensuring a free-Phe reserve for the maintenance of basic functions such as protein synthesis. According to this model, reduced accumulation of lignin in the stems of plants silenced for *NbPAT* (**Figure 4**) would be a direct consequence of a prioritized use of the Phe pool toward protein biosynthesis (**Figure 8**). According to our model, the prephenate/arogenate pathway would be predominant in the biosynthesis of Phe, although the plants can use a functional alternate pathway using phenylpyruvate as an intermediate.

In summary, the roles of *NbAsp5* and *NbPAT* were investigated in *N. benthamiana* using a VIGS approach. Phenotypic and metabolic analyses were conducted in silenced plants to investigate the specific roles of these enzymes in the biosynthesis of essential amino acids within the plastids. Our results indicate that *NbPAT* has an overlapping role with *NbAsp5* in the biosynthesis of Asp and a key role in the production of Phe for the biosynthesis of phenylpropanoids.

MATERIALS AND METHODS

cDNA Cloning

Full-length *NbPAT* and *NbAsp5* cDNAs were isolated from total RNA extracted from *Nicotiana benthamiana* leaves by reverse transcription-PCR using iScript Reverse Transcription Supermix (Bio-Rad) with NbPAT-1/NbPAT-2 and NbAsp5-1/NbAsp5-2, respectively (Supplemental Table S1). cDNA fragments were subcloned into the pJET1.2 vector and completely sequenced.

VIGS

cDNA fragments *Nb5'PAT*, *NbMPAT*, *Nb3'PAT*, *NbAsp5*, and *Nb3'Asp-5* (Supplemental Figure S1) were amplified using specific Gateway attb adapter primers (Nb5'PAT-F/Nb5'PAT-R, NbMPAT-F/NbMPAT-R, Nb3'PAT-F/Nb3'PAT-R, NbAsp5-F/NbAsp5-R, and Nb3'Asp-5-F/Nb3'Asp-5-R; Supplemental Table S1), recombined into pDONR207 (Invitrogen) and then cloned into the VIGS Gateway-adapted destination vector pTRVGW. Constructs were confirmed by sequencing using a CEQ 8000 (Beckman Coulter España). Subsequently, *Agrobacterium tumefaciens* strain C58C1 was transformed by electroporation with the new constructs. For silencing in *N. benthamiana*, seedlings with two emerging leaflets (10 d old) were syringe infiltrated with cultures containing pTRVGW-derived constructs mixed with cultures containing pTRV1, both with an optical density at 600 nm of 0.1, according to procedures described previously (Liu et al., 2002). Silencing of the corresponding genes was assessed by qPCR using specific primers (Supplemental Table S2). Plants were harvested for analyses 4 weeks after VIGS induction. Leaves and stems from silenced

plants were immediately frozen in liquid nitrogen, ground in a mortar with pestle, and the resulting powder was stored at -80°C until use.

Expression Analysis

Total RNA was isolated following the protocol described by Liao et al. (2004), with minor modifications. The RNA samples were treated with RQ1 RNase-free DNase (Promega) to eliminate traces of genomic DNA, and cDNA synthesis was performed as described previously (Canales et al., 2012). qPCR was performed on a CFX-384 Real Time System(Bio-Rad) with SsoFast EvaGreen Supermix (Bio- Rad) under the following conditions: 95°C for 3 min (one cycle), and then 95°C for 1 s and 60°C for 5 s (45 cycles). After the final cycle, a melting curve analysis was performed over a temperature range of 65°C to 95°C in 0.5°C increments to verify the reaction specificity. cDNAs corresponding to 20 ng of reverse transcribed RNA were used as a template for each reaction (Supplemental Table S2). The raw fluorescence data from each reaction were fitted to the Mass Action Kinetic2 model, which requires no assumptions about the amplification efficiency of a qPCR assay (Boggy and Woolf, 2010). The initial target concentrations (D_0 parameter) for each gene were deduced from the Mass Action Kinetic2 model using the qpcR package for the R environment (Ritz and Spiess, 2008) and normalized to *NbActin2*.

In Gel AAT/PAT Activity

N. benthamiana tissues were frozen in liquid nitrogen, ground in a mortar and pestle, and the resulting powder was transferred into a tube containing extraction buffer consisting of 50 mM Tris-Cl, pH 7.5, 10% (v/v) glycerol, and 0.1% (v/v) Triton X-100. The samples were then subjected to native

PAGE in a discontinuous system with a 4% (w/v) stacking buffer (29:1 acrylamide:bisacrylamide and 125 mM Tris-Cl, pH 6.8) and a 7% (w/v) separating buffer (29:1 acrylamide: bisacrylamide and 375 mM Tris-Cl, pH 8.8) in a Mini-Protean Tetra Cell module (Bio-Rad). Staining for AAT activity was carried out using Fast Blue BB salt (Sigma), which turns blue in the presence of oxaloacetate, as described previously (Wendel and Weeden, 1989; de la Torre et al., 2009). Using similar procedures, oxaloacetate generated through PAT activity was also detected on native gels.

Determination of Photosynthetic Pigments

Chlorophylls *a* and *b* were extracted with 100% methanol from leaf discs of plants grown for 6 weeks under greenhouse conditions, and the content was determined according to Lichtenthaler and Wellburn (1983).

Lignin Staining and Quantification

Cell wall preparation and lignin quantification were performed following the method described by Lange et al. (1995). Briefly, tubes containing about 200 mg of frozen samples were treated with 1 mL of methanol, shaken for 1 h, and then centrifuged at 7,500*g* for 5 min. The supernatant was removed and the pellet treated successively with 1.5 mL of the following solvents: methanol, 1 M NaCl, 1% (w/v) SDS, deionized water, and chloroform:methanol (1:1). Finally, the pellet was resuspended in 1.5 mL of tert-butyl methyl ether and lyophilized. Ten milligrams of the resulting powder was resuspended in 1 mL of 2 M HCl and 0.2 mL of thioglycolic acid and incubated at 95°C for 4 h. The solution was centrifuged at 18,000 *g* for 10 min, and the resulting pellet was washed three times with deionized water, resuspended in 1.5 mL of 0.5 M NaOH, and shaken for 12

h. The samples were then centrifuged at 18,000 *g* for 20 min, and the supernatant was mixed with 0.3 mL of 12 M HCl and incubated for 4 h at 4°C. After centrifugation for 40 min, the resulting pellets were dried using a speed vacuum concentrator, resuspended in 1 mL of 0.5 M NaOH, and the A_{280} was determined and compared with a previously determined lignin standard made with pure lignin from Sigma-Aldrich. Histochemical analysis of the accumulated lignin was performed using phloroglucinol staining of hand-cut sections of *N. benthamiana* stems. Sections were incubated for 10 min in a 1% phloroglucinol solution in ethanol:HCl (2:1) and visualized by light microscopy using a Nikon Eclipse E 800 microscope.

Free Amino Acid Analysis

Frozen plant material (100 mg) was ground in a mortar and extracted in 50 mM Tris-HCl, pH 8, 2 mM EDTA, and 10 mm 2-mercaptoethanol. The extract was centrifuged at 20,000 *g* at 4°C, and the supernatant was recovered and centrifuged again as above. An aliquot of 400 mL was transferred to a fresh tube, to which 1 mL of methanol was added and mixed for 10 min at 4°C. After centrifugation as above, the supernatant was saved in a fresh tube and the pellet was resuspended in 200 mL of methanol:water (4:1, v/v), stirred for 10 min at 4°C, and centrifuged, and the supernatant was reserved. The last step was repeated twice, and all the reserved supernatants were combined. The volume was reduced to 200 mL by evaporation at 90°C in an oven. Finally, the extract was filtered through a 0.2-mm pore filter.

To determine the free amino acid pool content of the samples, amino acids were separated with no derivatization with a System Gold HPLC BioEssential high-performance liquid chromatograph (Beckman-Coulter)

using a lithium citrate buffer system, followed by a postcolumn ninhydrin reaction detection system. For the identification and quantification of amino acids, the corresponding standards were used.

Carbohydrate Analysis

Carbohydrates were extracted from frozen powder corresponding to *N. benthamiana* leaves with 80% ethanol at 80°C for 30 min, followed by further washing with 50% ethanol at 80°C for 20 min and two additional washes with water. Combined supernatants were centrifuged to remove debris and lyophilized. The resulting powder was resuspended in water, and Suc, Glc, and Fru were measured enzymatically following the reduction of NADP at 340 nm after successive additions of the coupling enzymes Glc-6-P dehydrogenase (4 units mL^{-1}), hexokinase (10 units mL^{-1}), phosphoglucose isomerase (5 units mL^{-1}), and invertase (Sekin, 1978). Starch was measured as Glc from the extracted pellet, following incubation at 37°C for 4 h with a-amylase (4 units mL^{-1}) and amyloglucosidase (8 units mL^{-1}). All enzymes were obtained from Roche Diagnostics. Cellulose content was determined by the anthrone method (Updegraff, 1969). Briefly, 100 mg of tissue powder was boiled in 1 mL acetic-nitric reagent (acetic acid:nitric acid:water, 8:1:2) for 30 min to remove lignin and hemicellulosic carbohydrates. After centrifugation at 4,500 *g* for 20 min, the supernatant was removed and the remaining material was washed twice with 1 mL of distilled water. The cellulose samples were then hydrolyzed in 67% (v/v) sulfuric acid for 1 h at 25°C, and the Glc content of the samples was determined as follows: 5 mL of the sulfuric acidhydrolyzed samples was mixed with 495 mL of water and 1 mL of 0.2% (w/v) anthrone in concentrated sulfuric acid on ice. The samples were boiled for 10 min, and then the absorbance was measured as optical density at 630 nm. Cellulose

content was determined based on a previously determined standard using commercial cellulose from Sigma-Aldrich.

Supplemental Data

The following materials are available in the online version of this article:

Supplemental Figure S1. Open reading frame cloning of *NbPAT*, *NbAsp5*, and derivative VIGS constructs.

Supplemental Figure S2. Phenotypes of *N. benthamiana* leaves subjected to VIGS of *NbPDS*.

Supplemental Figure S3. Phenotypes of *N. benthamiana* leaves subjected to VIGS of *NbPAT* and/or *NbAsp5* genes.

Supplemental Figure S4. Fresh weight-to-dry weight ratio in *N. benthamiana* plants silenced for *NbPAT* and/or *NbAsp5*.

Supplemental Figure S5. Expression analysis of the AAT gene family in plants silenced for *NbPAT* and/or *NbAsp5*.

Supplemental Figure S6. Analysis of carbohydrate content in *N. benthamiana* plants silenced for *NbPAT* and/or *NbAsp5*.

Supplemental Table S1. Amino acid sequence identity between *Arabidopsis* and *N. benthamiana* AAT.

Supplemental Table S2. Oligonucleotides used in this work.

Acknowledgements

We thank Savithramma P. Dinesh-Kumar (University of California [Davis]) and Dr. Olga del Pozo (Instituto Bioquímica Vegetal y Fotosíntesis, Consejo Superior de Investigaciones Científicas) for the pTRVGW silencing vector. We also thank Javier Canales (Universidad de Málaga) for technical advice with qPCR analysis.

References

Azevedo RA, Lancien M and Lea PJ (2006) *The aspartic acid metabolic pathway, an exciting and essential pathway in plants*. Amino Acids, 30:143-162.

Boggy GJ and Woolf PJ (2010) *A mechanistic model of PCR for accurate quantification of quantitative PCR data*. PLoS ONE, 5:e12355.

Brauc S, De Vooght E, Claeys M, Höfte M and Angenon G (2011) *Influence of over-expression of cytosolic aspartate aminotransferase on amino acid metabolism and defence responses against Botrytis cinerea infection in* Arabidopsis thaliana. J Plant Physiol, 168:1813-1819.

Buchanan BB, Gruissem W and Jones RL (2000) *Biochemistry and Molecular Biology of Plants*. American Society of Plant Physiologists, Rockville, MD.

Canales J, Rueda-López M, Craven-Bartle B, Ávila C and Cánovas FM (2012) *Novel insights into regulation of asparagine synthetase in conifers*. Front Plant Sci, 3:100.

Craven-Bartle B, Pascual MB, Cánovas FM and Ávila C (2013) *A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine*. Plant J, 74:755-766.

Dal Cin V, Tieman DM, Tohge T, McQuinn R, de Vos RC, Osorio S, Schmelz EA, Taylor MG, Smits-Kroon MT, Schuurink RC, et al (2011) *Identification of genes in the phenylalanine metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit*. Plant Cell, 23:2738-2753.

de la Torre F, De Santis L, Suárez MF, Crespillo R and Cánovas FM (2006) *Identification and functional analysis of a prokaryotic-type aspartate aminotransferase: implications for plant amino acid metabolism*. Plant J, 46:414-425.

de la Torre F, Suárez MF, De Santis L and Cánovas FM (2007) *The aspartate aminotransferase family in conifers: biochemical analysis of a prokaryotictype enzyme from maritime pine*. Tree Physiol, 27:1283-1291.

de la Torre F, Moya-García AA, Suárez MF, Rodríguez-Caso C, Cañas RA, Sánchez-Jiménez F and Cánovas FM (2009) *Molecular modeling and sitedirected mutagenesis reveal essential residues for catalysis in a prokaryotetype aspartate aminotransferase*. Plant Physiol, 149:1648-1660.

Dubouzet JG, Ishihara A, Matsuda F, Miyagawa H, Iwata H and Wakasa K (2007) *Integrated metabolomic and transcriptomic analyses of hightryptophan rice expressing a mutant anthranilate synthase alpha subunit*. J Exp Bot, 58:3309-3321.

Forde BG and Lea PJ (2007) *Glutamate in plants: metabolism, regulation, and signalling*. J Exp Bot, 58:2339-2358.

Galili G (2011) *The aspartate-family pathway of plants: linking production of essential amino acids with energy and stress regulation*. Plant Signal Behav, 6:192-195.

Graindorge M, Giustini C, Jacomin AC, Kraut A, Curien G and Matringe M (2010) *Identification of a plant gene encoding glutamate/aspartateprephenate*

aminotransferase: the last homeless enzyme of aromatic amino acids biosynthesis. FEBS Lett, 584:4357-4360.

Hudson AO, Singh BK, Leustek T and Gilvarg C (2006) *An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants*. Plant Physiol, 140:292-301.

Ireland RJ and Joy KW (1985) *Plant transaminases*. In: P Christen, DE Metzler, eds, *Transaminases, Vol 2*. Wiley, New York, pp. 376-384.

Jander G and Joshi V (2009) *Aspartate-derived amino acid biosynthesis in* Arabidopsis thaliana. The Arabidopsis Book, 7:e0121.

Lange BM, Lapierre C and Sandermann H Jr (1995) *Elicitor-induced spruce stress lignin: structural similarity to early developmental lignins*. Plant Physiol, 108:1277-1287.

Lea PJ and Azevedo RA (2003) *Primary products: plant amino acids*. In: B Thomas, SJ Murphy, BG Murray, eds, *Encyclopaedia of Applied Plant Sciences, Vol 3*. Elsevier, Amsterdam, pp. 871-883.

Lea PJ, Sodek L, Parry MAJ, Shewry PR and Halford NG (2007) *Asparagine in plants*. Ann Appl Biol, 150:1-26.

Liao Z, Chen M, Guo L, Gong Y, Tang F, Sun X and Tang K (2004) *Rapid isolation of high-quality total RNA from taxus and ginkgo*. Prep Biochem Biotechnol, 34:209-214.

Lichtenthaler HK and Wellburn AR (1983) *Determination of total carotenoids and chlorophylls a and b of leaf extracts in different solvents*. Biochem Soc Trans, 603:591-592.

Liepman AH and Olsen LJ (2004) *Genomic analysis of aminotransferases in* Arabidopsis thaliana. Crit Rev Plant, Sci 23:73-89.

Liu Y, Schiff M and Dinesh-Kumar SP (2002) *Virus-induced gene silencing in tomato*. Plant J, 31:777-786.

Maeda H and Dudareva N (2012) *The shikimate pathway and aromatic amino acid biosynthesis in plants*. Annu Rev Plant Biol, 63:73-105.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, Cooper BR, Rhodes D, Pichersky E and Dudareva N (2010) *RNAi suppression of Arogenate Dehydratase1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22:832-849.

Maeda H, Yoo H and Dudareva N (2011) *Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate*. Nat Chem Biol, 7:19-21.

Miesak BH and Coruzzi GM (2002) *Molecular and physiological analysis of Arabidopsis mutants defective in cytosolic or chloroplastic aspartate aminotransferase*. Plant Physiol, 129:650-660.

Pagnussat GC, Yu HJ, Ngo QA, Rajani S, Mayalagu S, Johnson CS, Capron A, Xie LF, Ye D and Sundaresan V (2005) *Genetic and molecular identification of genes required*

for female gametophyte development and function in Arabidopsis. Development, 132:603-614.

Prabhu PR and Hudson AO (2010) *Identification and partial characterization of an L-tyrosine aminotransferase (TAT) from* Arabidopsis thaliana. Biochem Res Int, 2010:549572.

Rippert P, Puyaubert J, Grisollet D, Derrier L and Matringe M (2009) *Tyrosine and phenylalanine are synthesized within the plastids in* Arabidopsis. Plant Physiol, 149:1251-1260.

Ritz C and Spiess AN (2008) *qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis*. Bioinformatics, 24:1549-1551.

Schultz CJ and Coruzzi GM (1995) *The aspartate aminotransferase gene family of* Arabidopsis *encodes isoenzymes localized to three distinct subcellular compartments*. Plant J, 7:61-75.

Schultz CJ, Hsu M, Miesak B and Coruzzi GM (1998) Arabidopsis *mutants define an* in vivo *role for isoenzymes of aspartate aminotransferase in plant nitrogen assimilation*. Genetics, 149:491-499.

Sekin S (1978) *Enzymatic determination of glucose, fructose and sucrose in tobacco*. Tobacco Science, 23:75-77.

Tzin V and Galili G (2010) *New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants*. Mol Plant, 3:956-972.

Tzin V, Malitsky S, Aharoni A, Galili G (2009) *Expression of a bacterial bifunctional chorismate mutase/prephenate dehydratase modulates primary and secondary metabolism associated with aromatic amino acids in* Arabidopsis. Plant J, 60:156-167.

Updegraff DM (1969) *Semimicro determination of cellulose in biological materials*. Anal Biochem, 32:420-424.

Wendel JF and Weeden NF (1989) *Visualization and interpretation of plant isoenzymes*. In: DE Soltis, PS Soltis, eds, *Isoenzymes in Plant Biology*. Dioscorides Press, Portland, OR, pp. 5-45.

Wilkie SE and Warren MJ (1998) *Recombinant expression, purification, and characterization of three isoenzymes of aspartate aminotransferase from* Arabidopsis thaliana. Protein Expr Purif, 12:381-389.

Paper 2

Identification of a small protein domain present in all plant lineages that confers high prephenate dehydratase activity

Jorge El-Azaz, Fernando de la Torre, Concepción Ávila and Francisco M. Cánovas

The Plant Journal. July, 2016. 87:215-229.

Summary

L-Phenylalanine serves as a building block for the biosynthesis of proteins, but also as a precursor for a wide range of plant-derived compounds essential for plants and animals. Plants can synthesize Phe within the plastids using arogenate as a precursor; however, an alternative pathway using phenylpyruvate as an intermediate, described for most microorganisms, has recently been proposed. The functionality of this pathway requires the existence of enzymes with prephenate dehydratase (PDT) activity (EC 4.2.1.51) in plants. Using phylogenetic studies, functional complementation assays in yeast and biochemical analysis, we have identified the enzymes displaying PDT activity in *Pinus pinaster*. Through sequence alignment comparisons and site-directed mutagenesis we have identified a 22-amino acid region conferring PDT activity (PAC domain) and a single Ala^{314} residue critical to trigger this activity. Our results demonstrate that all plant clades include PAC domain-containing ADTs, suggesting that the PDT activity, and thus the ability to synthesize Phe using phenylpyruvate as an intermediate, has been preserved throughout the evolution of plants. Moreover, this pathway together with the

arogenate pathway gives plants a broad and versatile capacity to synthesize Phe and its derived compounds. PAC domain-containing enzymes are also present in green and red algae, and glaucophytes, the three emerging clades following the primary endosymbiont event resulting in the acquisition of plastids in eukaryotes. The evolutionary prokaryotic origin of this domain is discussed.

INTRODUCTION

The biosynthesis of Phe in plants is an exceptionally important pathway because this amino acid is required for protein biosynthesis but also as a precursor for a wide range of metabolites that are essential for plants, such as lignin, suberin, flavonoids, isoflavonoids or anthocyanins. The absence of a Phe biosynthetic pathway in animals also makes this pathway an attractive target for herbicide design. The shikimate pathway located in plastids connects central carbon metabolism with the biosynthesis of aromatic amino acids, Phe, Tyr and Trp. Thereby, phospho*enol*pyruvate, from glycolisis, and erythrose-4-phosphate, from the pentose phosphate pathway, are converted to chorismate, a direct precursor for the biosynthesis of aromatic amino acids and other compounds, such as vitamins K1 and B9 (**Figure 1**; Tzin and Galili, 2010). Chorismate is alternatively used by anthranilate synthase, the first enzyme in the Trp biosynthetic pathway, or by chorismate mutase to generate prephenate for the biosynthesis of Phe and Tyr (Cotton and Gibson, 1965; Maeda and Dudareva, 2012). Two alternative routes of Phe biosynthesis have been reported: (i) in the arogenate pathway, prephenate is transaminated by prephenate-aminotransferase (PAT) to generate arogenate, which is then decarboxylated and dehydrated by arogenate dehydratase (ADT) to give

Phe (Bonner and Jensen, 1987a); (ii) in the phenylpyruvate pathway, present in most microorganisms, the decarboxylation and dehydratation of prephenate results in the generation of phenylpyruvate in a reaction catalyzed by prephenate dehydratase (PDT). Subsequently, phenylpyruvate is transaminated to Phe through a phenylpyruvate aminotransferase (Fischer and Jensen, 1987a; **Figure 1**). Prephenate and arogenate are also precursors for the biosynthesis of Tyr through the activity of prephenate dehydrogenase and arogenate dehydrogenase, respectively (**Figure 1**; Fischer and Jensen, 1987b; Bonner and Jensen, 1987b).

The arogenate pathway has been proposed as the main route of Phe biosynthesis in plants (Maeda et al., 2010, 2011); however, the detection of PDT activity in etiolated *Arabidopsis thaliana* seedlings (Warpeha et al., 2006) and recently in *Petunia hybrida* petals (Yoo et al., 2013) suggests the existence of a functional phenylpyruvate pathway in plants. Moreover, the exogenous supply of shikimate in *Petunia hybrida* petals silenced for ADT activity led to phenylpyruvate accumulation and a restoration of Phe levels, indicating the existence of a functional PDT pathway (Maeda et al., 2010). The precise identity of the enzyme/s responsible for this activity is unclear, but several studies have revealed that some ADTs can use prephenate at low catalytic efficiencies (Cho et al., 2007; Maeda et al., 2010; Bross et al., 2011). Recombinant expression of the *Arabidopsis* enzymes AtADT1 and AtADT2 reverts the growth-arrest phenotype induced by the deletion of the endogenous PDT activity in the yeast mutant *pha2* (Bross et al., 2011). In addition, Yoo et al. (2013) have described a cytosolic phenylpyruvate-aminotransferase in *Petunia hybrida* that catalyzes the efficient interconversion of phenylpyruvate and Tyr to 4-hydroxyphenylpyruvate and Phe. Through feeding experiments with labelled phenylpyruvate as well as

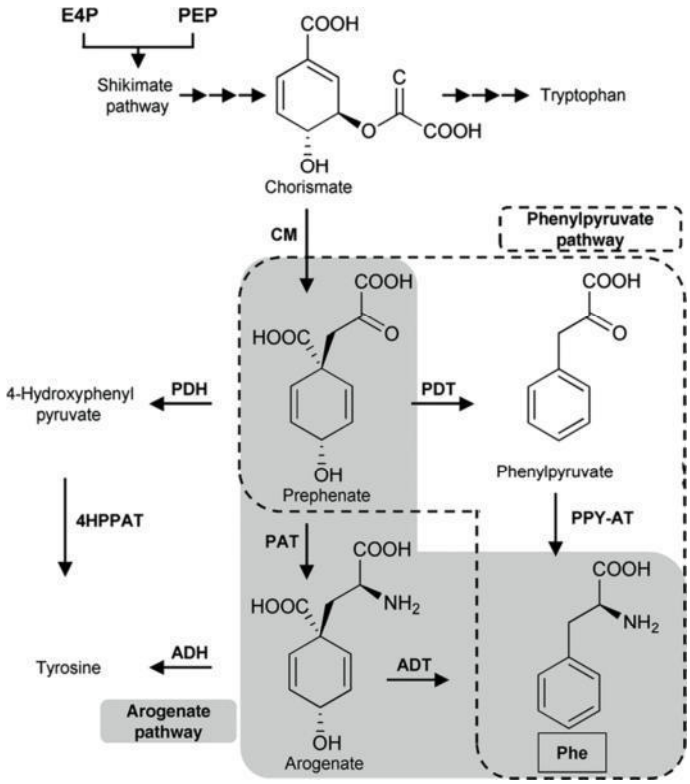


Figure 1. The phenylalanine biosynthetic pathways in plants. Shaded in gray, the arogenate pathway; encircled by dashed line, the phenylpyruvate pathway. Abbreviations: ADH, arogenate dehydrogenase; ADT, arogenate dehydratase; CM, chorismate mutase; E4P, erythrose 4-phosphate; 4HPPAT, 4-hydroxyphenylpyruvate aminotransferase; PAT, prephenate aminotransferase; PDH, prephenate dehydrogenase; PDT, prephenate dehydratase; PEP, phosphoenol pyruvate; and PPY-AT, phenylpyruvate aminotransferase.

silencing and overexpression analysis, these authors proposed the existence of a functional phenylpyruvate pathway that has enhanced flux when blocking the arogenate pathway. Consistently, the virus-induced gene silencing of PAT in *Nicotiana benthamiana* resulted in the transcriptional activation of two genes encoding for ADT/PDT enzymes (de la Torre et al., 2014a); however, no PDT active enzyme was assigned to this route in any of these studies.

In regard to the regulation of Phe biosynthesis in plants, only a minor proportion is currently known. Together, pathways for the biosynthesis of Phe, Tyr or Trp appear to be regulated developmentally (Benfey and Chua, 1989), and in response to biotic (Keith et al., 1991; Görlach et al., 1995) and abiotic stresses (Dyer et al., 1989; Keith et al., 1991; Janzik et al., 2005). The role of different transcription factors regulating the biosynthesis of aromatic amino acids has been recently reviewed by Maeda and Dudareva (2012).

As described for different microorganisms, the biosynthesis of Phe, Tyr and Trp is subject to multiple levels of post-transcriptional regulation in plants. The alternative metabolic bifurcation at chorismate level is allosterically controlled by the feedback inhibition of anthranilate synthase by Trp, or by the inhibition of chorismate mutase by Phe and Tyr (Romero et al., 1995; Bohlmann et al., 1996). The accumulation of Trp enhances the biosynthesis of Phe and Tyr as described in *Solanum tuberosum* (Kuroki and Conn, 1988). Both postchorismate pathways leading to the biosynthesis of Phe are also allosterically regulated through a feedback inhibition mechanism triggered by Phe upon ADT and PDT activities (Siehl and Conn, 1988; Rippert and Matringe, 2002; Yamada et al., 2008). This allosteric regulation is executed when Phe is bound to the C-terminal ACT (from aspartate kinase, chorismate mutase and TyrA) regulatory domain. This binding is suggested to induce a conformational change that results in reduced access to the active site and the subsequent blocking of its activity (Tan et al., 2008). Accordingly, free Phe levels are dramatically increased in *A. thaliana* (Huang et al., 2010) and Oryza sativa (rice; Yamada et al., 2008) plants carrying point mutations in the regulatory ACT domain of ADTs, which result in feedback-insensitive enzymes. In this study, we report the identification of enzymes displaying PDT activity using the conifer *Pinus*

pinaster as a model plant. Gymnosperms exhibit a highly complex secondary metabolism derived from aromatic amino acids, and therefore represent a very interesting model for studying these pathways (Warren et al., 2015). Using genetic, biochemical and molecular strategies we have identified the protein domain responsible for this activity, and have contributed to the understanding of the evolutionary history of Phe biosynthesis in plants.

RESULTS

Identification of the ADT gene family in *Pinus pinaster*

In the present study, we have identified sequences encoding putative ADT/PDTs in *Pinus pinaster*. Based on the SustainPine database (Canales et al., 2014; http:// www.scbi.uma.es/sustainpine), we identified nine members within this family: *PpADT-A* to *-I* (Appendix S1), which is consistent with the occurrence of eight and nine members in *Picea glauca* and *Pinus taeda*, respectively. There are more genes for ADT/PDTs in gymnosperms than in angiosperms (between three and six), lycophytes (three) or mosses (four). The corresponding cDNA sequences were isolated, and qPCR analysis performed in pine seedlings showed specific and diferentiated expression profiles of PpADT genes (Figure S1). N-terminal plastid localization sequences were predicted for all these proteins, using different protein subcellular localization prediction methods, such as ChloroP (Emanuelsson et al., 1999), PSORT (Nakai and Horton, 1999) or TargetP (Emanuelsson et al., 2000). C-terminal GFP fusions of all PpADTs were transiently expressed in N. benthamiana leaves via agroinfiltration, in order to analyze its subcellular localization. PpADT-A, PpADT-B, PpADT-C,

PpADT-G and PpADT-H were successfully expressed and their plastid localization confirmed (**Figure 2**). Unfortunately, no expression was detected for PpADT-D, PpADT-E, PpADT-F and PpADT-I, and consequently we were unable to confirm the subcellular localization of their expression products. These data are consistent with the previously described plastid localization of the six *A. thaliana* proteins (Rippert et al., 2009).

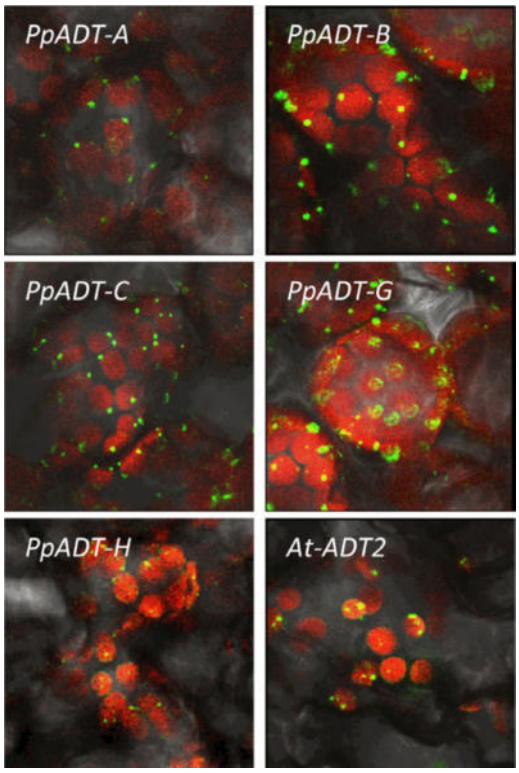
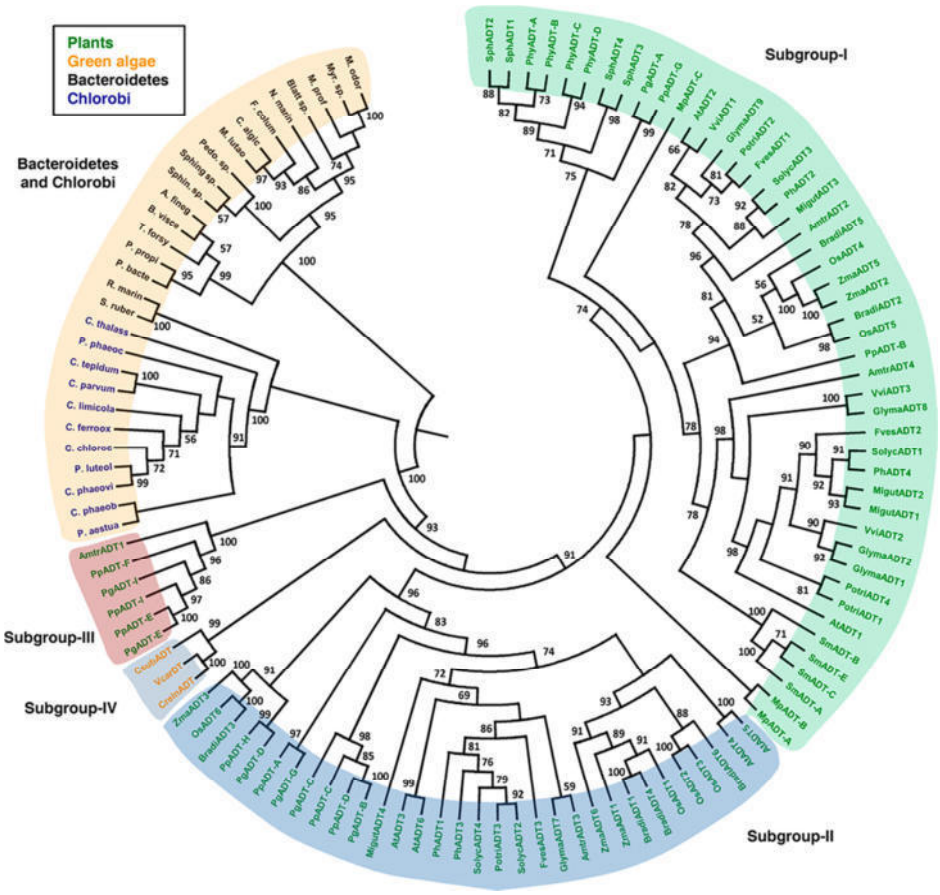


Figure 2. Subcellular localization of the *Pinus pinaster* arogenate dehydratase (ADT) isoforms in *Nicotiana benthamiana* leaves. Transient expression in *N. benthamiana* leaves was achieved by agroinfiltration using pGWB5 plasmids harbouring C-terminal GFP fusions to PpADTs. GFP fluorescence was monitored 48 h after agroinfiltration using confocal laser scanning microscopy. Subcelullar localization of AtADT2::GFP was employed as a positive control for chloroplast targeting.

Phylogenetic relationship among ADTs from different plant clades

We have conducted a deep phylogenetic analysis of ADT protein sequences from different and distant plant clades, including lycophytes (*Selaginella moellendorffii*), mosses (*Physcomitrella patens* and *Sphagnum fallax*), liverworts (*Marchantia polymorpha*), gymnosperms (*Pinus pinaster* and *Picea glauca*), basal angiosperms (*Amborella trichopoda*) and angiosperms (monocots *Brachypodium distachyon*, *Oryza sativa* and *Zea mays*, and dicots *Arabidopsis thaliana*, *Fragaria vesca*, *Glycine max*,



◄ **Figure 3. Maximum-likelihood phylogeny of plant arogenate dehydratase (ADT) proteins**. Phylogeny of the 89 plant ADT protein sequences inferred using maximum likelihood. Sequences were aligned using MUSCLE (Edgar, 2004) and the tree was produced using PHYML. Bootstrapping was performed with 1000 replicates. Abbreviations: Amtr, *Amborella trichopoda*; At, *Arabidopsis thaliana*; Bradi, *Brachipodium distachyon*; Crein, *Chlamydomonas reinhardtii*; Csub, *Coccomyxa subellipsoidea*; Fves, *Fragaria vesca*; Glyma, *Glycine max*; Migut, *Mimulus guttatus*; Mpol, *Marchantia polymorpha*; Os, *Oryza sativa*; Pg, *Picea glauca*; Ph, *Petunia hybrida*; Phy, *Physcomitrella patens*; Potri, *Populus trichocarpa*; Pp, *Pinus pinaster*; Sm, *Selaginella moelindorffii*; Solyc, *Solanum lycopersicum*; Sph, *Sphagnum fallax*; Vcart, *Volvox carteri* f. *nagariensis*; Vvi, *Vitis vinífera*; and ZMa, *Zea mays*. Sequences from *Bacteroidetes/Chlorobi* were included as the root of the plant ADT/prephenate dehydratase (PDT) phylogenetic analysis: A. fineg, *Alistipes finegoldii*; B. visce, *Barnesiella viscericola*; Blatt sp., *Blattabacterium* sp.; C. algic, *Cellulophaga algicola*; C. chloroc, *Chlorobium chlorochromatii*; C. ferroox, *Chlorobium ferrooxidans*DSM 13031; C. limícola, *Chlorobium limícola*; C. parvum, *Chlorobaculum parvum*; C. phaeobI, *Chlorobium phaeobacteroides* I; C. phaeobII, *Chlorobium phaeobacteroides*; C. phaeovi, *Chlorobium phaeovibrioides*; C. thalass, *Chloroherpeton thalassium*; C. tepidum, *Chlorobium tepidum*TLS; F. colum, *Flavobacterium columnare*; M. lutao, *Muricauda lutaonensis*; M. odor, *Myroides odoratimimus*; M. prof, *Myroides profundi*; Myr sp., *Myroides* sp.; N. marin, *Nonlabens marinus*; P. aestua, *Prosthecochloris aestuarii*; P. bacte, *Porphyromonadaceae bacterium*; P. luteol, *Pelodictyon luteolum*; P. phaeoc, *Pelodictyon phaeoclathratiforme*; P. propi, *Paludibacter propionicigenes*; Pedo sp., *Pedobacter* sp.; R. marin, *Rhodothermus marinus*; S. rubber, *Salinibacter ruber*; Sphin sp., *Sphingobacterium* sp.; Sphing sp. *Sphingobacterium* sp.; T. forsy, *Tannerella forsythia*.

Mimulus guttatus, *Petunia hybrida*, *Populus trichocarpa*, *Solanum lycopersicum* and *Vitis vinifera*). ADT sequences from chlorophytes (*Chlamydomonas reinhardtii*, *Coccomyxa subellipsoidea* and *Volvox carteri* f. *nagariensis*) were also included in the study. Dornfeld et al. (2014) reported that the *Bacteroidetes/Chlorobi* phyla have the closest bacterial homologs to plant ADTs, and therefore a set of protein sequences corresponding to species within this phyla were included to root the phylogenetic analysis. Our results indicate that ADTs are phylogenetically

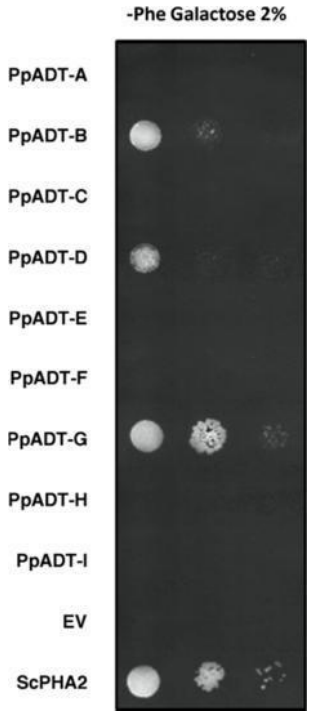
divided into four subgroups I–IV: I, II and III, exclusive to plants; and IV, exclusive to green algae (**Figure 3**). Subgroup I is the only subgroup that includes sequences from all plant species considered, and interestingly includes all ADT protein sequences from the most ancient plant clades: lycophytes, mosses and liverworts. Subgroup II includes only sequences from gymnosperms and angiosperms, and finally subgroup III comprises only sequences from gymnosperms and a sequence from the basal angiosperm *Amborella trichopoda* (**Figure 3**). According to our analysis, gymnosperms are the only plant clade that retains representatives in all subgroups. Interestingly, the sequences from chlorophytes, the ancestors of green land plants, form an independent group preceding the genetic diversification observed in plants.

Functional identification of PDT active enzymes in *Pinus pinaster*

To identify functional enzymes in pine, we developed a genetic approach based on the complementation of the yeast *pha2* mutant. The *PHA2* gene encodes a single, monofunctional PDT enzyme involved in the unique Phe biosynthesis pathway in yeast, and its deletion results in a strict requirement of exogenous Phe for growth (Bross et al., 2011). Coding sequences from *PpADT-A* to *-I* were subcloned in the yeast galactose-inducible expression vector pYES260, and *ScPHA2* was subcloned in the same vector as the positive control.

Mutant *pha2* yeast strains harboring full-length sequences from *PpADT* genes and the appropriate controls were grown in YNB media plates lacking Phe and uracil. Heterologous expression under the control of the *GAL1* promoter was induced by adding galactose and detected by western blotting (Figure S2). Only yeast strains expressing PpADT-B or -G enzymes

were able to restore the *pha2* phenotype after 2 days of growth in the absence of Phe, indicating that these enzymes exhibit PDT activity *in vivo* (**Figures 4** and S3). Partially delayed growth was also detected after 4 days in strains expressing PpADT-D, suggesting that this enzyme can rescue the mutant with low efficiency. Interestingly, PpADT-B and PpADT-G are phylogenetically located in subgroup I, together with other enzymes described as having PDT activity, such as *Arabidopsis thaliana* AtADT1 and AtADT2 (Cho et al., 2007) or *Petunia hybrida* PhADT2 (Maeda et al., 2010). By contrast, other enzymes for which PDT activity has also been described, PpADT-D (this work), AtADT6 (Cho et al., 2007) and PhADT3 (Maeda et al., 2010), are located in subgroup II.



◄ **Figure 4. Identification of functional prephenate dehydratase (PDT) enzymes from *Pinus pinaster* in yeast**. *PpADT* coding sequences were cloned into the pYES260 *Saccharomyces cerevisiae* expression vector. For yeast complementation assays, 10 ml of serial decimal dilutions of the *pha2* mutant yeast strains expressing each of the PpADTs were spotted onto YNB plates without Phe. Galactose was added to the medium to induce recombinant expression. Plates were incubated at 28°C and photographed after 3 days. Yeast growth assays were repeated three times with similar results.

Table 1. Kinetic parameters for the prephenate dehydratase activity of the *Pinus pinaster* recombinant proteins.

	K_m (mM)	V_{max} (pkat µg^{-1})	K_{cat} (s^{-1})	K_{cat}/K_m (Ms)$^{-1}$
PpADT-A	nd	ND	nd	nd
PpADT-B	0.34	9.06	0.3539	1040.8
PpADT-C	2.72	1.04	0.04	14.7
PpADT-D	nd	11.14	nd	nd
PpADT-G	0.15	18.58	0.7158	4772.1
Cmut1	0.30	0.78	0.031	103.3
Gmut 9	2.15	1.05	0.04	18.6
Gmut11	2.11	0.95	0.036	17.14

ND, not detected; nd, not determined.

With the aim of carrying out a deeper biochemical characterization of these enzymes, we proceeded to its recombinant production in *Escherichia coli*. PpADT- E, PpADT-F and PpADT-I, the three members within subgroup III, showed low expression levels although various plasmid vectors and *E. coli* strains were used. PpADT-H purification could not be carried out because of the degradation cause by *E. coli* host cells. In contrast, PpADT-A, PpADT-B, PpADT-C, PpADT-D and PpADT-G were successfully expressed and purified. PDT kinetic analyses were conducted using prephenate as substrate. Consistently with our previous results in yeast no PDT activity could be detected for PpADT-A; however, PpADT-B, PpADT-C, PpADT-D and PpADT-G enzymes displayed detectable PDT activity. PpADT-G showed the highest activity per microgram of enzyme, 18.58 pkat µg^{-1}, whereas PpADT-C showed the lowest detectable value, 1.04 pkat µg^{-1} (**Table 1**). V_{max} (maximum velocity) values of PpADT-B and PpADT-D were

9.06 and 11.14 pkat μg^{-1}, respectively (**Table 1**). Subsequent kinetic characterization demonstrated that these enzymes strongly differ in their apparent K_m for prephenate, showing PpADT-B and PpADT-G the lowest values, 0.34 and 0.15 mM (**Table 1**). These results are in line with our previous PDT activity screening in yeast, suggesting that only enzymes with low apparent Michaelis constant (K_m) values for prephenate, PpADT-B and PpADT-G are capable of restoring growth in the pha2 mutant.

Mutagenic identification and functional characterization of the PDT-activity conferring (PAC) domain

To identify the residues responsible for PDT activity, we performed an amino acid sequence alignment using the enzymes from Arabidopsis thaliana and *Pinus pinaster* (Figure S4), because in vivo PDT activity has been analyzed only in these two species. Based on this alignment we proceeded to identify sequences that were common to enzymes displaying efficient PDT activity, PpADT-B, PpADT-G, AtADT1 and AtADT2, but absent in the rest of sequences. Five conserved motifs were selected and single mutants were generated for the PDT-active enzyme PpADT-G, by replacing selected residues with those present in the equivalent positions in PpADT-C, an enzyme without PDT activity (**Figure 5A**). In Gmut^1, Leu^{151} and Val^{152} from PpADT-G were substituted with the equivalent residues Ile^{178} and Ala^{179}, according to the sequence of PpADT-C. Similarly, Gmut^2 and Gmut^3 involved the substitution of Glu^{203} and Gln^{239} with the equivalent residues Tyr^{230} and Glu^{268}, respectively. Gmut^4 involved the substitution of a region containing 22 amino acids from Thr^{303} to Asn^{324} with the equivalent region from PpADT-C, which ranged from Ala^{332} to Ser^{353}. Finally, in Gmut^5, three residues within

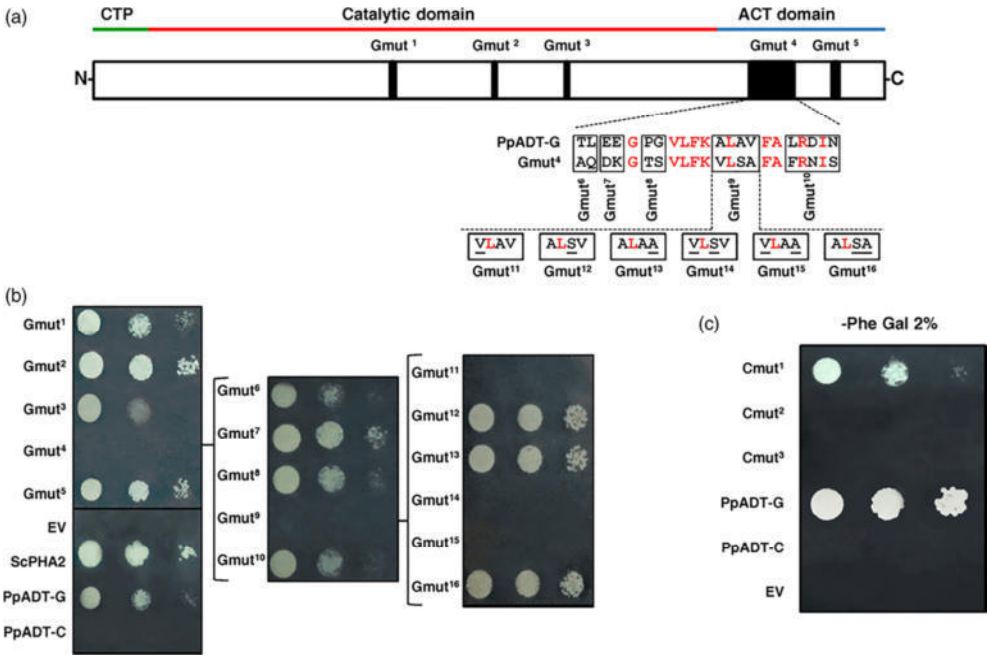


Figure 5. A small domain within the ACT-regulatory domain determines prephenate dehydratase (PDT) activity housed by arogenate dehydratase (ADT) enzymes. **(A)** Procedure followed for the identification of residues involved in PDT activity. Filled black boxes correspond to site-directed mutagenesis of PpADT-G: Gmut^{1}–Gmut^{5}. Lower part, Gmut^{6}–Gmut^{10} correspond to the mutagenic dissection of the protein region encompassed by Gmut^{4}. Gmut^{11}–Gmut^{16} correspond to the molecular dissection of the region encompassed by Gmut^{9}. **(B)** Serial decimal dilutions of *pha2* mutant yeast strains carrying each of the mutants and controls were grown without Phe. Yeast was grown on YNB plates containing 2% galactose to induce heterologous expression. EV,*pYES260*; ScPHA2, *pYES260-ScPHA2*. PpADT-C and PpADT-G were employed as negative and positive growth controls. **(C)** The PDT-activity conferring (PAC) domain enhances PDT activity of ADT enzymes. *pha2* yeast strains harboring PpADT-C derivative mutants Cmut^{1}, Cmut^{2} and Cmut^{3} (described in the text) were spotted onto YNB plates without Phe. Galactose was added to the medium to induce recombinant expression. PpADT-G and PpADT-C were used as positive and negative growth controls, respectively. In all experiments plates were incubated at 28°C and photographed after 3 days. Yeast growth assays were repeated three times with similar results.

the C-terminal region of PpADT-G (Gly^{374}, His^{375} and Leu^{376}) were replaced with the equivalent residues from PpADT-C, Gln^{404}, Glu^{405} and Val^{406} (**Figures 5A** and S4). $Gmut^1$, $Gmut^2$ and $Gmut^3$ contained residues within the catalytic domain, whereas $Gmut^4$ and $Gmut^5$ included residues located in the C-terminal regulatory ACT domain (**Figure 5A**). ACT domains are involved in end product- mediated allosteric regulation (Grant, 2006), and are linked to a wide range of metabolic enzymes, including PDTs and ADTs from multiple organisms.

Mutant *pha2* yeast strains harboring each of the PpADTG mutant proteins, $Gmut^1$–$Gmut^5$, were grown in media lacking Phe in order to determine whether these mutations could affect the PDT activity housed by wild-type PpADTG. $Gmut^1$, $Gmut^2$, $Gmut^3$ and $Gmut^5$ mutagenesis did not affect yeast growth. In contrast, $Gmut^4$ mutagenesis completely abolished the rescue of the mutant, indicating that the region encompassed by $Gmut^4$ is critical for PDT activity (**Figure 5B**). To further confirm the important role of this region for PDT activity, we generated a mutant protein, $Cmut^1$, in which the Ala^{332}–Ser^{353} domain in PpADT-C was substituted with the 22 residues domain (Thr^{303}–Asn^{324}) from PpADT-G. As shown in Figure 5C, $Cmut^1$ complemented the *pha2* mutation, demonstrating that the single exchange of this domain conferred PDT activity (hereafter referred to as the PDT activity-conferring domain, PAC domain). The PAC domains are present in all ADT enzymes belonging to subgroup I (**Figure 3**), and encompass a region of 22 amino acids, which includes 11 residues that are highly conserved in most plant ADT enzymes, and 11 residues that are highly conserved in PDTs with low apparent K_m values, and are consequently putatively involved in this activity. To indisputably identify the residues involved in PDT activity, a molecular dissection of the PAC domain was performed through site-directed mutagenesis. Five different mutations were

generated in which selected residues in PpADT-G were substituted with equivalent positions in PpADT-C: PpADT-G$^{T303A+L304Q}$ (Gmut^{6}), PpADT-G$^{E305D+E306K}$ (Gmut^{7}), PpADT-G$^{P308T+G309S}$ (Gmut^{8}), PpADT-G$^{A314V+A316S+V317A}$ (Gmut9) and PpADT-G$^{L320F+D322N+N324S}$ (Gmut^{10}) (**Figure 4A**). As shown **in Figure 5B**, yeast growth was completely abolished in the Gmut^{9} enzyme but remained unchanged in Gmut^{6}, Gmut^{7}, Gmut^{8} and Gmut^{10} enzymes, indicating that the single and simultaneous mutagenesis of A^{314}, A^{316} and V^{317} of PpADT-G was sufficient to impact PDT activity. Next, we conducted

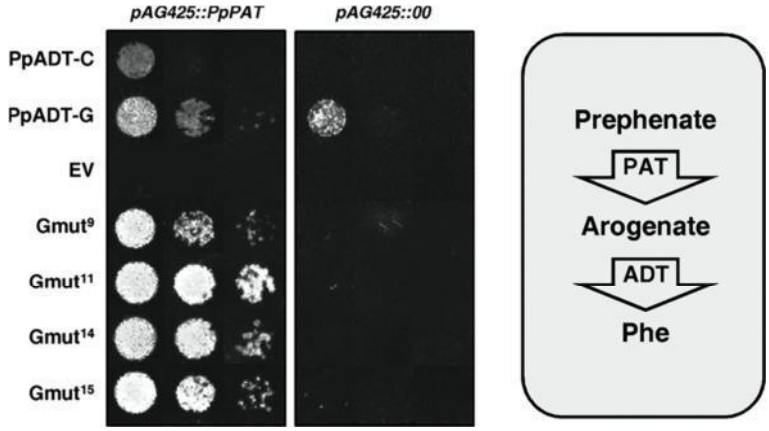


Figure 6. Suppression of prephenate dehydratase (PDT) activity does not affect arogenate dehydratase (ADT) activity in yeast. *pha2* yeast strains expressing each of the wild-type (wt) or mutant ADT proteins from pine were co-transformed with a pAG425 yeast expression vector carrying the open reading frame (ORF) from PpPAT (*pAG425::PpPAT*). In parallel, the same strains were co-transformed with the pAG425 empty vector (*pAG425::00*). Yeast was grown in YNB plates containing 2% galactose to induce recombinant expression. Plates were incubated at 28°C and photographed after 3 days. Yeast growth assays were repeated three times with similar results. PpADT-G and PpADT-C were used as positive and negative growth controls, respectively. EV, pYES260 empty vector. Scheme on the right, Phe biosynthetic metabolic bypass in yeast *pha2* via heterologous expression of PpPAT and PpADT.

a more detailed analysis aimed at identifying whether each of these three residues are critical for PDT activity. For this purpose, mutagenesis of each of the three amino acids was performed: PpADT-G^{A314V} ($Gmut^{11}$), PpADT-G^{A316S} ($Gmut^{12}$) and PpADT-G^{V317A} ($Gmut^{13}$). We also considered the combined mutagenesis of these residues: PpADTG $^{A314V+A316S}$ ($Gmut^{14}$), PpADT-$G^{A314V+V317A}$ ($Gmut^{15}$) and PpADT-$G^{A316S+V317A}$ ($Gmut^{16}$) (**Figure 5A**). Our results showed that mutant *pha2* yeast strains harboring $Gmut^{12}$, $Gmut^{13}$ or $Gmut^{16}$ were completely able to grow in media lacking Phe, suggesting that changes introduced to the mutant enzymes did not affect PDT activity (**Figure 5B**). By contrast, $Gmut^{11}$, $Gmut^{14}$ and $Gmut^{15}$ resulted in enzymes unable to rescue the *pha2* mutant, demonstrating that A^{314} ($Gmut^{11}$) displays an important role for PDT activity (**Figure 5B**). Consistently, the combined mutation of A^{314} with A^{316} ($Gmut^{14}$) or V^{317} ($Gmut^{15}$) also suppressed mutant growth (**Figure 5B**). As we demonstrated for the $Cmut^{1}$ mutant, the single introduction of a PAC domain results in an enhancement of the PDT activity that allows the growth of the *pha2* mutant. We next examined whether the same effect could be achieved through a more specific modification within this domain. For this purpose the PpADT-C enzyme was mutagenized to introduce, at equivalent positions, A^{314} ($Cmut^{2}$) or combined A^{314} + A^{316} + V^{317} residues ($Cmut^{3}$). **Figure 5C** shows that none of these proteins were able to restore the mutant growth, suggesting that both a single amino acid but also a suitable molecular environment are required to promote PDT activity.

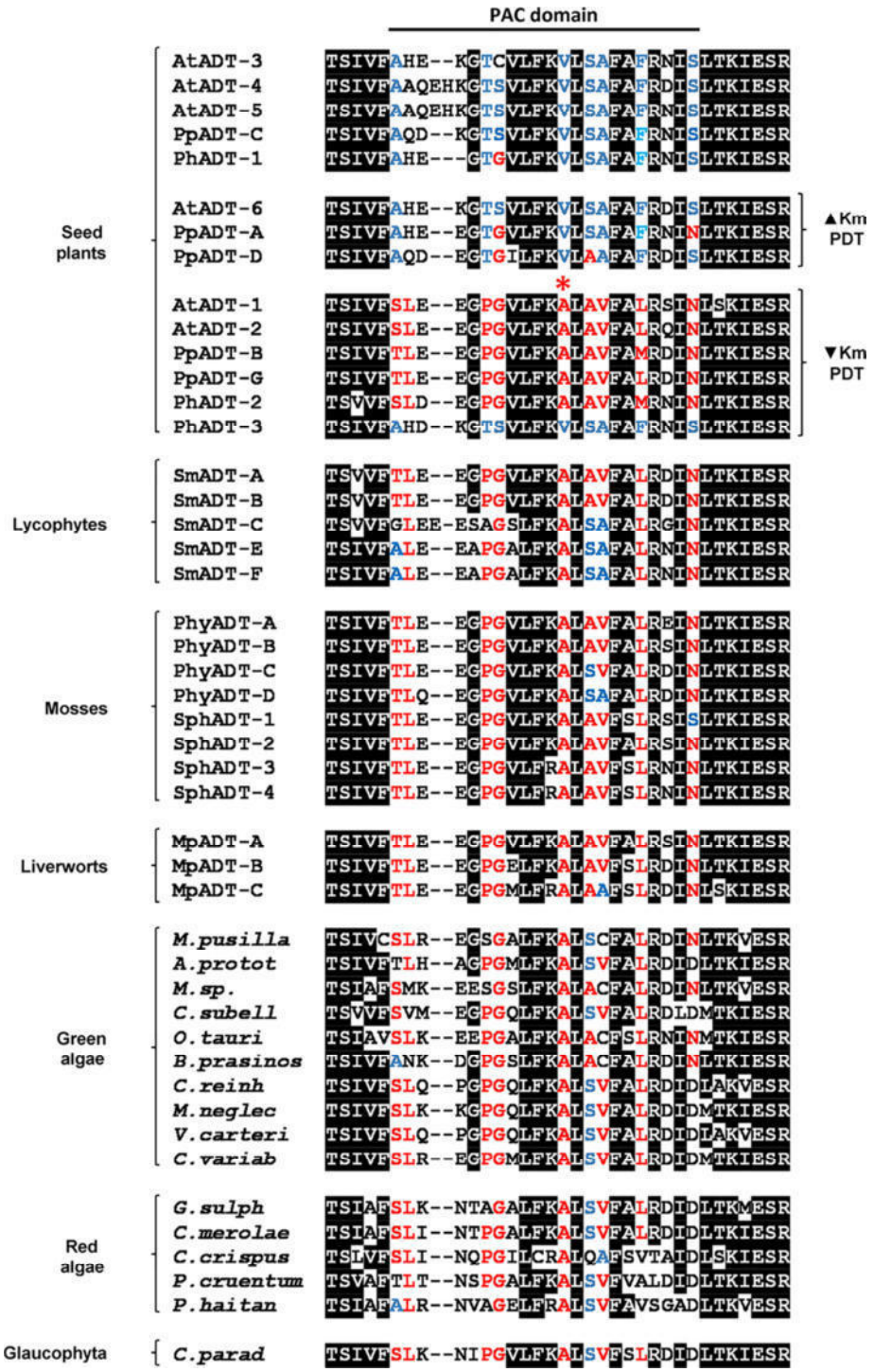
We developed a method to determine whether these residues are exclusively critical for PDT activity or whether they are also important for ADT activity. Mutant yeast strains expressing each of the PDT activity-defective mutants were co-transformed with a pAG425GAL vector containing the open reading frame (ORF) of the pine prephenate

aminotransferase minus the region encoding for the chloroplast transit peptide (de la Torre et al., 2006). The simultaneous expression of PpPAT with an active ADT enzyme resulted in the reconstitution of a functional arogenate pathway in yeast that was able to synthesize Phe. As shown in **Figure 6**, Gmut^{9}, Gmut^{11}, Gmut^{14} and Gmut^{15} mutant enzymes are defective in their PDT activity but retained their ADT activity.

To determine how mutations in the PAC domain affect PDT activity we proceeded to the recombinant production and purification of Gmut^{9} and Gmut^{11}, and subsequent kinetic characterization of the mutant enzymes. Following identical procedures, we also obtained recombinant Cmut1 protein. Apparent K_m values for Gmut^{9} and Gmut^{11} were 2.15 and 2.11 mM, respectively, more than an order of magnitude higher than the value calculated for PpADT-G (**Table 1**). Moreover, catalytic efficiencies (K_{cat}/K_m) of Gmut^{9} and Gmut^{11} were 256- and 278-fold lower than PpADT-G, suggesting that the PAC domain is critical for an efficient catalysis of the PDT reaction by reducing the apparent K_m values of ADT enzymes for prephenate (**Table 1**). The apparent K_m of Cmut^{1} towards prephenate is an order of magnitude lower than that of PpADT-C (**Table 1**), demonstrating that the single presence or absence of this domain, respectively, determines a low or high apparent affinity for prephenate. Interestingly, the low apparent K_m towards prephenate observed in Cmut^{1} compared with PpADT-C is accompanied by a seven-fold increase in the catalytic efficiency (K_{cat}/K_m) of the enzyme (**Table 1**).

The PAC domain is present along the entire *Archaeplastida* clade

The presence of ADT/PDT enzymes containing a PAC domain in other organisms outside of land plants was analyzed based on evolutionary criteria. Using multiple BLAST searches we identified up to 10 ADT sequences corresponding to different species of green algae, which together with land plants make up the *Viridiplantae* group. Multiple sequence alignment showed that most of the PAC domain characteristic residues present in plants are highly conserved in all of these sequences (**Figure 7**). Remarkably, the essential Ala^{314} is preserved in all of them. We next examined whether this domain is also preserved in the remaining groups belonging to the *Archaeplastida* clade: red algae and glaucophytes. Blast searches allowed the identification of five sequences corresponding to ADTs from red algae. With respect to the glaucophyta clade only an ADT sequence corresponding to *Cyanophora paradoxa* could be detected. This result is probably related to the fact that this is a poorly studied group with a greatly reduced number of species. Sequence alignment showed that throughout evolution both red algae and glaucophyta has also conserved most of the PAC domain characteristic residues, including Ala^{314} (**Figure 7**). Through BLAST searches of the publicly available prokaryotic protein databases (NCBI, Ensembl and UniProt) we were able to observe that, surprisingly, plant ADTs appear to be more closely related to enzymes from the *Bacteroidetes/Chlorobi* phyla than those of cyanobacteria. To further investigate the evolutionary origin of plant ADTs, we analyzed the level of conservation of the characteristic residues within the PAC region through multiple sequence alignment, including 21 sequences reflecting the biodiversity within the cyanobacteria and 30 sequences corresponding to the *Bacteroidetes/Chlorobi* phyla (Figure S5). These comparisons showed that the key residues within the PAC domain, including Ala^{314}, are more



◄ **Figure 7. Prephenate dehydratase (PDT) activity conferring (PAC) domain-containing arogenate dehydratase (ADT) enzymes are present in all Archaeplastida clades**. Multiple sequence alignment of the ADT/PDT PAC domain. Abbreviations in addition to those indicated in Figure 2: A. protot, *Auxenochlorella protothecoides*; B. prasinos, *Bathycoccus prasinos*; C. crispus, *Chondrus crispus*; C. merolae, *Cyanidioschyzon merolae*; C. parad, *Cyanophora paradoxa*; C. reinh, *Chlamydomonas reinhardtii*; C. subell, *Coccomyxa subellipsoidea*; C. variab, *Chlorella variabilis*; G. sulph, *Galdieria sulphuraria*; M. sp., *Micromonas* sp.; M. neglec, *Monoraphidium neglectum*; M. pusilla, *Micromonas pusilla*; O. tauri, *Ostreococcus tauri*; P. cruentum, *Porphyridium cruentum*; P. haitan, *Pyropia haitanensis*; and V. carteri, *Volvox carteri f. nagariensis*. Residues characteristics to active and non-active PDT enzymes are underlined in red and blue, respectively.

highly preserved in *Bacteroidetes/Chlorobi* sequences than in cyanobacteria. Remarkably, Ala^{314} is absent in cyanobacteria and Bacteroidetes, but is present in all sequences corresponding to *Chlorobi*, with the single exception of a sequence corresponding to *Chlorobium phaeobacteroides*; however, this organism has an additional sequence in which Ala^{314} is present.

DISCUSSION

Aromatic amino acids Phe, Tyr and Trp are essential compounds for all living organisms since a wide range of important products are synthesized from them. Plants and most microorganisms have developed specific routes for their biosynthesis, whereas animals depend on their acquisition through diet (Mathews et al., 2012; de la Torre et al., 2014b). In this study, we report the identification of the complete ADT family in *Pinus pinaster* and the characterization of enzymes displaying PDT activity, and thus

putatively involved in a prokaryotic-type pathway of Phe biosynthesis. We propose that the evolutionary history of a small protein domain conferring PDT activity (PAC domain) may have contributed to preserving a functional phenylpyruvate-dependent pathway for Phe biosynthesis in all plant lineages.

To cope with the main objective of this study we opted for the use of *Pinus pinaster* as a model species, considering that Phe is the precursor of a highly complex secondary metabolism (Franceschi et al., 2005; Bonello et al., 2006; Warren et al., 2015). We have found that the maritime pine ADT gene family is composed of at least by nine members: *PpADT-A* to *-I* (Appendix S1). The genomes of other conifers such as *Picea glauca* and *Pinus taeda* also contains a similar number of ADT genes (eight or nine), whereas a lower number of genes is found in the genomes of angiosperms such as *Arabidopsis thaliana* (six), *Nicotiana benthamiana* (four) and *Petunia hybrida* (3). In addition, our results indicate that at least five of these genes (*PpADT-A*, *PpADT-B*, *PpADT-C*, *PpADT-G* and *PpADT-H*) encode proteins of plastidial localization (**Figure 2**), and may contribute to Phe biosynthesis in the plastid. The *ADT/PDT* gene family from gymnosperms is phylogenetically diverse, with its members distributed into three different subgroups I–III (**Figure 3**). ADT/PDT sequences from most ancient plant clades, including lycophytes, mosses or liverworts, are exclusively found in subgroup I, whereas angiosperms are included in subgroups I and II (**Figure 3**). Interestingly, we have observed that sequences corresponding to green algae, which together with land plants comprise the sub-kingdom *Viridiplantae*, are exclusively located in subgroup IV, pre-dating the divergence observed in plants (**Figure 3**).

We hypothesize that a more complex secondary metabolism, a feature of conifer trees (Franceschi et al., 2005; Bonello et al., 2006; Warren et al., 2015), is corelated to a larger and more diverse ADT/PDT family. These findings suggest that Phe has many different metabolic fates in these woody plants. Consistently with this possibility, the phenylalanine ammonia-lyase gene family in gymnosperms, encoding for the first and committed step in the phenylpropanoid pathway, is composed of a greater and more diverse set of genes than in angiosperms (Bagal et al., 2012). Moreover, a recent analysis has shown that compared with angiosperms, conifer genomes include more genes encoding for key enzymes in the biosynthesis of Phe-derived compounds (Warren et al., 2015).

Using complementation analysis in yeast we determined that two among the pine genes, *PpADT-B* and *PpADT-G*, encode functional PDT enzymes (**Figure 4**). Minor complementation of PDT activity was detected for PpADT-D. The biochemical characterization of recombinantly generated pine enzymes showed that PpADT-B, PpADT-G, PpADT-D but also PpADT-A are capable of catalyzing the PDT reaction, showing different catalytic efficiencies and affinities for prephenate (**Table 1**). PpADT-B and PpADT-G enzymes, which resulted in a more effective complementation of the yeast mutant, showed the lowest apparent K_m values and the highest catalytic efficiencies. Similarly, only the AtADT1 and AtADT2 from *Arabidopsis* were able to complement the same yeast mutants (Bross et al., 2011). Furthermore, only these two ADT enzymes in the Arabidopsis family have low apparent K_m values for prephenate (Cho et al., 2007). Together, these results indicate that the functional rescue of yeast mutants defective in PDT activity is achieved by ADT enzymes with a high apparent affinity for prephenate. Additionally, Maeda et al. (2010) demonstrated that *Petunia*

PhADT2 and PhADT3, but not PhADT1, display PDT activity showing a low apparent K_m for prephenate.

Phylogenetic analysis clearly showed that all characterized ADTs exhibiting low apparent K_m values for prephenate are included in subgroup I, with the exception of PhADT3, which is located into subgroup II (**Figure 3**). This phylogenetic grouping, together with the fact that throughout evolution all plant clades have conserved at least one member within subgroup I, suggest that all plant ADTs emerged from a common subgroup I-type ancestor displaying ADT and PDT activities, and thus with the ability to participate in a Phe-biosynthetic pathway using PPY as an intermediary. Although we have reported the existence in pine of enzymes displaying PDT activity at high catalytic efficiencies (**Table 1**), future analyses will be necessary to study the in planta contribution of the phenylpyruvate-mediated Phe biosynthetic pathway in different tissues and physiological scenarios.

The above findings prompted us to search characteristic residues and/or domains exclusive to enzymes belonging to subgroup I. Through multiple rounds of mutation, functional analysis in yeast and biochemical assays, we identified a region of 22 amino acids within the C-terminal ACT regulatory domain that is responsible for conferring low apparent K_m for prephenate, designated as the PAC domain (**Figure 4**). The importance of this domain as an enhancer of PDT activity was confirmed by its ability to reduce the apparent K_m towards the substrate when introduced in an enzyme exhibiting low apparent affinity such as PpADT-C. Moreover, we identified a single residue within the PAC domain, Ala^{314}, which is essential for conferring a low apparent K_m for prephenate, but also for enhancing V_{max}. We have found that PDT activity is determined fundamentally on the

regulatory domain, and not through specificities in the catalytic domain. Very interestingly, the Ala^{314} residue is precisely localized between two tetrapeptides GVLF and ESRP that are highly conserved in all ADTs and PDTs, and are involved in the allosteric binding of Phe, as demonstrated in *Escherichia coli* and *Chlorobium tepidum* (Pohnert et al., 1999; Tan et al., 2008; Figure S6). This observation suggests a correlation between the Phe-mediated allosteric regulation and PDT activity that requires a deeper structural analysis to address the exact mechanism determining PDT activity. Interestingly, the PAC domain is exclusive and characteristic for ADTs belonging to subgroups I and IV. This domain is present in all examined ADTs from green algae, from ancient plant clades such as mosses, liverworts or lycophytes, but is also present in a subset of ADTs from angiosperms and gymnosperms. Based on these observations we propose that at the origin and during the early evolution of land plants, all ADTs retained PAC domains and thus the ability to display both ADT and PDT activities. Later, throughout evolution, gene duplication events gave rise to the emergence in plants of subgroup II and III ADTs. This appearance of different types of plant ADTs was likely accompanied by increased specialization in the use of arogenate rather than prephenate; however, all plant species analyzed have retained the subgroup-I bifunctional ADT/ PDT enzymes, suggesting an essential role for PDT throughout the life cycle of plants.

Considering the above findings we extended our sequence analysis to green and red algae and glaucophytes, the three evolutionary lineages that emerged after the primary endosymbiotic event that led to the emergence of plastids (Rodríguez-Ezpeleta et al., 2005; Keeling, 2013). Through these analyses we detected a high level of conservation of the PAC domain characteristic residues and, remarkably, the Ala^{314} residue is conserved in

all sequences. These observations suggest that PDT activity was present in the ancestor of all *Archeaplastidia*. In agreement with previous results from Dornfeld et al. (2014), we have observed that the closest bacterial homologs to plant ADTs are found in the *Bacteroidetes/Chlorobi* phyla, and not in the *Cyanobacteria phylum*. We have also noticed that the characteristic residues in the PAC domain are better preserved in sequences from *Bacteroidetes/Chlorobi* than in Cyanobacteria. Moreover, the critical residue of Ala^{314} is preserved in sequences from Chlorobiales but is absent in Bacteroidetes or Cyanobacteria (Figure S5). Interestingly, among the 12 sequences from Chlorobiales identified in our study, only a single sequence corresponding to *Chlorobium phaeobacteroides* lacks the residue Ala^{314}, however, this organism possesses a second sequence containing this residue. These data suggest an important role for maintaining a residue that is critical for a high apparent affinity towards prephenate.

Whether plant ADT enzymes may have an endosymbiotic origin or were present in the eukaryotic ancestor remains unknown. In this regard, very recently, during the preparation of this manuscript, the complete sequencing of the genome of the Chloroflexi species *Ardenticatena maritima* has been released (Kawaichi et al., 2015), showing that it holds an ADT protein with greater identity (60%) to plant ADTs than that described for the closest *Bacteroidetes/Chlorobi* (46–49%) or cyanobacteria (36–39%) sequences. In contrast, other sequenced *Chloroflexi* species do not show these levels of identity. The rapid increase in the number of sequenced bacterial genomes will be critical to the future resolution of this issue.

In summary, our results indicate that over the course of evolution, various subfamilies of ADTs have evolved from the subfamily I-type ADTs existing in most ancient plant clades. The availability of a more diverse catalog of a critical enzyme for the biosynthesis of Phe has allowed a more specialized ability to deal with the multiple physiological situations that rely on the biosynthesis of Phe-derived compounds.

EXPERIMENTAL PROCEDURES

cDNA cloning

Full-length *PpADT-A* to *-I* cDNAs were isolated from total RNA extracted from Pinus pinaster seedlings by reverse transcription- PCR using iScript Reverse Transcription Supermix (Bio-Rad, http://www.bio-rad.com) with PpADT-AFW/PpADT-ARV, PpADT-BFW/ PpADT-BRV, PpADT-CFW/PpADT-CRV, PpADT-DFW/PpADT-DRV, PpADT-EFW/PpADT-ERV, PpADT-FFW/PpADT-FRV, PpADT-GFW/ PpADT-GRV, PpADT-HFW/PpADT-HRV, PpADT-I FW/PpADT-I RV, respectively (Table S1). Then, cDNAs were subcloned into the pJET1.2 (ThermoFisher ScientificTM, https://www.thermofisher.com) vector and completely sequenced. Later, cDNAs were PCR amplified using forward and reverse oligonucleotides containing partial attB1 and attB2 sites on their 50 ends, respectively: PpADTAattB1/ PpADT-AattB2, PpADT-BattB1/PpADT-BattB2, PpADT-CattB 1/PpADTCattB2, PpADT-DattB1/PpADT-DattB2, PpADT-EattB1/PpA DT-EattB2, PpADT-FattB1/PpADT-FattB2, PpADT-GattB1/PpADTGattB2, PpADT-HattB1/PpADT-HattB2, PpADT-IattB1/PpADT-IattB2 (Table S1). Complete attB1/attB2 sites were introduced through PCR amplification with oligonucleotides FattB1 and

FattB2 (Table S1), and products were cloned into the Gateway-compatible vector pDONR207 using BP Clonase II enzyme mix (Thermo-Fisher Scientific).

Expression analysis

Total RNA isolation and cDNA synthesis was obtained as described previously (de la Torre et al., 2014a). qPCR was performed on a CFX-384 Real Time System (Bio-Rad) with SsoFast EvaGreen Supermix (Bio-Rad) under the following conditions: 95°C for 2 min (one cycle), followed by 95°C for 1 s and 60°C for 5 s (45 cycles). cDNAs corresponding to 10 ng of reverse-transcribed RNA were used as a template for each reaction. Raw fluorescence data from each reaction were fitted to the Mass Action Kinetic 2 model, which requires no assumptions about the amplification efficiency of a qPCR assay (Boggy and Woolf, 2010). The initial target concentrations (D_0 parameter) for each gene were deduced from the Mass Action Kinetic 2 model using the qPCR package for the R environment (Ritz and Spiess, 2008), and normalized to *PpActin2* and *PpEF1*.

Subcellular localization of pine ADTs

Pine ADTs were subcloned into the pGWB5 vector via Gateway Technology (Invitrogen) to produce full-length PpADT proteins fused to GFP at their C-termini under the control of the CaMV *35S* promoter. The recombinant plasmids were introduced into the *Agrobacterium tumefaciens* C58C1 strains through electroporation. GFP-tagged ADT proteins were expressed through agroinfiltration (OD_{600} = 0.5) in *Nicotiana benthamiana* leaves, as previously described (He et al., 2004). All assays included the coexpression of the silencing suppressor p19 (Voinnet et al., 2003). GFP

fluorescence was monitored 48 h after agroinfiltration using a Leica TCS Sp5 II confocal microscope. For GFP detection, the excitation source was an argon ion laser at 488 nm, and emission was observed between 505 and 525 nm. Chloroplast autofluorescence was collected between 680 and 700 nm.

Maximum likelihood phylogeny analysis

ADT phylogenetic analysis was conducted with 89 sequences corresponding to multiple species from different plant clades and green algae: *Amborella trichopoda, Arabidopsis thaliana, Brachipodium distachyon, Chlamydomonas reinhardtii, Coccomyxa subellipsoidea Fragaria vesca, Glycine max, Marchantia polymorpha, Mimulus guttatus, Oryza sativa, Petunia hybrida, Physcomitrella patens, Picea glauca, Pinus pinaster, Populus trichocarpa, Selaginella moelindorffii, Solanum lycopersicum, Sphagnum fallax, Vitis vinifera, Volvox carteri* f. *nagariensis* and *Zea mays*. Yeast PHA2 protein was used to out-group root the tree. Sequences were aligned using MUSCLE (Edgar, 2004), and tree topology was inferred using maximum likelihood with PhyML (Guindon and Gascuel, 2003; Guindon et al., 2005). Bootstrapping was performed with 1000 replicates. The phylogenetic trees were drawn using MEGA 5.0 (Tamura et al., 2011).

Plasmid DNA constructs

PpADT-A, *-B*, *-C*, *-D*, *-E*, *-F*, *-H* and *-I* were PCR amplified from the corresponding pDONR207 derivative vectors using forward, Xh207FW, and reverse, Xh207RV, flanking oligonucleotides containing XhoI recognition sequences. XhoI-digested fragments were subcloned into the pYES260

modified vector, pYES260X, constructed by introducing a XhoI site downstream of the *GAL1* promoter through site-directed mutagenesis using pY2X1 and pY2X2 oligonucleotides (Table S1). Similarly, *PpADT-G* was PCR amplified from the corresponding pDONR207 derivative vector using forward and reverse flanking oligonucleotides containing BamHI sites: Bm207FW and Bm207RV. The BamHI-digested fragment was subcloned into the pYES260 modified vector, pYES260B, constructed by introducing a BamHI site downstream of the *GAL1* promoter through site-directed mutagenesis using pY2B1 and pY2B2 oligonucleotides (Table S1). *ScPHA2* coding DNA sequence (CDS) was amplified from yeast genomic DNA using oligonucleotides ScPHA2FW and ScPHA2RV (Table S1), and cloned into the pYES260B vector. *PpPAT* was amplified using oligonucleotides PpPATattB1 and PpPATattB2 (Table S1). The PCR product was later re-amplified with oligonucleotides FattB1 and FattB2, and cloned into the pDONR207 vector using BP Clonase II enzyme mix, and subsequently recombined into the pAG425 vector using the Gateway LR reaction.

Recombinant expression and purification of ADT enzymes in *Escherichia coli*

Open reading frames (ORFs) corresponding to pine ADTs minus the region encoding the chloroplast transit peptide were PCR-amplified using oligonucleotides containing the appropriate restriction sites, as described in Table S1. Thereafter, PCR products were digested and subcloned into the pET30b at NdeI and XhoI sites, with the single exception of *PpADT-G*, which was subcloned into NdeI and NotI sites. Plasmid constructs were sequenced to confirm that no undesirable alterations had occurred, and were subsequently transformed into *Escherichia coli* strain BL21-AI

(ThermoFisher Scientific). Cells were grown at 37°C with shaking in Luria Bertani broth containing 50 lg mL^{-1} kanamycin until OD_{600} = 0.6; 0.2% arabinose (w/v) and 1 mM IPTG were added, and cultures were incubated for 20–24 h at 15°C with shaking. Cells were pelleted by centrifugation at 4500 *g* and frozen. Proteins were purified by binding onto Ni-agarose resin (Protino NiNTA; Macherey-Nagel, http://www.mn-net.com). Protein aliquots were analyzed by SDS-PAGE to confirm purification.

PDT activity assay

Prephenate dehydratase activity was determined by the conversion of prephenate to phenylpyruvate according to the method described previously (Fischer and Jensen, 1987a). The 200 µL reaction mixture, containing Tris-HCl 50 mM, pH 8, prephenate 1 mM (P2384; Sigma-Aldrich) and 1 µg of the corresponding purified protein, was incubated at 37°C for 30 min. The reaction was stopped by adding 800 µL of NaOH 2.5 M. The absorbance was immediately measured at 321 nm. The assay was linear throughout the 30-min period. Apparent K_m values were determined from Lineweaver–Burk plots. Kinetic parameters were determined used the following range of prephenate concentrations: 10 µM, 50 µM, 100 µM, 200 µM, 500 µM, 1 mM, 2 mM, 5 mM and 10 mM.

Site-directed mutagenesis

Site-directed mutagenesis was performed to generate $Cmut^2$ and $Cmut^3$, in addition to $Gmut^1$–$Gmut^{16}$, but $Gmut^4$ was obtained through a PCR-based strategy using iProof high-fidelity DNA polymerase (Bio-Rad). Briefly, 20 ng of the corresponding pYES260B: PpADT-G or pYES260X-PpADT-C vectors were used for PCR amplification using the corresponding

oligonucleotides described in Table S1 under the following conditions: 2 min at 98°C, followed by 23 cycles (10 s at 98°C, 20 s at 55°C and 4 min at 72°C) and a final elongation of 4 min at 72°C. Samples were treated with DpnI (ThermoScientificTM) to eliminate native methylated plasmid DNA, and the PCR product was transformed into *Escherichia coli* DH5a with selection on 100 μg ml^{-1} ampicillin.

Chimeric $Gmut^4$ was constructed through the overlapping PCR method. First, using pDONR207-PpADT-G as the template, sequences encompassing CDS regions 1–902 and 979–1200 were amplified separately by PCR using oligonucleotides Bm207FW/Gmut4RV1 and Gmut4FW2/Bm207RV, respectively (Table S1). Second, using pDONR207-*PpADT-C* as the template, a sequence encompassing the CDS region 966–1089 was amplified by PCR using oligonucleotides Gmut4FW1/Gmut4RV2 (Table S1). All three fragments were subsequently fused through overlapping PCR with oligonucleotides Bm207FW/Bm207RV (Table S1). The resulting chimeric fragment was digested with BamHI restriction enzyme and subcloned into the pYES260B vector. Similarly, for the construction of chimeric Cmut1 two fragments encompassing CDS regions 1–989 and 1066–1320 of *PpADT-C* were PCR-amplified using oligonucleotides Xh207FW/Gmut4RV1 and Gmut4FW2/ Xh207RV, respectively (Table S1), with pDONR207-PpADT-C as the template. In parallel, a CDS fragment of *PpADT-G* was PCR-amplified using oligonucleotides Gmut4FW1/Gmut4RV2 (Table S1) and pDONR207-*PpADT-G* as the template. All three fragments were fused through overlapping PCR using oligonucleotides Xh207FW/Xh207RV (Table S1). The resulting chimeric fragment was digested with XhoI and subcloned into the pYES260X vector.

Yeast complementation assays

Complementation assays were performed using a haploid pha2 knock-out *Saccharomyces cerevisiae* strain (YNL316c: Mata, his3Δ1, leu2Δ0, lys2Δ0, ura3Δ0, YNL316c:kanMX4) obtained from the Saccharomyces Genome Database (http://www.yeastgenome.org). *Saccharomyces cerevisiae pha2* strain Phe autoxotrophy was confirmed by comparing growing rates in liquid glucose media with or without Phe (50 mg L^{-1}). Mutant yeast *pha2* cells were transformed with pYES260 plasmid carrying each PpADT gene, mutant derivatives or empty vector by the LiAc/SS-Carrier DNA/PEG method, as described by Gietz and Schiestl (2007). Depending on the experiment, cells were alternatively co-transformed with the pAG425 plasmid carrying *PpPAT* or with the empty vector. Positive transformants were identified by PCR colony testing and by selecting for uracil (pYES260) or leucine (pAG425) prototrophy in a solid synthetic minimal media, as described elsewhere (Ausubel., 1994). Positive colonies were grown in liquid media supplemented with Phe (50 mg L^{-1}) and raffinose (2%) to reach stationary phase. Aliquots of each culture were taken and pelleted at 11,000 *g* for 30 s, the media was removed and the cells were then washed with sterile distilled water (SDW) to eliminate remnant Phe. After that, cells were resuspended in SDW at four different concentrations of 10,000, 1000, 100 and 10 cells L^{-1}. A volume of 10 µL corresponding to each sample was spotted onto plates containing either 2% galactose (*GAL1* inductor) or 2% glucose (*GAL1* repressor), and with or without Phe (50 mg L^{-1}). Plates were incubated at 28°C for 3 days and photographed to monitor growth.

Protein determination and western blot

Protein concentrations were determined by the Bradford dye-binding method (Bradford, 1976). SDS-PAGE and western blots were carried out as described previously (Cánovas et al., 1984). The following antibodies were used: 1:5000 anti-His 6 (Roche Life Science, http://lifescience.roche.com) and 1:20,000 anti-mouse HRP-conjugate (SantaCruz Biotechnology, http://www.scbt.com).

Accession numbers

GenBank accession numbers: *PpADT-A* (KX010923), *PpADT-B* (KX010924), *PpADT-C* (KX010925), *PpADT-D* (KX010926), *PpADT-E* (KX010927), *PpADT-F* (KX010928), *PpADT-G* (KX010929), *PpADT-H* (KX010930) and *PpADT-I* (KX010931).

Acknowledgments

We are grateful to Dr. Gutiérrez-Beltran for the pAG425 vector and Kinia Ameztoy for assistance with yeast manipulations. We are indebted to Dr. Kohchi and Dr. Yamaoka for providing the *Marchantia polymorpha* ADT sequences. We are thankfull for grants provided by the Spanish MINECO (BIO2015-69285-R) and Junta Andalucía (BIO2012-0474). JEA is the recipient of a Formación Profesorado Universitario fellowship (Ministerio Educación, Spain).

Supplemental data

Additional Supporting Information may be found in the online version of this article:

Figure S1. Expression profiles of ADT genes in *Pinus pinaster* seedlings.

Figure S2. Heterologous expression of PpADT in yeast.

Figure S3. Control yeast growth conditions.

Figure S4. Protein sequence alignment of ADT from *Pinus pinaster* and *Arabidopsis thaliana*.

Figure S5. Sequence analysis of PAC domain-containing enzymes in certain bacterial groups.

Figure S6. Phe allosteric binding sites within the ACT domain of ADT/PDTs.

Table S1. Oligonucleotides used in this work.

Appendix S1. Coding DNA and protein sequences from the *Pinus pinaster* ADT family.

References

Ausubel FM (1994) *Current Protocols in Molecular Biology*, Vol 1, Ch 5. New York, (John Wiley & Sons) Inc.

Bagal UR, Leebens-Mack JH, Lorenz WW and Dean JF (2012) *The phenylalanine ammonia lyase (PAL) gene family shows a gymnospermspecific lineage*. BMC Genom, 11(13 Suppl 3):S1.

Benfey PN and Chua NH (1989) *Regulated genes in transgenic plants*. Science, 244:174-181.

Boggy GJ and Woolf PJ (2010) *A mechanistic model of PCR for accurate quantification of quantitative PCR data*. PLoS ONE, 5:e12355.

Bohlmann J, Lins T, Martin W and Eilert U (1996) *Anthranilate synthase from* Ruta graveolens. *Duplicated AS alpha genes encode tryptophansensitive and tryptophan-*

insensitive isoenzymes specific to amino acid and alkaloid biosynthesis. Plant Physiol, 111:507-514.

Bonello P, Gordon TR, Herms DA, Wood DL and Erbilgin N (2006) *Nature and ecological implications of pathogen-induced systemic resistance in conifers: a novel hypothesis*. Physiol Mol Plant Pathol, 68:95-104.

Bonner C and Jensen R (1987a) *Prephenate aminotransferase*. Methods Enzymol, 142:479-487.

Bonner C and Jensen R (1987b) *Arogenate dehydrogenase*. Methods Enzymol, 142:488-894.

Bradford MM (1976) *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 72:248-254.

Bross CD, Corea OR, Kaldis A, Menassa R, Bernards MA and Kohalmi SE (2011) *Complementation of the* pha2 *yeast mutant suggests functional differences for arogenate dehydratases from* Arabidopsis thaliana. Plant Physiol Biochem, 49:882-890.

Canales J, Bautista R, Label P et al. (2014) *De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology*. Plant Biotechnol J, 12:286-299.

Cánovas F, Valpuesta V and de Castro IN (1984) *Characterization of tomato leaf glutamine-synthetase*. Plant Science Letters, 37:79-85.

Cho MH, Corea OR, Yang H et al. (2007) *Phenylalanine biosynthesis in* Arabidopsis thaliana. *Identification and characterization of arogenate dehydratases*. J Biol Chem 282:30827-30835.

Cotton RG and Gibson F (1965) *The biosynthesis of phenylalanine and tyrosine: enzymes converting chorismic acid into prephenic acid and their relationships to prephenate dehydratase and prephenate dehydrogenase*. Biochim Biophys Acta, 100:76-88.

de la Torre F, De Santis L, Suárez MF, Crespillo R and Cánovas FM (2006) *Identification and functional analysis of a prokaryotic-type aspartate aminotransferase: implications for plant amino acid metabolism*. Plant J, 46:414-425.

de la Torre F, El-Azaz J, Ávila C and Cánovas FM (2014a) *Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism*. Plant Physiol, 164:92-104.

de la Torre F, Cañas RA, Pascual MB, Ávila C and Cánovas FM (2014b) *Plastidic aspartate aminotransferases and the biosynthesis of essential amino acids in plants*. J Exp Bot 65:5527-5534.

Dornfeld C, Weisberg AJ, Ritesh KC, Dudareva N, Jelesko JG and Maeda HA (2014) *Phylobiochemical characterization of class-Ib aspartate/prephenate aminotransferases reveals evolution of the plant arogenate phenylalanine pathway*. Plant Cell 26:3101-3114.

Dyer WE, Henstrand JM, Handa AK and Herrmann KM (1989) *Wounding induces the first enzyme of the shikimate pathway in* Solanaceae. Proc Natl Acad Sci USA, 86:7370-7373.

Edgar RC (2004) *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucleic Acids Res, 32:1792-1797.

Emanuelsson O, Nielsen H and Von Heijne G (1999) *ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites*. Protein Sci, 8:978-984.

Emanuelsson, O, Nielsen H, Brunak S and Heijne G (2000) *Predicting subcellular localization of proteins based on their N-terminal amino acid sequence*. J Mol Biol, 300:1005-1016.

Fischer R and Jensen R (1987a) *Prephenate dehydratase (monofunctional)*. Methods Enzymol, 142:507-512.

Fischer R and Jensen R (1987b) *Prephenate dehydrogenase (monofunctional)*. Methods Enzymol, 142: 503-507.

Franceschi VR, Krokene P, Christiansen E and Krekling T (2005) *Anatomical and chemical defenses of conifer bark against bark beetles and other pests*. New Phytol, 167:353-375.

Gietz RD and Schiesti RH (2007) *Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method*. Nat Protoc, 2:38-41.

Görlach J, Raesecke HR, Rentsch D, Regenass M, Roy P, et al. (1995) *Temporally distinct accumulation of transcripts encoding enzymes of the prechorismate pathway in elicitor-treated., cultured tomato cells*. Proc Natl Acad Sci USA, 92:3166-3170.

Grant GA (2006) *The ACT domain: a small molecule binding domain and its role as a common regulatory element*. J Biol Chem, 281:33825- 33829.

Guindon S and Gascuel O (2003) *A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood*. Syst Biol 52:696-704.

Guindon S, Lethiec F, Duroux P and Gascuel O (2005) *PHYML Online–a web server for fast maximum likelihood-based phylogenetic inference*. Nucleic Acids Res, 33:557-559.

He X, Anderson JC, del Pozo O, Gu YQ, Tang X and Martin GB (2004) *Silencing of subfamily I of protein phosphatase 2A catalytic subunits results in activation of plant defense responses and localized cell death*. Plant J, 38:563-577.

Huang T, Tohge T, Lytovchenko A, Fernie AR and Jander G (2010) *Pleiotropic physiological consequences of feedback-insensitive phenylalanine biosynthesis in* Arabidopsis thaliana. Plant J, 63:823-835.

Janzik I, Preiskowski S and Kneifel H (2005) *Ozone has dramatic effects on the regulation of the prechorismate pathway in tobacco* (Nicotiana tabacum L. cv. Bel W3). Planta, 223: 20-27.

Kawaichi S, Yoshida T, Sako Y and Nakamura R (2015) *Draft Genome Sequence of a Heterotrophic Facultative Anaerobic Thermophilic Bacterium*, Ardenticatena maritima Strain 110ST. Genome Announc, 3: pii: e01145-15. doi: 10.1128/genomeA.01145-15.

Keeling PJ (2013) *The number, speed, and impact of plastid endosymbiosis in eukaryotic evolution*. Annu Rev Plant Biol 64:583-607.

Keith B, Dong XN, Ausubel FM and Fink GR (1991) *Differential induction of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase genes in* Arabidopsis thaliana *by wounding and pathogenic attack*. Proc Natl Acad Sci USA, 88:8821-8825.

Kuroki GW and Conn EE (1988) *Purification and characterization of an inducible aromatic amino acid-sensitive form of chorismate mutase from* Solanum tuberosum L. *tubers*. Arch Biochem Biophys, 260:616-621.

Maeda H and Dudareva N (2012) *The shikimate pathway and aromatic amino Acid biosynthesis in plants*. Annu Rev Plant Biol, 63:73-105.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, et al. (2010) *RNAi suppression of Arogenate Dehydratase1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22:832-849.

Maeda H, Yoo H and Dudareva N (2011) *Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate*. Nat Chem Biol 7:19-21.

Mathews CK, van Holde KE, Appling DR and Anthony-Cahill SJ (2012) *Biochemistry, 4th edn*. New Jersey: Prentice Hall.

Nakai K and Horton P (1999) *PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization*. Trends Biochem Sci, 24:34-36.

Pohnert G, Zhang S, Husain A, Wilson DB and Ganem B (1999) *Regulation of phenylalanine biosynthesis. Studies on the mechanism of phenylalanine binding and feedback inhibition in the* Escherichia coli *P-protein*. Biochemistry, 38:12212-12217.

Rippert P and Matringe M (2002) *Purification and kinetic analysis of the two recombinant arogenate dehydrogenase isoforms of* Arabidopsis thaliana. Eur J Biochem 269:4753-4761.

Rippert P, Puyaubert J, Grisollet D, Derrier L and Matringe M (2009) *Tyrosine and phenylalanine are synthesized within the plastids in* Arabidopsis. Plant Physiol, 149: 1251-1260.

Ritz C and Spiess AN (2008) *qPCR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis*. Bioinformatics, 24:1549-1551.

Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, et al. (2005) *Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes*. Curr Biol, 15:1325-1330.

Romero RM, Roberts MF and Phillipson JD (1995) *Anthranilate synthase in microorganisms and plants*. Phytochemistry, 39:263-276.

Siehl DL and Conn EE (1988) *Kinetic and regulatory properties of arogenate dehydratase in seedlings of* Sorghum bicolor (L.) Moench Arch Biochem Biophys, 260:822-829.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011) *MEGA5: molecular evolutionary genetic analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods*. Mol Biol Evol 28:2731-2739.

Tan K, Li H, Zhang R, Gu M, Clancy ST and Joachimiak A (2008) *Structures of open (R) and close (T) states of prephenate dehydratase (PDT)–implication of allosteric regulation by L-phenylalanine*. J Struct Biol 162:94-107.

Tzin V and Galili G (2010) *New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants*. Mol Plant, 3:956-972.

Voinnet O, Rivas S, Mestre P and Baulcombe D (2003) *An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus*. Plant J, 33:949-956. Erratum in: Plant J, 2015, 83:752.

Warpeha KM, Lateef SS, Lapik Y, Anderson M, Lee BS and Kaufman LS (2006) *G-protein-coupled receptor 1, G-protein Galpha-subunit 1, and prephenate dehydratase 1 are required for blue light-induced production of phenylalanine in etiolated* Arabidopsis. Plant Physiol, 140:844-855.

Warren RL, Keeling CI, Yuen MM et al. (2015) *Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism*. Plant J, 83:189- 212.

Yamada T, Matsuda F, Kasai K, Fukuoka S, Kitamura K, et al. (2008) *Mutation of a rice gene encoding a phenylalanine biosynthetic enzyme results in accumulation of phenylalanine and tryptophan*. Plant Cell, 20:1316-1329.

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, et al. (2013) *An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase*. Nat Commun, 2833. doi:10.1038/ncomms3833.

Paper 3

New insights into the metabolic regulation of phenylalanine biosynthesis in plants

Fernando de la Torre, Jorge El-Azaz, Concepción Ávila and Francisco M. Cánovas.

Short communication draft

INTRODUCTION

Aromatic amino acids (AAAs) phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) are of paramount importance for all forms of live. However, only some living organisms, including bacteria, fungi and plants, have the necessary biochemical pathways for their biosynthesis. In accordance with the importance of these pathways, evolution has selected alternative and overlapping regulatory mechanisms enabling a fine control of the metabolic flux through these pathways. Furthermore, the biosynthesis of AAAs requires a tight coordination with downstream metabolic pathways to guarantee a proper supply of precursors. In plants, Phe serves as precursor for the biosynthesis of a wide range of phenolic derivatives such as phenylpropanoids, flavonoids, anthocyanins or lignin.

Two alternative routes of the Phe biosynthesis have been described (**Figure 1**): i) in the arogenate pathway, prephenate is transaminated by prephenate-aminotransferase (PAT) to generate arogenate which is next decarboxylated and dehydrated by arogenate dehydratase (ADT) to give Phe (Bonner et al., 1987); ii) in the phenylpyruvate pathway, present in

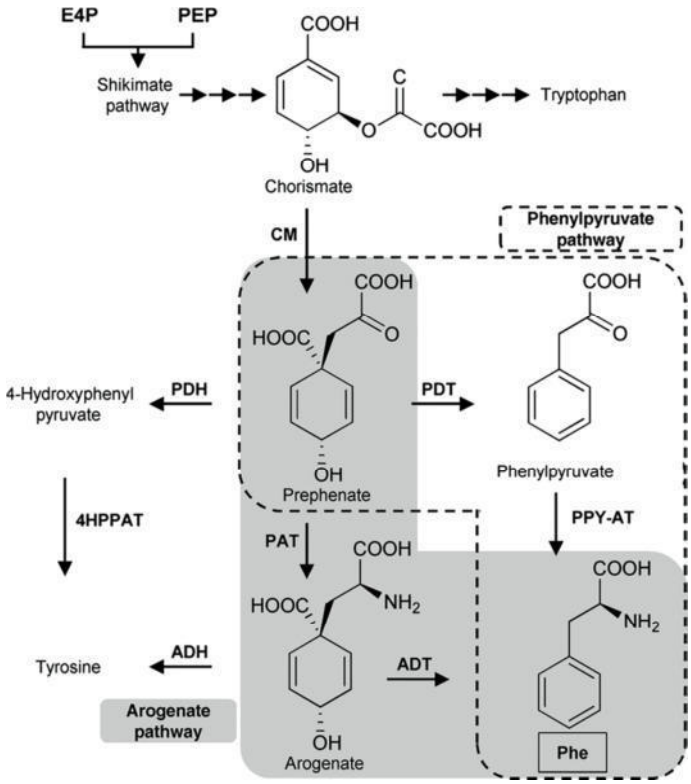
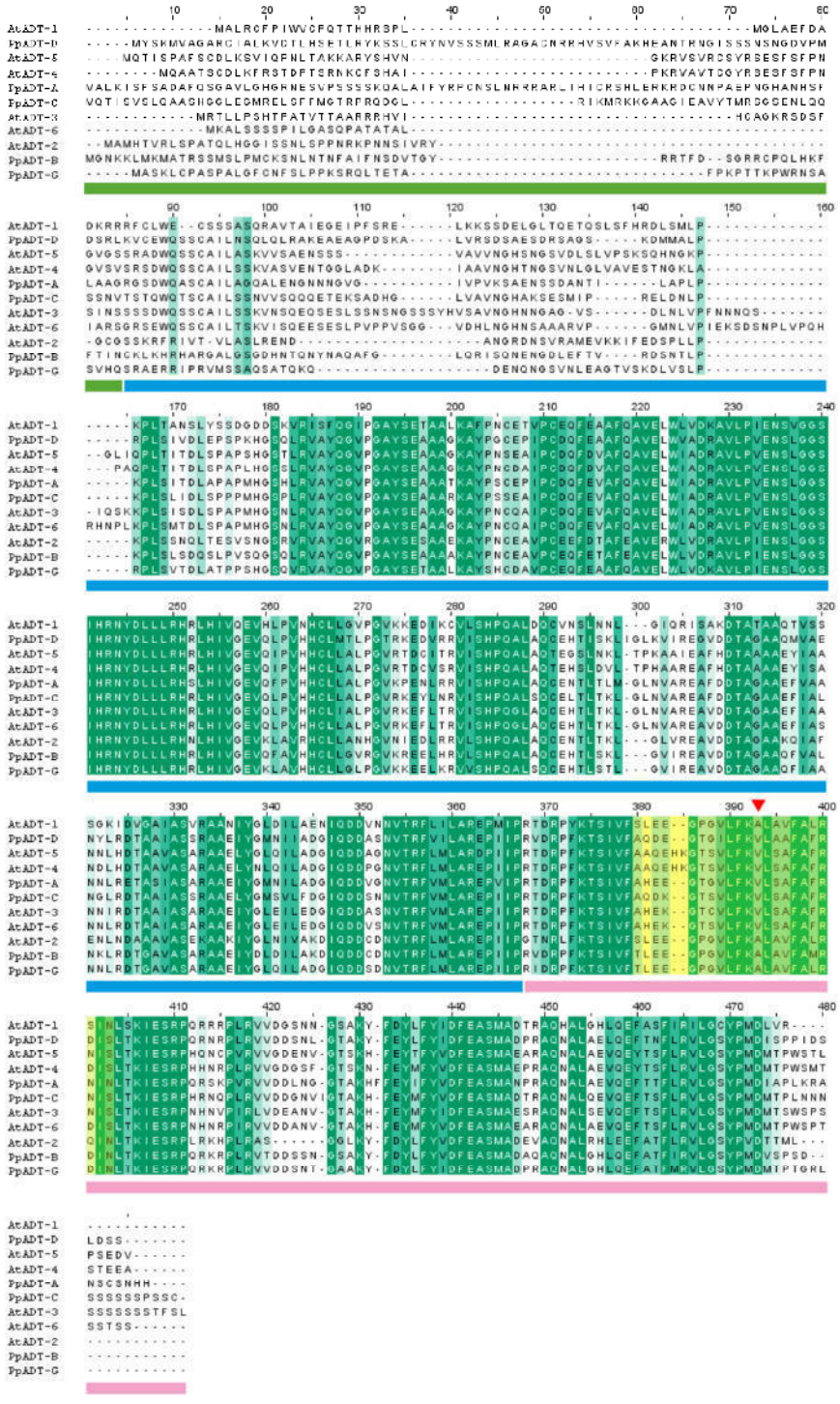


Figure 1. The phenylalanine biosynthetic pathways in plants. Shaded in gray, the arogenate pathway; encircled by dashed line, the phenylpyruvate pathway. Abbreviations: ADH, arogenate dehydrogenase; ADT, arogenate dehydratase; CM, chorismate mutase; E4P, erythrose 4-phosphate; 4HPPAT, 4-hydroxyphenylpyruvate aminotransferase; PAT, prephenate aminotransferase; PDH, prephenate dehydrogenase; PDT, prephenate dehydratase; PEP, phosphoenol pyruvate; and PPY-AT, phenylpyruvate aminotransferase.

most microorganisms, prephenate dehydratase (PDT) catalyzes the decarboxylation and dehydratation of prephenate rendering phenylpyruvate. Phenylpyruvate is later transaminated to Phe through a phenylpyruvate-aminotransferase (Fischer et al., 1987a). Prephenate and arogenate are also employed as precursors for the biosynthesis of Tyr through the activity of prephenate dehydrogenase and arogenate dehydrogenase, respectively (Fischer et al., 1987b; Bonner et al., 1987).

Plant ADTs have two well-defined domains (**Figure 2**): an N-terminal catalytic domain and a C-terminal ACT regulatory domain firstly characterized in aspartate kinase-chorismate mutase- TyrA. The ACT-domains are directly associated to a wide range of metabolic enzymes that are regulated by changes in amino acid concentration. In plants, ADT activity is positively regulated by Tyr and negatively by Phe (Tzin and Galili, 2010). Consistently, Phe levels are largely increased in *A. thaliana* (Huang et al., 2010) and *Oryza sativa* (Yamada et al., 2008) plants carrying point mutations in the regulatory ACT domain of ADTs, which result in feedback-insensitive enzymes.

Different reports have shown that some plant ADTs are also able to catalyze the PDT reaction *in vitro* (Cho et al., 2007; Maeda et al., 2010; Bross et al., 2011; El-Azaz et al., 2016). In a previous work we have demonstrated that in these bifunctional ADT/PDT enzymes, PDT activity is determined by a small protein domain named PAC-domain, for PDT Activity Conferring-domain (El-Azaz et al., 2016). Interestingly, the PAC-domain is included in the C-terminal ACT regulatory domain out of the catalytic domain (**Figure 2**). Furthermore, we demonstrated that the single exchange of this domain conferred an enhanced PDT activity to a monofunctional enzyme that exclusively exhibits ADT activity *in vivo* (El-Azaz et al., 2016). Therefore, the catalytic activity of the protein was altered without modifying its catalytic domain. Moreover, the single mutation of a critical residue (Ala^{314}) within the PAC domain was sufficient to abolish PDT but not ADT activity *in vivo* in the bifunctional enzyme (El-Azaz et al., 2016).



◄ Figure 2. Protein sequence aligment and domain structure of ADTs/PDTs from *Arabidopsis thaliana* and *Pinus pinaster*. Protein domains have been marked as described by Cho et al. (2007) in: green, putative chloroplast transit peptide; blue, ADT/PDT catalytic domain; pink, ACT regulatory domain. Residues with a high degree of conservation are shown in teal colour. PDT activity conferring domain (PAC-domain) is marked in yellow over the ACT domain. The residue Ala^{314} in the PAC-domain, which as been proposed to be essential for an efficient PDT catalytic activity (El-Azaz et al., 2016), is marked with a red triangle. Plants abbreviations: At, *Arabidopsis thaliana*; Pp, *Pinus pinaster*.

RESULTS AND DISCUSSION

To clearly identify the role displayed by the PAC domain and Ala^{314} in conferring PDT activity we proceeded to the computer modeling of the 3D structure of *Pinus pinaster* PpADT-G, a bifunctional PDT/ADT enzyme. For this purpose, we have selected the previously characterized 3D structure of the T-state of *Chlorobium tepidum* PDT (PDB: 2QMX) (Tan et al., 2008) as a template for the protein modeling since the amino acid sequence of this protein is 49% identical to PpADT-G. The SWISS-MODEL automated protein structure homology-modelling server was employed (Arnold et al., 2006; Kiefer et al., 2009). The complete sequence for PpADT-G was larger than that of the template CtPDT. Therefore, the N-terminal of the protein, including the plastid-targeting peptide, and 10 residues in the C terminus were not included in the model. The region under analysis consisted of the amino acid residues from Val^{111} to Tyr^{389} in the PpADT-G primary sequence. Along this polypeptide fragment, PpADT-G and CtPDT are 50,2% identical. The Ramachandran plot analysis showed that most of the amino acids within the model, 93%, are included in the most favored region (Chen et al., 2010).

Our model represents the first molecular model for an ADT/PDT in plants. According to it, the enzyme exists as a homodimer and the way the two subunits might fit together is illustrated in **Figure 3A**. Each monomer consists of an N-terminal catalytic domain (residues 111-285) and a smaller C-terminal regulatory ACT domain (residues 295-389). The ACT domains of both monomers contact each other through a small hydrophobic interface. The overall folding of the PpADT-G model resembles the known 3D structures for PDTs from *C. tepidum* and *Staphylococcus aureus* (Tan et al., 2008).

Based on this model we proceeded to the *in silico* simulation of the Gmut^{11} PpADT-G (El-Azaz et al., 2016) mutant protein by replacing Ala^{314} with Val^{314} using the mutagenesis tool in PyMOL, followed by energy-minimizations. Although this involves only a minor change, since Ala and Val are both small aliphatic amino acids, the resulting mutant protein does not exhibit an efficient PDT activity, whereas the ADT activity housed by this enzyme does not seem to be significantly affected (**Figure 3B**). Using the PyMOL tools for predicting collision between atoms, we investigated putative alterations within the ACT domain as a result of a minor increase in the side chain of the residue at position 314, derived from the Ala to Val mutation. Our calculations indicate that the mutant Val^{314} in Gmut^{11} is located in a hydrophobic environment within the ACT domain and clashes with Phe^{302}, Leu^{311}, Leu^{376} and Phe^{379}, four hydrophobic residues corresponding to the same polypeptide (**Figure 3B**). Among these, Phe^{302}, Leu^{311} and Phe^{379} are conserved in all ADTs while Leu^{376} is exclusive of ADTs displaying PDT activity. However, an essential role for this residue on PDT activity is not likely since its mutation to Val in Gmut^{5} had no effect on that activity (El-Azaz et al., 2016). Leu^{311} is located within the PAC domain

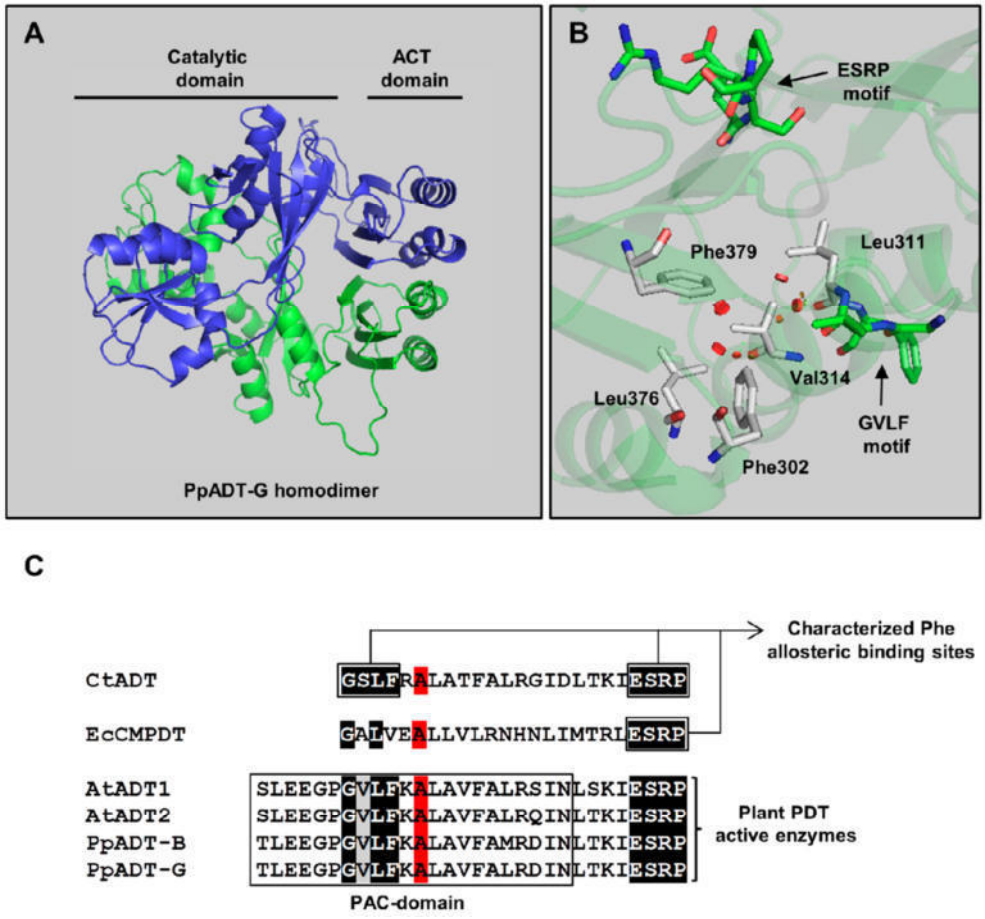


Figure 3. Computer modelling of PpADT-G. (A) Cartoon representation of tertiary and quaternary structure of PpADT-G; each color, green and blue, corresponds to a protein chain. **(B)** Structural simulation of mutant protein PpADT-G^{A314V}. Red disks indicate significant van der Waals overlap between mutant Val^{314} and other amino acids. GVLF and ESRP; allosteric Phe-binding sites. **(C)** Sequence alignment of the PAC domain region within the ACT domain of plants ADTs that exhibit PDT activity, ADT from *Chlorobium tepidum* (CtADT) and bifunctional chorismate mutase/prephenate dehidratase from *Escherichia coli* (EcCMPDT). The residues from putative Phe binding sites that are identical to CtADT are indicated in black; the residue proposed as essential for an efficent PDT activity, Ala^{314}, is marked in red.

and is part of a small protein sub-domain, GVLF, highly conserved in ADT enzymes from plants but also from bacteria such as *Escherichia coli* or *C. tepidum* (**Figure 3C**) (Tan et al., 2008). Different studies have shown in *E. coli* (Pohnert et al., 1999) and *C. tepidum* (Tan et al., 2008) that this domain plays a critical role as an allosteric binding site for Phe. Additionally, a second allosteric binding site for Phe, ESRP, within the ACT domain of *E. coli* (Pohnert et al., 1999) and *C. tepidum* (Tan et al., 2008) is conserved in all plant ADTs and PDTs and is also located closely to the PAC domain, based on the 3D modeling of PpADT-G (**Figure 3C**). According to the proposed model of Phe-mediated feedback inhibition of CtADT, upon Phe binding at the ACT dimer, this domain rearranges from an open to a closed conformation. These changes are propagated to the catalytic domain, which is connected through extensive interactions with the ACT domain, resulting in a reduction of the access of the substrate to the catalytic site (Tan et al., 2008).

Our *in silico* modelization indicates that the mutation of Ala^{314} could alter the structure of the GVLF domain involved in Phe recognition by ACT domain, suggesting a correlation between the strong decrease in PDT activity and the allosteric regulation of the enzyme. By contrast, mutation of the Gly residue in the GVLF sequence of PpADT-G, had no effect on PDT activity (El-Azaz et al., 2016).

We have demonstrated that the Ala^{314} point mutation strongly reduces enzyme affinity towards prephenate and PDT activity, but does not seem to have the same effect on ADT activity, indicating that this mutation is not affecting the functionality of the catalytic domain. This result strongly suggests that ADT and PDT activities are determined through different regulatory mechanisms within the ACT domain. *In vitro* determination of

PDT activity is routinely carried out with purified recombinant proteins and in the presence of the single reaction substrate, prephenate. Thus, we hypothesize that prephenate displays a dual role as substrate and allosteric activator of PDT activity through an unknown sensing mechanism that involves the PAC domain. In accordance with this mechanism, Phe would be synthesized in plants mainly through the arogenate pathway. However, under metabolic conditions resulting in increased prephenate levels, as occurs when PAT is not present, this metabolite would act as a trigger for PDT activity and, therefore, of its own conversion into phenylpyruvate for Phe biosynthesis. Consistently, Maeda et al., (2011) have shown that the blocking of the arogenate pathway as a consequence of *PAT* silencing in *P. hybrida* results in a minor reduction of Phe (15-20%) with a simultaneous 4-fold accumulation of prephenate. Similarly, we observed in *N. benthamiana* that the blocking of the arogenate pathway was not corelated to a reduction in Phe levels (de la Torre et al., 2014).

We have previously described that all plant species analyzed have retained bifunctional ADT/PDT enzymes, suggesting an essential role for PDT activity throughout the life cycle of plants (El-Azaz et al., 2016). In this paper, we propose that plants have preserved PDT activity and, subsequentely, an alternative pathway for Phe biosynthesis from phenylpyruvate, through the conservation of a positive allosteric regulation mechanism mediated by prephenate.

References

Arnold K, Bordoli L, Kopp J and Schwede T (2006) *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling*. Bioinformatics 22:195-201.

Bonner C and Jensen R (1987) *Prephenate aminotransferase*. Methods Enzymol, 142:479-487.

Bross CD, Corea OR, Kaldis A, Menassa R, Bernards MA and Kohalmi SE (2011) *Complementation of the pha2 yeast mutant suggests functional differences for arogenate dehydratases from* Arabidopsis thaliana. Plant Physiol Biochem, 49:882-890.

Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, et al. (2010) *MolProbity: all-atom structure validation for macromolecular crystallography*. Acta Crystallogr D Biol Crystallogr, 66:12-21.

Cho MH, Corea OR, Yang H, Bedgar DL, Laskar DD, et al. (2007) *Phenylalanine biosynthesis in* Arabidopsis thaliana. *Identification and characterization of arogenate dehydratases*. J Biol Chem, 282:30827-30835.

de la Torre F, El-Azaz J, Avila C and Cánovas FM (2014) *Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism*. Plant Physiol, 164:92-104.

El-Azaz J, de la Torre F, Ávila C and Cánovas FM (2016) *Identification of a small protein domain present in all plant lineages that confers high prephenate dehydratase activity*. Plant J, 87:215-229.

Fischer R and Jensen R (1987a) *Prephenate dehydratase (monofunctional)*. Methods Enzymol, 142:507-512.

Fischer R and Jensen R (1987b) *Prephenate dehydrogenase (monofunctional)*. Methods Enzymol, 142:503-507.

Huang T, Tohge T, Lytovchenko A, Fernie AR and Jander G (2010) *Pleiotropic physiological consequences of feedback-insensitive phenylalanine biosynthesis in* Arabidopsis thaliana. Plant J, 63:823-835.

Kiefer F, Arnold K, Künzli M, Bordoli L and Schwede T (2009) *The SWISS-MODEL Repository and associated resources*. Nucleic Acids Res, 37(Database issue):D387-92.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, et al. (2010) *RNAi suppression of Arogenate Dehydratase1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22:832-849.

Maeda H, Yoo H and Dudareva N (2011) *Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate*. Nat Chem Biol, 7:19-21.

Pohnert G, Zhang S, Husain A, Wilson DB and Ganem B (1999) *Regulation of phenylalanine biosynthesis. Studies on the mechanism of phenylalanine binding and feedback inhibition in the* Escherichia coli *P-protein*. Biochemistry, 38:12212-12217.

Tan K, Li H, Zhang R, Gu M, Clancy ST and Joachimiak A (2008) *Structures of open (R) and close (T) states of prephenate dehydratase (PDT)--implication of allosteric regulation by L-phenylalanine*. J Struct Biol, 162:94-107.

Tzin V and Galili G (2010) New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. Mol Plant, 3:956-972.

Yamada T, Matsuda F, Kasai K, Fukuoka S, Kitamura K, et al. (2008) *Mutation of a rice gene encoding a phenylalanine biosynthetic enzyme results in accumulation of phenylalanine and tryptophan*. Plant Cell, 20:1316-1329.

Paper 4

Identification and transcriptional regulation of arogenate dehydratase genes linked to lignin biosynthesis in maritime pine

Jorge El-Azaz et al.,

Research paper draft

Summary

Plants rely on the biosynthesis of phenylalanine as building block for the synthesis of proteins but also as precursor for a great variety of plant-derived compounds essential for their grown, development and defense. Polymerization of secondary cell wall in trees involves the massive biosynthesis, among others, of the phenylalanine-derived polymer lignin. Thus, these plants require an accurate metabolic coordination between phenylalanine and lignin biosynthesis to ensure its normal development. In this study, we have found that the enzyme arogenate dehydratase, that catalizes the last step in phenylalanine biosynthetic pathway in plants, is transcriptionally regulated through direct interaction with the transcription factor PpMYB8. The transcriptional analysis of pine plants silenced for *PpMyb8* suggest that this transcription factor is directly involved in secondary cell wall biogenesis and cell death processes. Taken together, these results indicate that PpMYB8 is a strong candidate to act as a coordinator of lignin accumulation and the proper biosynthesis of its essential precursor, phenylalanine.

INTRODUCTION

Phenylalanine (Phe) biosynthesis in plants is an exceptionally important process, since this amino acid has a doubly essential function as building block for proteins and as the main precursor for the biosynthesis of phenylpropanoids, a wide range of aromatic compounds that play a key role in plant growth, development and stress responses. Land colonization by the first terrestrial plants, one of the most important evolutionary steps in the history of life on Earth, would not have been possible without the emergence of specialized metabolic pathways for the biosynthesis of these secondary metabolites (Lowry et al., 1980). The metabolism of Phe is essential in the carbon channeling from photosynthesis to phenylpropanoids biosynthesis. It is particularly relevant in trees, such as conifers, that divert large amounts of carbon into these Phe-derived compounds, mainly lignin, an important constituent of wood (Craven-Bartle et al. 2013). A strong regulation of the Phe metabolic flux should be expected depending on its alternative use for protein biosynthesis versus phenylpropanoid biosynthesis. In vascular plants this second fate involves a massive carbon flux with more than 30% of photosynthetically fixed carbon directed via the phenylpropanoid pathway towards the synthesis of specialized metabolites, mainly lignin (Haslam, 1993).

In the early steps of the phenylpropanoid pathway, Phe is deaminated into trans-cinnamic acid by phenylalanine ammonia-lyase (PAL). Subsequently, p-coumaric acid is synthesized by cinnamate-4-hydroxylase (C4H); p-coumaric acid can be esterified to a coenzyme A (CoA) by coumarate ligase activity (4CL) giving p-coumaroyl-CoA, which can be converted into p-coumaroyl alcohol though two consecutive reactions (CCD and CAD). Alternatively, p-coumaric acid can be hydroxylated on the C3 position of the

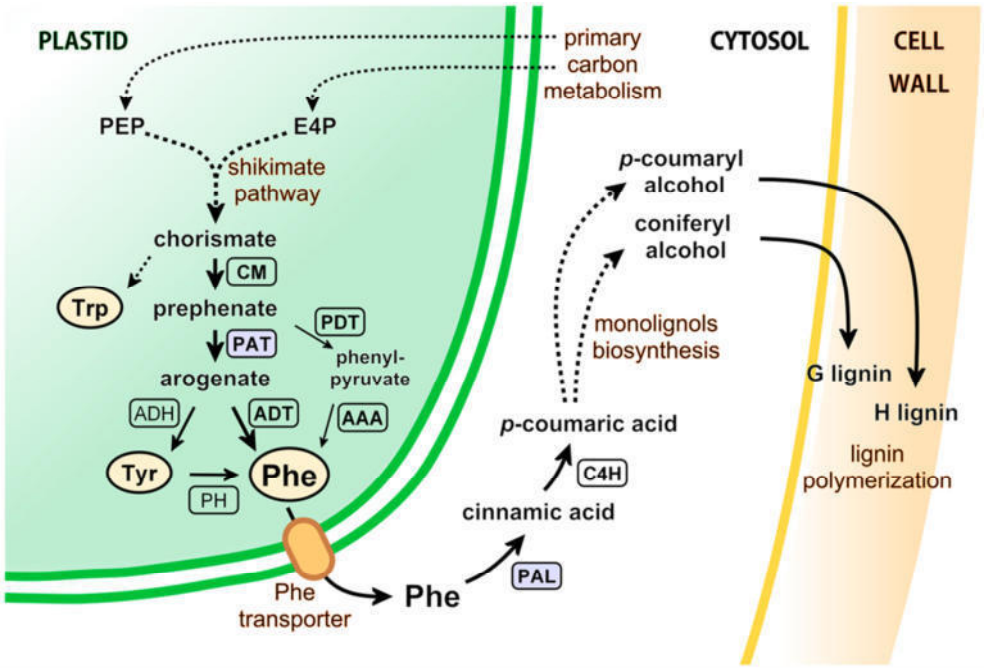


Figure 1. **General scheme of aromatic amino acids and lignin biosynthesis in conifers**. Enzymes names appear inside rectangles. Blue rectangles indicate enzymes coded by genes regulated by PpMYB8 as published by Craven-Bartle et al., 2013. PEP, phospho*enol*pyruvate; E4P, erythrose 4-phosphate; CM, chorismate mutase; PAT, prephenate aminotransferase; PDT, prephenate dehydratase; ADT, arogenate dehydratase; ADH, arogenate dehydrogenase; AAA, aromatic amino acid aminotransferase; PH, phenylalanine hydroxylase; PAL, phenylalanine ammonia-lyase; C4H, cinnamic acid 4-hydroxylase.

aromatic ring to render caffeic acid, a reaction that is considered takes place mainly at the level of the coumaroyl shikimate/quinate ester (Vanholme et al 2013). The caffeate moiety, either as free caffeic acid or as caffeoyl-CoA, can be methylated on the hydroxyl group to yield ferulic acid or feruloyl-CoA, which through additional reactions finally renders coniferyl alcohol. Both monolignols, p-coumaroyl alcohol and coniferyl alcohol, will be respectively the building blocks for H lignins and G lignins biosynthesis

in conifers (**Figure 1**). Due to lack of feruloyl-5-hydroxylase activity (F5H) S-lignins monomer, sinapyl alcohol, is absent in conifers lignin (Wagner et al., 2015; Pascual et al., 2016).

As mentioned above, Phe metabolism plays a central role in the channeling of carbon from photosynthesis to the biosynthesis of phenylpropanoids. Inside the plastids, plants primary carbon metabolism is connected with the biosynthesis of the three aromatic amino acids, Phe, Tyr and Trp, via the shikimate pathway (**Figure 1**). This pathway starts from two compounds: phospho*enol*pyruvate, from glycolisis, and erythrose-4-phosphate, from pentose phosphate pathway, which are finally converted to chorismate, a direct precursor for the biosynthesis of aromatic amino acids and other compounds, such as vitamins K1 and B9 (**Figure 1**; Tzin and Galili, 2010). Chorismate is alternatively used by anthranilate synthase in the starting reaction of the Trp biosynthetic pathway, or by chorismate mutase to generate prephenate for the biosynthesis of Phe and Tyr (Cotton and Gibson, 1965; Maeda and Dudareva, 2012). Phe and Tyr are synthesized from prephenate by two concatenated reactions: decarboxylation and transamination. For the Phe biosynthetic branch, both reactions can equally occur alternatively: i) transamination from prephenate to arogenate by the enzyme prephenate aminotransferase (PAT), and then decarboxylation and dehydratation by arogenate dehydratase (ADT) to give Phe (Bonner and Jensen, 1987a); ii) first decarboxylation and dehydration from prephenate to phenylpyruvate by the enzyme prephenate dehydratase (PDT), and then transamination from phenylpyruvate to phenylalanine through an aromatic amino acid aminotransferase (Fischer and Jensen, 1987a; **Figure 1**). Similarly, prephenate and arogenate are also precursors for the biosynthesis of Tyr through the activity of prephenate dehydrogenase and arogenate dehydrogenase, respectively (**Figure 1**; Fischer and Jensen,

1987b; Bonner and Jensen, 1987b). While in most microorganisms and fungi the biosynthesis of Phe occurs through the phenylpyruvate-dependent pathway, the arogenate pathway has been proposed as predominant for Phe biosynthesis in plants (Maeda et al., 2010; 2011) (**Figure 1**). Nevertheless, in the last years some authors have proposed the co-existence of a functional phenylpyruvate pathway in plants (Yoo et al., 2013; El-Azaz et al., 2016). Some members of the ADT enzyme family can also display PDT activity, although with lower catalytic efficiency, and therefore are able to be involved in both alternative pathways.

Regulation of phenylpropanoids and lignin biosynthesis in plants is a complex process that involves the coordinated expression of genes encoding enzymes located in different subcellular compartments and cellular types. Vascular plants from herbaceous species to trees share a relatively conserved ancestral xylem transcriptome (Li et al 2010). Thus, our current knowledge highlights a general pyramid-shaped regulatory scenario in which a limited set of transcription factors (TFs), placed at the top of the regulatory network, act as master regulators of a larger set of downstream TFs with more specific regulatory functions over particular cell wall biosynthetic genes (Nakano et al., 2015). This hierarchical model is based in two extensively characterized TF families: NACs and MYBs. In this model, NACs TFs act as master regulators, controlling the expression of MYBs TFs, one of the largest TF families found in plants (Martin and Paz-Ares, 1997).

Members of the R2R3-MYB family regulate lignification through interaction to AC-enriched elements present in the promoter regions of phenylpropanoid and lignin biosynthetic genes (Zhong and Ye, 2009). In trees, several studies have demonstrated the implication of various

transcription factors in wood formation. EgMYB2 from *Eucalyptus grandis* bind AC elements in the promoter of their target genes in developing wood and very recently, two additional eucalyptus MYB TFs, EgMYB88 and EgMYB1, have been associated with the biosynthesis of lignin (Soler et al., 2016; Soler et al., 2017). In poplar, a deep dissectional analysis has also revealed the implication of multiple NAC and MYB TFs regulating secondary wall biosynthesis during wood formation (Zhong et al., 2011). Regarding to conifers, PtMYB1, PtMYB4 and PtMYB8 regulates phenylpropanoid metabolism and secondary cell wall biogenesis in *Pinus taeda* by regulating multiple genes encoding phenylpropanoid enzymes involved in lignin monomer synthesis (Bomal et al., 2008).

Despite of its importance in plant metabolism as a potential bottleneck of carbon flux towards secondary metabolism, little attention has been paid to the importance of Phe biosynthesis (Corea et al., 2012). This affirmation is particularly true in woody plants such as conifers. Craven-Bartle et al. (2013) reported the capacity of the R2R3-MYB TF PpMYB8 from *Pinus pinaster* to co-activate the expression of PAT and PAL in pine protoplasts, showing its ability to specifically bind a conserved AC-II element in the promoter region of these genes. Consequently, *PpMyb8* expression level was shown to be higher in lignifying stem tissues. These results support the role of PpMYB8 as a regulator of Phe biosynthesis during wood formation.
According to Maeda et al (2011), ADT catalyzes the rate-limiting step of phenylalanine biosynthesis in plants. However, the regulatory mechanisms underlying the expression of ADT genes have been poorly characterized, specifically those connected with the channeling of Phe that is necessary for lignification. In this study we have characterized two ADT genes, *PpADT-A* and *PpADT-D,* highly expressed during lignification in maritime pine. Transcriptomic analysis in lignifying tissues of transgenic trees and *in*

vivo and *in vitro* interaction experiments indicate that PpMYB8 coordinates the expression of *PpADT-A* and *PpADT-D* with the lignification process. This constitutes the first direct association between two *ADT* genes and lignin biosynthesis in plants.

RESULTS

Transcriptional profile of *PpADT* genes in compression wood

Xylogenesis involves the biosynthesis of large amounts of lignin and therefore requires a constant Phe biosynthetic flux. In gymnosperms, an specific woody tissue that is known as compression wood (CW) develops on the underside of branches and leaning stems, while opposite wood (OW) develops on the opposite side (Groover, 2016). Compared to OW, CW contains more lignin and less cellulose and represents a good model to study the transcriptional regulation of pathways providing substrates during xylogenesis. In this study we have investigated which *PpADT* isoforms are predominant in CW of *P. pinaster* and therefore potential candidates to be involved in lignin biosynthesis. The expression levels of *ADT* genes were analyzed by qPCR in xylem from CW and OW of 25-year-old trees mechanically stressed to produce compression wood (Villalobos et al., 2012) (**Figure 2**). Interestingly, 6 of the 9 ADT genes, *PpADT-A*, *PpADT-D*, *PpADT-E*, *PpADT-F*, *PpADT-G* and *PpADT-I* showed significantly higher expression in CW compared to OW. Among these, *PpADT-A, PpADT-D* and *PpADT-G* showed the highest increases (**Figure 2**). Enhanced expression levels in CW *vs* OW were also found for other genes encoding for enzymes directly involved in the biosynthesis and metabolic utilization of Phe such as chorismate mutase isoforms 1 and 2 (*PpCM1*, *PpCM2*),

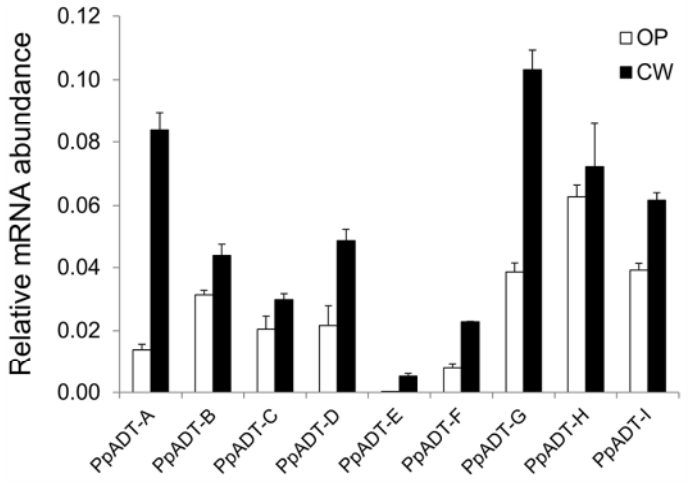


Figure 2. ***PpADT*** **genes relative mRNA levels in opposite wood (OP; white bars) compared to compression wood (CW; black bars) in** ***Pinus pinaster*** **wood samples**. Total RNA was extracted from OP and CW tissue and transcript levels were determined by qPCR, as described in the materials and methods section. Values represent means of three technical replicates for the same biological sample. Error bars represent SD. Asterisks indicate *t*-test significant differences between OP and CW (p-Value < 0.01).

prephenate aminotransferase (*PpPAT*), phenylalanine hydroxylase (*PpPH*) or phenylalanine ammonia lyase *PpPAL1* (Supplementary Figure 1).

Regulation of *PpADTs* by PpMYB8

In a previous study, we demonstrated a fundamental role of PpMYB8 in the regulation of genes involved in Phe metabolism during the xylogenesis process induced in compression wood (Craven-Bartle et al., 2013). In order to identify *ADT* genes putatively involved in this process, maritime pine plants silenced for *PpMyb8* were generated via somatic embryogenesis. qPCR analysis of 16-months-old plants showed a significant silencing in

PpMyb8 of 80-90% when compared to controls (**Figure 3A**). We proceeded to check whether the expression level of *PpADT* genes was modified in stems silenced for *PpMyb8*. As result, we observed that the silencing of *PpMyb8* resulted in a parallel and significant down-regulation of three *ADT* genes: *PpADT-A*, *PpADT-D* and *PpADT-I* (**Figure 3B**). Interestingly, *PpADT-A* and *PpADT-D* are also included among the most up-regulated *ADT* genes in CW suggesting a direct link between *ADT* regulation by PpMYB8 and xylogenesis. To further explore the impact of *PpMyb8* silencing, the expression levels of other genes encoding for enzymes directly involved in Phe biosynthesis were also determined. The results showed a significant downregulation of *PpPAT*, *PpCM2* and

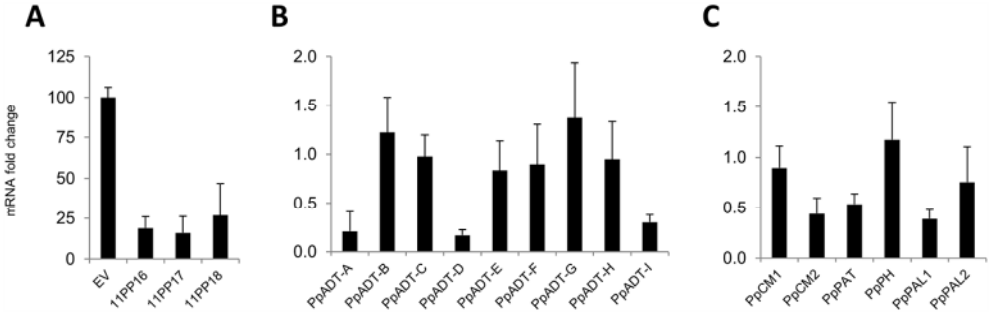


Figure 3. qPCR expression analysis in pine seedlings silenced for *PpMyb8*. Total RNA was extracted from stems of 16-months-old plants (*n* = 3 biological replicates). **(A)** *PpMyb8* mRNA suppression in pine stems from the independent silenced lines analyzed (11PP16, 11PP17, 11PP18) relative to corresponding levels in control plants (EV, *n* = 2), which were set as 100%. **(B)** Expression profiling for the ADT family. mRNA abundancy is expressed as fold change in decimals compared to control. **(C)** mRNA levels fold change for other Phe metabolism closed related genes. *PpCM1*, chorismate mutase 1; *PpCM2*, chorismate mutase 2; *PpPAT*, prephenate aminotransferase; *PpPH*, phenylalanine hydroxylase; *PpPAL1*, phenylalanine ammonia-lyase 1; *PpPAL2*, phenylalanine ammonia-lyase 2. Analysis was done with three different biological samples. Error bars represent SD. Asterisks indicate *t*-test significant differences respect to control plants (p-Value < 0.01).

PpPAL1 in *PpMyb8* silenced plants (**Figure 3C**), according with a hypothetical role of PpMYB8 as co-regulator of multiple genes involved in the Phe synthetic pathway and the convertion of this amino acid into phenylpropanoids.

In silico identification and analysis of *PpADT* regulatory regions

Considering the lack of a fully assembled genome of *P. pinaster*, the reference genome of the closely related conifer *Pinus taeda (*V1.01 Genomic Scaffolds: http://dendrome.ucdavis.edu/resources/blast/; Wegrzym et al., 2014) was searched in order to identify the 5'-flanking regions of the *P. taeda ADT* genes based on their *P. pinaster* homologues. The 5'-flanking regions of the *P. taeda* homologues were used to perform a nucleotide BLAST search on the non-assembled *P. pinaster* genomic reads (SustainPineDB: http://www.scbi.uma.es/sustainpine/), which were manually assembled using the *P. taeda* sequences as template (**Figure 4**). The length of the sequences assembled ranged from a minimum of 517 bp for *PpADT-B*, to a maximum of 4.804 bp for *PpADT-E*.

In silico and manual analysis of the *PpADT* 5'-flanking regions showed the presence of putative R2R3-MYB binding sites in all the sequences collected, with the sole exception of *PpADT-B* and *PpADT-H.* A close-up analysis of these 5'-flanking sequences showed an enrichment of putative R2R3-MYB biding sites in the proximal region of *PpADT-A* and *PpADT-D* (**Figure 4**). The identified 5'-flanking region of *PpADT-A* contains 1907 bp upstream the predicted translation start codon, presenting three AC elements in its proximal region: from -307 to -301 (AC-III element), from -252 to -247 (AC-II element) and from -209 to -182 (tandem repeated AC-II element). The 5'-flanking region of *PpADT-D* comprises 1226 bp from the

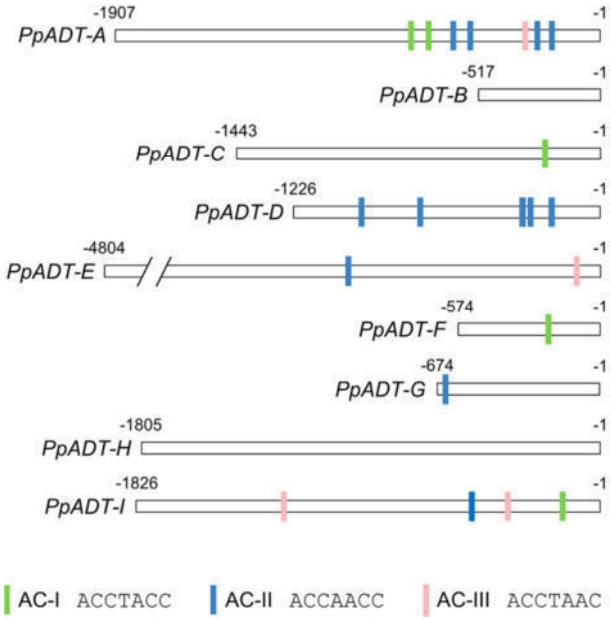


Figure 4. **Comparative *in silico* analysis of *P. pinaster ADT* putative promoters sequences**. Genomic DNA sequences upstream of the initiation codon were manually assembled from the information available at *SustainPineDB*, using their *P. taeda* homologues as a reference. Different colours, as indicated in the key, were used to indicate the three different AC elements previously described as putative PpMYB8 binding sites. Every *cis*- element is presented relative to its position within each promoter.

predicted translation start codon, with three AC elements in its proximal region, all of them belonging to AC-II class: from -305 to -298, from -288 to -283 and from -213 to -208 (**Figure 4**).

Isolation of *PpADT-A* and *PpADT-D* regulatory regions and binding assay of PpMYB8.

PpADT-A and *PpADT-D* 5'-flanking regions were amplified from genomic DNA of *P. pinaster* plantlets, sequenced and cloned into the pJET1.2 vector. To investigate whether the putative cis-AC elements found in these 5'-upstream regions are crucial for PpMYB8 recognition, we used an electrophoretic mobility shift assay (EMSA) to establish a physical link between PpMYB8 two partially overlapping DNA probes for each *–A* and *–D* 5'-upstream regions, spanning from positions: -225 to -329 (probeA1 fragment, 105-bp lenght), -281 to -370 (probeA2, 90-bp), -253 to -331 (probeD1, 66-bp) and -192 to -318 (probeD2, 127-bp).

A clear band shift was observed for the probeA2 fragment in the presence of PpMYB8 protein, which was strongly reduced when unlabeled competitor DNA is added (**Figure 5A**). In contrast, no shift was detected for probeA1 fragment. EMSAs also revealed a clear interaction of PpMYB8 with probeD1 but not probeD2 fragment (**Figure 5A**). Similarly, the intensity of the band shift was reduced in the presence of unlabeled competitor DNA. These results strongly suggest the existence of a physical interaction between PpMYB-8 and the *PpADT-A* and *–D* 5'-upstream regulatory regions at the positions selected for the analysis.

Interaction of PpMYB8 and *PpADT* regulatory regions in yeast *LacZ* reporter assay

To confirm the above described specific interaction between PpMYB8 and *PpADT* putative regulatory regions *–A* and *–D*, and to further delimit the PpMYB8 binding site, we performed a β-galactosidase reporter assay using

baker's yeast (*Saccharomyces cerevisiase*). For this purpose, sequential deletion derivatives from the 5'-border of *PpADT-A* and *PpADT-D* putative regulatory regions were cloned into the Gateway® yeast promoterless vector pMW#3 (Reece-Hoyes and Walhout, 2012) upstream of the *LacZ* reporter gene. *PpMyb8* was cloned into the Gateway® yeast expression plasmid pDEST22 generating a PpMYB8 recombinant protein with the GAL4 activation domain fused to its N-terminus. *S. cerevisiae* W303 strain was co-transfected with different combinations of the *LacZ* repoter gene under control of the *PpADT –A* and *–D* regulatory regions and the *PpMyb8* expression construct, along with its respective controls. To test how sequential deletions affect PpMYB8 binding capability *in vivo*, β-galactosidase activity was measured in liquid culture using ortho-Nitrophenyl-β-galactoside (ONPG) as the substrate.

A full length construction for *PpADT-A* (PA.1) and successive 5'- deletion sequences at -638 (PA.2), -394 (PA.3) and -246 (PA.4) were assayed (**Figure 5B**). Thus, significant increases in β-galactosidase activity were detected for *PpADT-A* sequential truncations PA.2 (-638) and PA.3 (-394) (**Figure 5B**). By contrast, β-galactosidase activity was strongly reduced in yeast cells carrying the PA.4 (-246) deletion. These results are consistent with previous EMSAs indicating that AC elements at -301 and -247 displays a critical role in the transcriptional regulation of *PpADT-A* by PpMyb8. No significant increase in β-galactosidase activity was detected for the PA.1 construct due to the high background levels provided by the *PA1:LacZ* construct (data not shown). Similarly, the 5'-upstream region of *PpADT-D* (PD.1) was deleted at -331 (PD.2) and -205 (PD. 3) to analyze its interaction to PpMYB8. No significant β-galactosidase activity was detected for any of the constructs, being unable to confirm the *in vivo* interaction between the *PpADT-D* 5'-upstream region and PpMYB8 (data not shown).

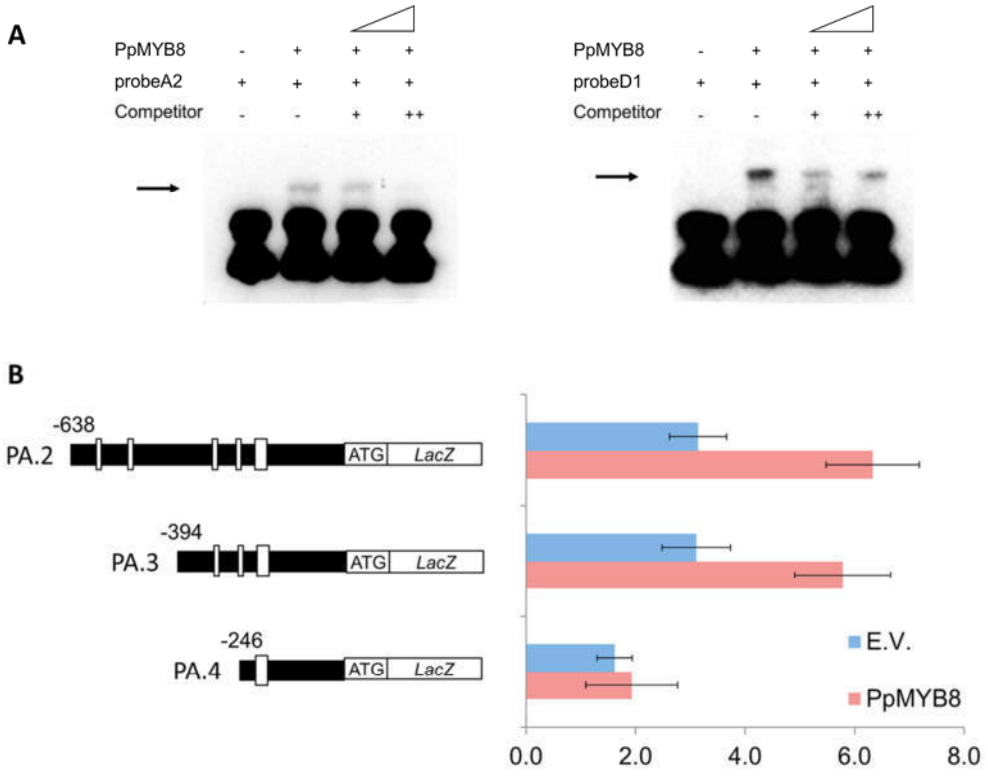


Figure 5. Physical interaction between PpMYB8 and the putative regulatory regions of *PpADT-A* and *PpADT-D*. (A) EMSA using nucleotide probes from *PpADT-A* (-370/-281) and *PpADT-D* (-331/-253) 5' upstream regulatory region and PpMYB8. The formation of a Protein/DNA complex in the presence of recombinant PpMYB8 is indicated with an arrow. Lane 1 indicates probe alone; lane 2 probe with recombinant PpMYB8; lane 3 and 4, competitition with excess unlabeled probe in two different ratios. **(B)** PpMYB8-mediated activation of the *PpADT-A* promoter as determined in yeast. White boxes are indicating putative AC biding elements in the *PpADT-A* promoter. *E.V.*, empty prey vector; *PpMYB8*, *MYB8* expression construct**.** β-galactosidase reporter assays were performed in the *Saccharomyces cerevisiae* W303 strain. Three different deletions of the *PpADT-A* promoter region, named as -638, -394 and -246, were cloned into pMW#3 bait vector and assayed in the presence of *pDEST22:PpMyb8* expression construct under control of the ADH yeast promoter.

Microarray analysis of plants overexpressing *PpMyb8*

To more deeply identify the transcriptional network under the control of PpMYB8 we performed transcriptomic analysis of transgenic plants of maritime pine over-expressing this TF. A 60-mer oligonucleotide microarray (PINARRAY3) developed in our lab and manufactured by Agilent (Santa Clara, CA, USA; Cañas et al., 2015) was used. Six microarray fields were hybridized with 3 biological replicates corresponding to stems from wt and transgenic plants. Our analysis render 1,134 spots considered to represent differentially expressed (DE) genes (*adjP.value* < 0.05; logFC > 0.58). Among these genes, 962 showed up-regulation in plants overexpressing *PpMyb8*, compared to controls, whereas 172 where down-regulated. We focus our analysis on the set of up-regulated genes in order to identify targets of PpMyb8 involved in the biosynthesis of Phe, monolignols/lignin and secondary cell wall. We identified the over-expression of additional transcription factors including a R2R3-Myb TF (sp_v3.0_unigene27138) or a mini-zinc finger (sp_v3.0_unigene30450). We also detected the induction of three genes encoding for WD40 repeat proteins (sp_v3.0_unigene103660, sp_v3.0_unigene57748 and sp_v3.0_unigene9308), putative modulators of the transcriptional activity of MYBs (Ramsay et al., 2005). In regard to the biosynthesis of Phe we identified up up-regulation ($Log_2FC=3.82$) of a gene encoding for 3-phosphoshikimate 1-carboxyvinyltransferase (EPSP) (sp_v3.0_unigene119387), a key enzyme in the shikimate pathway, and two aromatic amino acid transaminases (sp_v3.0_unigene94 and sp_v3.0_unigene135960) required for the phenylpyruvate-mediated Phe biosynthesis.

Very prominently we identified the overexpression of genes encoding for enzymes directly involved in the biosynthesis and maintenance of secondary *cell wall such as* cytochrome P450 monooxygenase (sp_v3.0_unigene167647), hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase (sp_v3.0_unigene37310), dirigent-like protein (sp_v3.0_unigene30744), O-methyltransferase (sp_v3.0_unigene18221) or xylosidase (sp_v3.0_unigene98277). In addition, we detected a clear enrichment in genes directly involved in the reorganization of the cell wall including pectate lyase (sp_v3.0_unigene120063), cellulase (sp_v3.0_unigene77632), cell wall hydrolase (sp_v3.0_unigene17845), endo-beta-1,4-D-mannanase (sp_v3.0_unigene153148), mannan endo-1,4-beta-mannosidase (sp_v3.0_unigene114483) or beta-1,3-glucanase (sp_v3.0_unigene2417). We also identified a set of up-regulated genes directly involved with cell death such as programmed cell death 6 (sp_v3.0_unigene157363), CAS_CSE1 domain protein (sp_v3.0_unigene75458), DCD (development and cell death) (sp_v3.0_unigene2467) or autophagy protein 5 (sp_v3.0_unigene32020).

DISCUSSION

Plants invest large amounts of metabolic resources in the biosynthesis of a wide range of phenolic compounds, essential in functions such as growth, reproduction, development or defense. Among others, this group of compounds includes flavonoids, anthocyanins, lignin, lignans or suberins. Phe is the common precursor to all these compounds and therefore, its biosynthesis requires a fine regulation at transcriptional, post-translational and metabolic levels.

The biosynthesis of lignin in trees represents a metabolic flux of global importance that requires a massive supply of its essential precursor Phe. During the last decade we have seen a notably increase in our current understanding of the processes associated with xylogenesis, and particularly with lignification in trees. However, the regulatory and metabolic mechanisms involved in the supply of Phe for lignin biosynthesis remains poorly studied. In this paper we have investigated the transcriptional regulation of ADTs during lignification in *P. pinaster*. The use of a conifer model species to perform these studies is particularly relevant due to the quantitative and qualitative importance of lignin biosynthesis in this group of woody plants. The research is particularly important, as ADTs have been proposed to be a rate-limiting step of phenylalanine biosynthesis in plants (Maeda and Dudareva, 2012). Moreover, the existence of a single *PAT* gene in most plants and multiple *ADT* isogenes suggests that the different ADT isoforms could be involved in the biosynthesis of Phe for different and specific metabolic purposes. In this regard, Corea et al. (2012) have demonstrated, by using different combinations of *ADT* knockouts, that the six *Arabidopsis* ADT isoforms profoundly differ in their contribution to lignin accumulation, being AtADT5 the primarily responsible. Also in *Arabidopsis*, ADT isoforms strongly differ in their contribution to anthocyanin accumulation being AtADT2 the major contributor (Chen et al., 2016).

The development of compression wood in conifers provides an excellent model for dissecting the regulatory elements that coordinate pathways providing substrates during xylogenesis, and particularly during lignification. CW is characterized by the deposition of a thicker secondary cell wall, an enhanced contain of lignin, enriched in p-hydroxyphenylpropane units, and reduced levels of cellulose (Groover, 2016). Various studies have also shown that the development of CW is corelated to the up-regulation of

genes encoding for enzymes involved in the building of secondary cell wall. CW induced xylogenesis is also directly related to cell death processes associated with the development of tracheary elements.

In order to identify which pine ADTs could be involved in the biosynthesis of the pool of Phe to be later channeled for the biosynthesis of monolignols and lignin, we compared the expression of these genes in CW *vs* OW. Our results showed increased expression in CW *vs* OW of 6 out of the 9 isoforms: *PpADT-A*, *PpADT-D*, *PpADT-E*, *PpADT-F*, *PpADT-G* and *PpADT-I*. However, *PpADT-A*, and to a lesser extent *PpADT-G* and *PpADT-D*, exhibited the highest induction rate in CW. To confirm how long our analysis was able to detect alterations in the expression levels of enzyme-coding genes involved in Phe biosynthesis, we also determined in the same samples the expression of a set of genes directly linked to this process. Genes encoding for chorismate mutase isoforms 1 and 2 (*PpCM1*, *PpCM2*), prephenate aminotransferase (*PpPAT*), phenylalanine hydroxylase (*PpPH*) or phenylalanine ammonia lyase (*PpPAL1*) were up-regulated in CW compared to OW. Together, these results support that the development of CW is intimately associated with Phe biosynthesis as previously suggested by Craven-Bartle et al 2013.

Previous studies have shown that the conifer MYB8 is a regulator of phenylpropanoid metabolism and secondary cell wall biogenesis. Overexpression of *PtMyb8*, from *P. taeda*, in *Picea glauca* resulted in the up-regulation of a set of genes involved in the biosynthesis of Phe and monolignols and the parallel down-regulation of other genes putatively involved in the same pathway (Bomal et al., 2008). Secondly, it is well documented that the *P. pinaster* PpMYB8 directly binds to an AC-II element in the promoter regions of *PAT* and *PAL1* activating their expression. In this

paper, we have investigated to what extent PpMYB8 is involved in the crucial regulation of ADTs during xylogenesis. We have shown that the silencing of *PpMyb8* in 16-month-old pine plants resulted in the concomitant down-regulation of three *ADT*s: *PpADT-A*, *PpADT-D* and *PpADT-I*. Interestingly, *PpADT-A* and *PpADT-D* genes are transcriptionally activated in CW compared to OW. Moreover, *in silico* analysis of the 5'-upstream region of pine ADT genes have shown the presence of putative R2R3-MYB binding sites in all of them with the exception of *PpADT-B* and *PpADT-H.* This result is consistent with the broadly characterized regulation of phenylpropanoid biosynthesis in plants through MYB TFs (Zhong and Ye, 2007; Nakano et al., 2015; Ye et al., 2015; Lamara et al., 2016). The particular enrichment of R2R3-MYB binding sites in the 5'-flanking region of *PpADT-A* and *PpADT-D* mimics the previously described promoter architecture, and PpMYB8-regulation, of *PpPAT* and *PpPAL1* genes coding for enzymes catalyzing the previous and the following reactions to ADT, respectively (Craven-Bartle et al., 2013). These observations strongly suggest that these ADT isoforms may be involved in the lignin-associated Phe biosynthetic pathway.

To determine whether the regulatory regions of these genes are physically interacting with PpMYB8, a double strategy using EMSAs and yeast reporter assays was carried out. Using both methodologies, it has been demonstrated that *PpADT-A* is directly regulated by PpMYB8 through direct binding to a small region containing AC-elements at positions -301 and -247 from the translation start codon. Using EMSAs we have also demonstrated that PpMYB8 directly binds to the 5'-upstream region of *PpADT-D*, but failed in the *in vivo* assays, suggesting the requirement of a particular cell environment that was not provided in this experiment.

To more deeply identify other genes co-regulated with *PpADT-A* and *PpADT-D* in the lignin-associated Phe biosynthesis pathway, maritime pine plants over-expressing *PpMyb8* were generated via somatic embryogenesis (Trontin et al. 2012). The effectiveness of the over-expression of *PpMyb8* strongly differs from one tissue to another suggesting the existence of alternative regulatory mechanisms of MYBs. In this regard, different studies have demonstrated the important regulatory role exerted by microRNAs on MYBs. The overexpression of miR858a in *Arabidopsis* led to the down-regulation of various MYBs involved in the phenylpropanoid pathway (Sharma et al., 2016). Also in *Arabidopsis*, miR156 negatively regulates anthocyanin accumulation by destabilization of a MYB-bHLH-WD40 transcriptional activation complex (Gou et al., 2011). Further studies are required to analyze the putative existence in conifers of a similar regulatory mechanism heading the biosynthesis of monolignols and lignin. Additional transcription factors including a R2R3-MYB TF (sp_v3.0_unigene27138) and a mini-zinc finger (sp_v3.0_unigene30450) were identified by transcriptome analysis of *PpMyb8* over-expressing lines. We have also found the up-regulation of up to three genes encoding for WD40 repeat proteins. These proteins are associated with MYBs and HLHs conforming complexes that promotes their transcriptional activity (Ramsay et al., 2005). An attractive hypothesis, that requires further investigation, is that these WD40 repeat proteins may participate in the modulation of the activity of the set of MYBs involved in the biosynthesis of Phe, monolignols and lignin. Interestingly, we have found strong up-regulation of a gene encoding for an enzyme essential for Phe biosynthesis, the shikimate enzyme EPSP. Verdonk et al. (2005) demonstrated that the *Petunia hybrida* R2R3-type MYB TF *ODO1* activates the promoter of *EPSP synthase* and its downregulation strongly reduces the volatile emission of the Phe-derived benzenoids by decreasing the synthesis of precursors from

the shikimate pathway. We hypothesize that PpMYB8 may also control Phe-biosynthesis in maritime pine regulating EPSP synthase activity. In a similar manner, PpMYB8 may also control the lignin-associated Phe biosynthesis regulating the expression of two aromatic amino acid aminotransferases putatively involved in the phenylpyruvate pathway.

Our transcriptomic analysis has revealed the overexpression of a broad set of genes encoding for enzymes directly involved in the biosynthesis of Phe, but also in the maintenance and reorganization of secondary cell wall. These results are consistent with an enhanced and/or accelerated xylogenesis process and agree those described by heterologous overexpression of PtMYB8 in *Picea glauca* (Bomal et al., 2008). More research efforts are required to determine whether these genes are directly regulated by PpMYB8, as shown for *PpADT-A* and *PpADT-D*, or their regulation is mediated by other TFs.

The identification of a set of genes directly involved with cell death processes is not surprising if we consider that cell death processes are intimately associated with the development of tracheary elements. Thus, we can conclude that an accelerated secondary cell wall formation, induced by over-expression of *PpMyb8*, leads to cell death events. Together, these results strongly suggest that PpMYB8 displays a complex role coordinating the biosynthesis and metabolic utilization of Phe for development of secondary cell wall but also regulates cell death processes associated with the development of tracheary elements.

EXPERIMENTAL PROCEDURES

Cloning of the *PpADT-A* and *PpADT-D* regulatory regions

The *PpADT-A* and *PpADT-D* 5' upstream regulatory regions sequence were obtained from SustainPineDB super-reads and assemble manually. Primers were designed over this assembled sequence. The promoter was cloned using 100 ng of genomic DNA of *P. pinaster* seedlings as template, through a nested-PCR approach. First PCR was performed using the primers 1830.5Fwd and 1830.7Rvs. This PCR was diluted to 1/50 and 1 µL was taken as template for the second PCR, using the primers 1830.4Fwd and 1830.6Rvs. The nested PCR product was purified from gel using the NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) commercial kit, cloned into pJET1.2/blunt (Thermo Fisher Scientific) and confirmed by sequencing.

Expression analysis

Total RNA isolation and cDNA synthesis was obtained as described previously (de la Torre et al., 2014a). qPCR was performed on a CFX-384 Real Time System (Bio-Rad) with SsoFast EvaGreen Supermix (Bio-Rad) under the following conditions: 95°C for 2 min (one cycle), followed by 95°C for 1 s and 60°C for 5 s (45 cycles). cDNAs corresponding to 10 ng of reverse-transcribed RNA were used as a template for each reaction. Raw fluorescence data from each reaction were fitted to the Mass Action Kinetic 2 model, which requires no assumptions about the amplification efficiency of a qPCR assay (Boggy and Woolf, 2010). The initial target concentrations (D_0 parameter) for each gene were deduced from the Mass Action Kinetic 2

model using the qPCR package for the R environment (Ritz and Spiess, 2008), and normalized to *PpActin2* and *PpEF1*.

Electrophoretic mobility shift assay

The recombinant protein PpMyb8 (FN868598), were produced in *Escherichia coli* BL21-AI™ (Thermo Fisher Scientific) overnight culture at 12º C. The oligonucleotide probes described in Figure 5, and containing the AC elements, were generated by annealing complementary biotinylated oligonucleotides designed to create 5′-biotinilated amplicons. At the end of the incubation period, the DNA–protein complexes were analysed by electrophoresis as previously described (Rueda-López *et al*., 2008). The binding specificity was evaluated using competition experiments with the corresponding non-biotinylated DNA probes.

Plasmid constructions for *PpADT-A* and promoter analysis in yeast

PpADT-A promoter deletions were amplified from the pJET1.2 construct using the reverse primer PADTAAttB1R paired with 4 different forward primers, from: P1AttB4 (-808), P2AttB4 (-638), P3AttB4 (-394) and P4AttB4 (-246). The corresponding PCR products were re-amplified using the primers AttB4 and AttB1R, and cloned into Gateway® pDONR-P4-P1R (Thermo Fisher Scientific) vector using BP Clonase® II enzyme mix. Subsequently, they were recombined into pMW#3 bait vector (Deplancke et al., 2006a; courtesy of Dr. Alberto de Marcos Serrano, from University of Castilla-La Mancha, Spain) using the *Gateway® LR* Clonase® II mix.

The *PpMyb8* prey construct was done amplifying PpMyb8 ORF from a construct previously abailable in our laboratory (Craven-Bartle et al., 2013)

using the primers MYB8AttB1 and MYB8AttB2. The PCR product was re-amplified using the primers AttB1 and AttB2, and cloned into Gateway® pDONR207, and then recombined into Gateway® pDEST22 vector, as described above for the *PpADT-A* promoter bait constructs.

Yeast manipulation and β-galactosidase assays

pMW#3 constructs (bait) carrying the 4 different deletions of the PpADT-A promoter were transformed into *Saccharomyces cerevisiae* W303 strain. For each transformation, 1 µg of each reporter constructs was linearized with FastDigest® ApaI restriction enzyme (Thermo Fisher Scientific) and integrated into yeast genome as described by Deplancke et al. (2006b). 12 yeast bait lines for each reporter construct were analyzed into Yeast Nitrogen Base (YNB) media plates without uracil supplemented with X-Gal, to determine the reporter self-activation level. After 4 days of incubation at 30º C, 2 bait lines with the lowest self-activation level were selected for each bait construct. The bait construct corresponding to the -808 promoter deletion was discarded due to the high self-activation level observed. Yeast colonies from the 3 remaining bait constructs were transformed with the prey construct pDEST22:*PpMyb8* and the corresponding empty vector as control. Yeast transformations were plated into YNB media without uracil and tryptophan. Plates were incubated at 30º C until yeast colonies were grown. A few colonies from each transformation were tested by PCR to confirm the presence of both bait and prey constructs or their respective control empty vectors.

For the β-galactosidase assay with ONPG as substrate, 2 colonies for each bait:prey combination were grown overnight at 30º C and 200 rpm into YNB liquid medium lacking uracil and tryptophan. 0.5 mL of the saturated

overnight cultures was used to inoculate 10 mL of Yeast Peptone Dextrose (YPD) liquid medium previously warm to 30º C. YPD tubes were incubated during 4 hours at 30º C and 200 rpm. After this incubation, 1.5 mL were taken from each tube to perform the β-galactosidase assay with ONPG as substrate, which was done as described by Clontech's Yeast Protocol Handbook (Various authors, 2009). Reactions were incubated from 45 to 90 min until yellow color was visible, stopped with Na_2CO_3 1 M, and absorbance measured at 420 nm.

Generation of *Pinus pinaster PpMyb8* transgenic lines

A cryopreserved embryogenic line (PN519; Breton et al., 2006) was used for genetic transformation and transgenic plant production. All transgenic lines and controls obtained or handled during this work were therefore derived from the same PN519 genetic background to facilitate data comparison. PN519 embryonal-suspensor masses (ESM) were reactivated from the cryopreserved stock following the method described by Morel et al. (2014). Subsequently to thawing and prior to genetic transformation assays, ESM were subcultured for 6 weeks on fresh mLV multiplication medium (Morel et al., 2014) using the plating method (Trontin et al. 2016) in order to stimulate a state of active ESM proliferation.

PpMyb8 cDNA was integrated into the constitutive overexpression cassette (maize ubiquitine promoter; CAMV35S terminator) of vector pMBb7Fm21GW-UBIL (Karimi et al., 2002). This binary vector is harbouring a *bar* gene cassette (CAMV35S promoter, nos terminator) which is effective for selection of transgenic events with phosphinothricin (PPT) in maritime pine (Trontin et al., 2007). The *PpMyb8* overexpression binary construct was transferred into the C58pMP90 strain of *A.*

tumefaciens (Koncz and Schell, 1986) following standard protocols. *Agrobacterium*-mediated genetic transformation of the PN519 embryogenic line was performed using DCR-based media (Gupta and Durzan, 1985). After decontamination with antibiotics (1-2 weeks), ESMs were cultivated for 13 weeks on selection medium. PPT-resistant embryogenic lines (PPT+) were collected and further proliferated on selection medium for 4 weeks and then on standard mLV proliferation medium for 4 to 6 weeks (Morel et al. 2014).

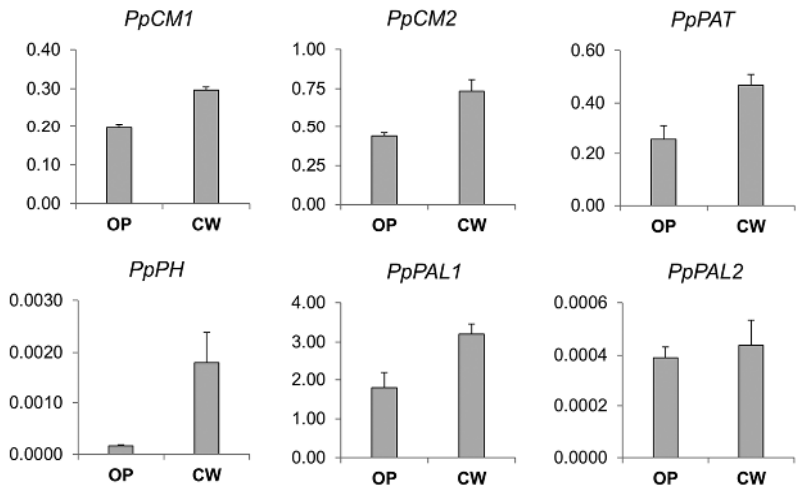
Somatic embryo (SE) development from selected transgenic embryogenic lines was achieved on mLV-based maturation medium within 12 weeks following the method reported in Morel et al. (2014). Cotyledonary SE were harvested under the binocular and stored in the dark at 4°C on maturation medium deprived of abscisic acid until germination assays. Cotyledonary SE from transgenic and control EV lines were germinated *in vitro* for ca. two weeks on a modified DCR medium. Germinated SE were then transplanted into plugs and grown for 4-8 more weeks to obtain rooted plantlets of sufficient size and quality to be progressively acclimatized to soil conditions into the greenhouse.

Acknowledgments

We are grateful to Jean-François Trontin, Sandrine Debille and Francis Canlet from FCBA Technological Institute, Biotechnology and Advanced Forestry Department (71 Route d'Arcachon, Pierroton, 33610 Cestas, France) for the generating of the *Pinus pinaster* transgenic lines. The samples of compression and lateral wood used in this paper were given by Dr. Francisco J. Ruiz Cantón (Universidad de Málaga, Málaga, España) and Dr. David Pacheco Villalobos. We are indebted to Dr. Javier Canales

Carrasco for his helpful advising with microarray analysis. We would also like to give thanks to Dr. Alberto de Marcos Serrano (Universidad de Castilla-La Mancha, España) for the pMW#3 vector. This work was supported by the grants provided by the Spanish MINECO (BIO2015-69285-R) and Junta Andalucía (BIO2012-0474). JEA is the recipient of a Formación Profesorado Universitario fellowship (Ministerio Educación, Spain).

Supplemental data



Supplemental figure 1. **qPCR expression analysis for some phenylalanine and phenylpropanoids biosynthesis genes comparing opposite wood (OP) and compression wood (CW) from *Pinus pinaster*.** *PpCM1*, chorismate mutase 1; *PpCM2*, chorismate mutase 2; *PpPAT*, prephenate aminotransferase; *PpPH*, phenylalanine hydroxylase; *PpPAL1*, phenylalanine ammonia-lyase 1; *PpPAL2*, phenylalanine ammonia-lyase 2. Total RNA was extracted from OP and CW tissue and transcript levels were determined by qPCR as described. Values represent means of three technical replicates for the same biological sample. Error bars represent SD. Asterisks indicate *t*-test significant differences between OP and CW (p-Value < 0.01).

References

Bomal C, Bedon F, Caron S, Mansfield SD, Levasseur C, et al. (2008) *Involvement of Pinus taeda MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis*. J Exp Bot, 59:3925-3939.

Bomal C, Duval I, Giguère I, Fortin É, Caron S, et al. (2014) *Opposite action of R2R3-MYBs from different subgroups on key genes of the shikimate and monolignol pathways in spruce*. J Exp Bot, 65:495-508.

Bonner C and Jensen R (1987a) *Prephenate aminotransferase*. Methods Enzymol, 142:479-487.

Bonner C and Jensen R (1987b) *Arogenate dehydrogenase*. Methods Enzymol, 142:488-494.

Breton D, Harvengt L, Trontin J-F, Bouvet A and Favre J-M (2006). *Long-term subculture randomly affects morphology and subsequent maturation of early somatic embryos in maritime pine*. Plant Cell Tissue Org Cult, 87:95-108.

Cañas RA, Feito I, Fuente-Maqueda JF, Ávila C, Majada J and Cánovas FM (2015) *Transcriptome-wide analysis supports environmental adaptations of two* Pinus pinaster *populations from contrasting habitats*. BMC Genomics, doi: 10.1186/s12864-015-2177-x.

Chen Q, Man C, Li D, Tan H, Xie Y and Huang J (2016) *Arogenate Dehydratase Isoforms Differentially Regulate Anthocyanin Biosynthesis in* Arabidopsis thaliana. Mol Plant, 9:1609-1619.

Corea OR, Ki C, Cardenas CL, Kim SJ, Brewer SE, et al. (2012) *Arogenate dehydratase isoenzymes profoundly and differentially modulate carbon flux into lignins*. J Biol Chem, 287:11446-11459.

Cotton RG and Gibson F (1965) *The biosynthesis of phenylalanine and tyrosine; enzymes converting chorismic acid into prephenic acid and their relationships to prephenate dehydratase and prephenate dehydrogenase*. Biochim Biophys Acta, 100:76-88.

Craven-Bartle B, Pascual MB, Cánovas FM and Avila C (2013*) A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine*. Plant J, 74:755-766.

de la Torre F, El-Azaz J, Avila C and Cánovas FM (2014) *Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism*. Plant Physiol, 164:92-104.

Deplancke B, Mukhopadhyay A, Ao W, Elewa AM, Grove CA, et al. (2006a) *A gene-centered* C. elegans *protein-DNA interaction network*. Cell, 125(6):1193-1205.

Deplancke B, Vermeirssen V, Arda HE, Martinez NJ and Walhout AJ (2006b) *Gateway-compatible yeast one-hybrid screens*. CSH Protoc, 2006(5). pii: pdb.prot4590.

El-Azaz J, de la Torre F, Ávila C and Cánovas FM (2016) *Identification of a small protein domain present in all plant lineages that confers high prephenate dehydratase activity*. Plant J, 87:215-229.

Fischer R and Jensen R (1987a) *Prephenate dehydratase (monofunctional)*. Methods Enzymol, 142:507-512.

Fischer R and Jensen R (1987b) *Prephenate dehydrogenase (monofunctional)*. Methods Enzymol. 142:503-507.

Gou JY, Felippes FF, Liu CJ, Weigel D and Wang JW (2011) *Negative regulation of anthocyanin biosynthesis in* Arabidopsis *by a miR156-targeted SPL transcription factor*. Plant Cell, 23:1512-1522.

Groover A (2016) *Gravitropisms and reaction woods of forest trees - evolution, functions and mechanisms*. New Phytol, 211(3):790-802.

Gupta PK and Durzan DJ (1985) *Shoot multiplication from mature trees of Douglas-fir (*Pseudotsuga menziesii*) and sugar pine (*Pinus lambertiana*)*. Plant Cell Rep, 4:177-179.

Haslam E (1993) *Shikimic Acid. Metabolism and Metabolites*. Von *E. Haslam.* Wiley, Chichester, 387 pp.

Karimi M, Inzé D and Depicker A (2002). *Gateway™ vectors for* Agrobacterium*-mediated plant transformation*. Trends in Plant Science, 7:193-195.

Koncz C and Schell J (1986) *The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of* Agrobacterium *binary vector.* Mol Gen Genet, 204:383-396.

Lamara M, Raherison E, Lenz P, Beaulieu J, Bousquet J and MacKay J (2016) *Genetic architecture of wood properties based on association analysis and co-expression networks in white spruce*. New Phytol, 210:240-255.

Li X, Wu HX and Southerton SG (2010) *Comparative genomics reveals conservative evolution of the xylem transcriptome in vascular plants*. BMC Evol Biol, doi: 10.1186/1471-2148-10-190.

Lowry B, Lee D and Hébant C (1980) *The Origin of Land Plants: A New Look at an Old Problem*. Taxon, 29:183-197.

Maeda H and Dudareva N (2012) *The shikimate pathway and aromatic amino Acid biosynthesis in plants*. Annu Rev Plant Biol, 63:73-105.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, et al. (2010) *RNAi suppression of Arogenate Dehydratase1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22:832-849.

Maeda H, Yoo H and Dudareva N (2011) *Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate*. Nat Chem Biol, 7:19-21.

Martin C and Paz-Ares J (1997) *MYB transcription factors in plants*. Trends Genet, 13:67-73.

Morel A, Trontin JF, Corbineau F, Lomenech AM, Beaufour M, et al. (2014). *Cotyledonary somatic embryos of* Pinus pinaster Ait. *most closely resemble fresh, maturing*

cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses. Planta, 240(5):1075-1095.

Nakano Y, Yamaguchi M, Endo H, Rejab NA and Ohtani M (2015) *NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants*. Front Plant Sci, doi: 10.3389/fpls.2015.00288.

Pascual MB, El-Azaz J, de la Torre FN, Cañas RA, Ávila C and Cánovas FM (2016) *Biosynthesis and Metabolic Fate of Phenylalanine in Conifers*. Front Plant Sci, doi: 10.3389/fpls.2016.01030.

Ramsay NA and Glover BJ (2005) *MYB-bHLH-WD40 protein complex and the evolution of cellular diversity*. Trends Plant Sci, 10:63-70.

Reece-Hoyes JS and Walhout AJ (2012) *Gene-centered yeast one-hybrid assays*. Methods Mol Biol, 812:189-208.

Sharma D, Tiwari M, Pandey A, Bhatia C, Sharma A and Trivedi PK (2016) *MicroRNA858 Is a Potential Regulator of Phenylpropanoid Pathway and Plant Development*. Plant Phys, 171:944-959.

Soler M, Plasencia A, Larbat R, Pouzet C, Jauneau A, et al. (2017) *The* Eucalyptus *linker histone variant EgH1.3 cooperates with the transcription factor EgMYB1 to control lignin biosynthesis during wood formation*. New Phytol, 213:287-299.

Soler M, Plasencia A, Lepikson-Neto J, Camargo EL, Dupas A, et al. (2016) *The Woody-Preferential Gene EgMYB88 Regulates the Biosynthesis of Phenylpropanoid-Derived Compounds in Wood*. Front Plant Sci, 7:1422.

Trontin J-F, Walter C, Klimaszewska K, Park YS and Lelu-Walter MA (2007) *Recent progress on genetic transformation of four* Pinus spp. Transgenic Plant J, 1:314-329.

Tzin V and Galili G (2010) New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. Mol Plant. 3, 956-972.

Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, et al. (2013) *Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in* Arabidopsis. Science, 341:1103-1106.

Various authors (2009) *Clontech: Yeast Protocol Handbook*. VI.D.:*Liquid Culture Assay Using ONPG as Substrate*. pp.: 27. Consulted online in the following link: *www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602*

Villalobos DP, Díaz-Moreno SM, Said el-SS, Cañas RA, Osuna D, Van Kerckhoven SH, et al. (2012) *Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-adenosylmethionine, lignin and lignan related genes*. BMC Plant Biol, doi: 10.1186/1471-2229-12-100.

Wagner A, Tobimatsu Y, Phillips L, Flint H, Geddes B, Lu F and Ralph J (2015) *Syringyl lignin production in conifers: Proof of concept in a Pine tracheary element system*. Proc Natl Acad Sci USA, 112:6218-6223.

Wegrzyn JL, Liechty JD, Stevens KA, Wu LS, Loopstra CA, Vasquez-Gross HA, et al. (2014) *Unique features of the loblolly pine* (Pinus taeda L.) *megagenome revealed through sequence annotation*. Genetics, 196:891-909.

Ye ZH and Zhong R (2015) *Molecular control of wood formation in trees*. J Exp Bot, 66:4119-4131.

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, et al. (2013) *An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase*. Nat Commun, doi: 10.1038/ncomms3833.

Zhong R, Lee C, McCarthy RL, Reeves CK, Jones EG and Ye ZH (2011) *Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors*. Plant Cell Physiol, 52:1856-1871.

Zhong R and Ye ZH (2007) *Regulation of cell wall biosynthesis*. Curr Opin Plant Biol, 10:564-572.

Zhong R and Ye ZH (2009) *Transcriptional regulation of lignin biosynthesis*. Plant Signal Behav, 4(11):1028-1034.

Paper 5

Biosynthesis and metabolic fate of phenylalanine in conifers

María B. Pascual, Jorge El-Azaz, Fernando N. de la Torre, Rafael A. Cañas, Concepción Ávila and Francisco M. Cánovas

Frontiers in Plant Science, July 2016. 7:1030

Summary

The amino acid phenylalanine (Phe) is a critical metabolic node that plays an essential role in the interconnection between primary and secondary metabolism in plants. Phe is used as a protein building block but it is also as a precursor for numerous plant compounds that are crucial for plant reproduction, growth, development, and defense against different types of stresses. The metabolism of Phe plays a central role in the channeling of carbon from photosynthesis to the biosynthesis of phenylpropanoids. The study of this metabolic pathway is particularly relevant in trees, which divert large amounts of carbon into the biosynthesis of Phe-derived compounds, particularly lignin, an important constituent of wood. The trunks of trees are metabolic sinks that consume a considerable percentage of carbon and energy from photosynthesis, and carbon is finally immobilized in wood. This paper reviews recent advances in the biosynthesis and metabolic utilization of Phe in conifer trees. Two alternative routes have been identified: the ancient phenylpyruvate pathway that is present in microorganisms, and the arogenate pathway that possibly evolved later during plant evolution. Additionally, an efficient nitrogen recycling mechanism is required to

maintain sustained growth during xylem formation. The relevance of phenylalanine metabolic pathways in wood formation, the biotic interactions, and ultraviolet protection is discussed. The genetic manipulation and transcriptional regulation of the pathways are also outlined.

Introduction

Animals, including humans, are unable to synthesize all the amino acids required for protein synthesis and metabolic nitrogen (N) homeostasis. These amino acids are termed essential and should be supplied in the diet direct or indirectly from a plant source. Essential amino acids are synthesized in the chloroplasts of photosynthetic organs and other non-green plastids of plants. The capacity of plants to synthesize a set of essential amino acids may be linked to the acquisition of a set of prokaryotic genes that are responsible for this function during the primary endosymbiosis event that resulted in the plastids (de la Torre et al., 2014). These findings may explain the inability of humans and many animals to synthesize essential amino acids and why they should be provided in their foods and feeds. One such dietary amino acid, phenylalanine (Phe), is a critical metabolic node that plays an essential role in the interconnection between the primary and secondary metabolism of plants. Phe is a protein building block and a precursor of numerous compounds that are crucial for plant reproduction, growth, development and defense against different types of stresses. Phe-derived compounds include phenylpropanoids, flavonoids, anthocyanins, lignin, lignans, stilbenes, condensed tannins and the important plant hormone salicylate (**Figure 1**).

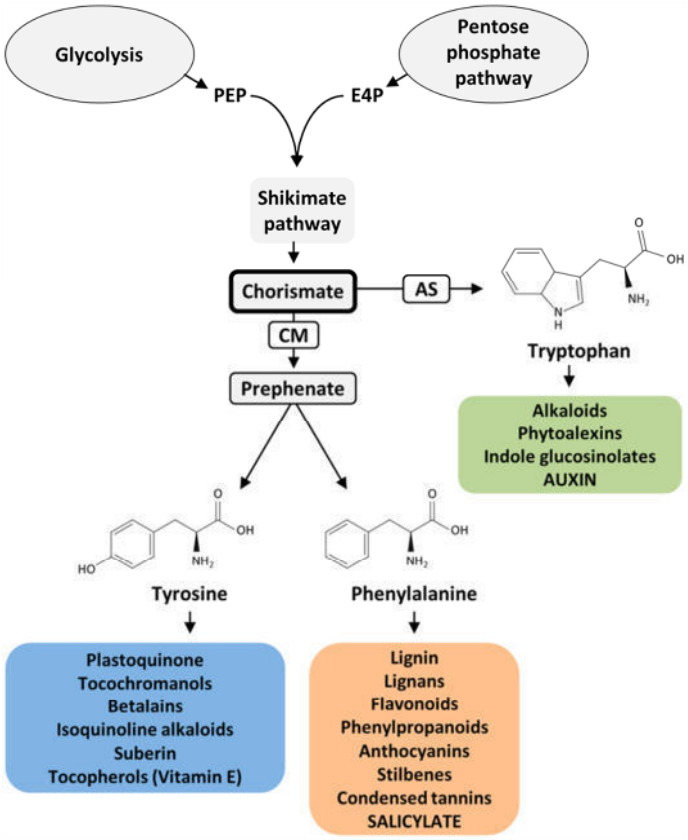


Figure 1. General pathway for aromatic amino acid biosynthesis and derived products. The shikimate pathway connects central carbon metabolism (glycolysis and pentose phosphate pathways) with the biosynthesis of aromatic amino acids and derived products. PEP, phospho*enol*pyruvate; E4P, erythrose 4-phosphate; CM, chorismate mutase; AS, anthranilate synthase.

The metabolism of Phe plays a central role in the channeling of carbon from photosynthesis to the biosynthesis of phenylpropanoids. The study of Phe metabolism is particularly relevant in trees, such as conifers, that divert large amounts of carbon into the biosynthesis of these Phe-derived compounds, particularly lignin, an important constituent of wood. It is estimated that nearly 30–40% of photosynthetically fixed carbon is

channeled through phenylalanine for the biosynthesis of lignin during wood formation. Consequently, the trunks of trees are powerful metabolic sinks of the fixed carbon in photosynthetic tissues that is immobilized in wood (Craven- Bartle et al., 2013). Conifers also exhibit a highly complex secondary metabolism derived from aromatic amino acids, and therefore these trees are interesting models for studying these pathways (Warren et al.,2015). Conifers are the most abundant group of extant gymnosperms and dominate large ecosystems in the northern hemisphere that make an important contribution to global carbon fixation. Coniferous trees are also of great economic importance because they are the primary source for timber and paper production worldwide (Farjon, 2010). Many comprehensive reviews of the biosynthesis of aromatic amino acids in plants were published recently (Tzin and Galili, 2010; Maeda and Dudareva, 2012).

However, less attention has been paid to the importance of Phe metabolism in woody plants. This review presents the biosynthesis and principal metabolic fates of Phe in conifers.

Biosynthesis of phenylalanine, deamination, and nitrogen recycling

The carbon skeletons required for the biosynthesis of aromatic amino acids are channeled from photosynthesis through the amino acids are channeled from photosynthesis through the shikimate pathway toward the biosynthesis of chorismate, a common precursor for the synthesis of Phe, Tyr, and Trp (Maeda and Dudareva, 2012). Chorismate may be used by the enzyme anthranilate synthase in the synthesis of Trp or converted to prephenate, which is the direct precursor for the biosynthesis of Tyr and Phe. Two alternative routes for the Phe biosynthesis have been identified (**Figure 2**). Prephenate is transaminated by prephenate-aminotransferase

(PAT) to generate arogenate in the arogenate pathway (Maeda and Dudareva, 2012). In a next step, arogenate is decarboxylated and dehydrated by the catalytic action of arogenate dehydratase (ADT) to yield Phe (Bonner and Jensen, 1987a). Alternatively, the decarboxylation and dehydratation of prephenate yields phenylpyruvate in a reaction catalyzed by the enzyme prephenate dehydratase (PDT) through the phenylpyruvate pathway. Subsequently, phenylpyruvate is transaminated to Phe by a phenylpyruvate-aminotransferase (Fischer and Jensen, 1987a). The phenylpyruvate pathway is utilized for Phe biosynthesis in microorganisms, and the arogenate pathway has been proposed to predominate in the biosynthesis of Phe in plants (Maeda et al., 2010). However, the functionality of these pathways was analyzed only in a small number of angiosperm species. Prephenate and arogenate are also precursors for the biosynthesis of Tyr via two alternative pathways mediated by prephenate dehydrogenase and arogenate dehydrogenase, respectively (Bonner and Jensen, 1987b; Fischer and Jensen, 1987b). Conifers, unlike angiosperms, possess a plastid-located folate-dependent phenylalanine hydroxylase that catalyzes the interconversion between Phe and Tyr to provide enhanced plasticity to the aromatic metabolism network (Pribat et al., 2010). Finally, Phe is transported outside the plastid to the cytosol and converted into trans-cinnamic acid, which is the first metabolite in the phenylpropanoid pathway, with the concomitant release of an ammonium molecule in the reaction catalyzed by phenylalanine ammonia-lyase (PAL).

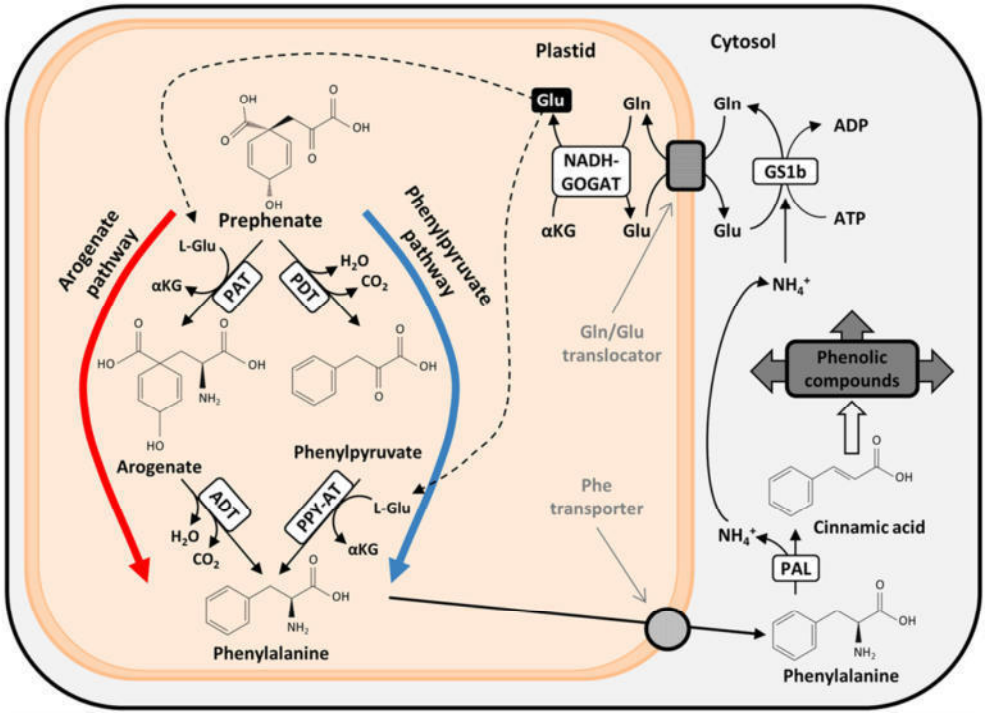


Figure 2. The two alternative pathways for Phe biosynthesis and associated N recycling. PAT, prephenate aminotransferase; PDT, prephenate dehydratase; ADT, arogenate dehydratase; PPY-AT, phenylpyruvate aminotransferase; PAL, phenylalanine ammonia lyase; GS1b, glutamine synthetase 1b; NADH-GOGAT, NADH-dependent glutamate synthase; L-Glu, L-glutamate; L-Gln, L-glutamine; 2-OG, 2-oxoglutarate.

An efficient N recycling mechanism is required to maintain the metabolic flux through the pathway and re-assimilate the released ammonium into glutamate and reincorporate it back into the Phe biosynthetic pathway by prephenate and phenylpyruvate aminotransferases (**Figure 2**; Cánovas et al., 2007; Craven-Bartle et al., 2013). Otherwise, plants undergoing active Phe biosynthesis would require acontinuous input of external N to avoid severe N deficiency. This N recycling is particularly important in trees to

maintain sustained growth in height during xylem differentiation and wood formation.

Phe biosynthesis in the vascular cells of conifers is located in the plastid, and Phe deamination occurs in the cytosol. N recycling involves enzymes in the cytosol (GS1b) and plastid (NADH-GOGAT). Therefore, the actions of a glutamine translocator and a phenylalanine transporter are necessary to provide Phe for a variety of metabolic pathways, including lignin biosynthesis during wood formation. A glutamine/glutamate translocator was characterized in maritime pine cells that may be involved in glutamine import into the chloroplast and glutamate export to the cytosol, which prevents N loss from this essential pathway (Claros et al., 2010). The activity of a phenylalanine transporter in the plastid membrane has not been demonstrated, and the molecular and kinetic characteristics of this amino acid transporter are not known. More knowledge of the regulation of Phe transport during xylem development and in response to a variety of environmental stresses is required to obtain a more complete view of Phe metabolism.

The existence in conifers of a very diverse set of Phe- demanding pathways makes these plants an attractive candidate for a thorough analysis to clarify the routes of biosynthesis of this important amino acid. A multigene family encodes the ADT enzyme in conifers, and this family includes more members in gymnosperms than angiosperms (El-Azaz et al., 2016). ADT genes in maritime pine are differentially expressed, which suggests that these genes play different biological roles. N-terminal plastid-targeting sequences were predicted for all maritime pine ADT enzymes and a plastid location was also experimentally confirmed for many of these enzymes. Some of these ADT enzymes also exhibited PDT activity, which

supports these enzymes as potential candidates in the alternative phenylpyruvate pathway for Phe biosynthesis. El-Azaz et al. (2016) identified a 22-amino acid region in the C-terminal of these enzymes as being responsible for conferring PDT activity (PAC-domain). This small protein domain is of prokaryotic origin and it has been shown to be present in all plant clades. Phylogenetic analyses suggest that various subfamilies of ADTs evolved from an ancestral PDT/ADT subfamily that existed in the most ancient plant clades (El-Azaz et al., 2016). Despite this functional diversification, the ancient pathway using phenylpyruvate as intermediary has been preserved throughout the evolution of plants for Phe biosynthesis. The availability of a diverse catalog of enzymes for the biosynthesis of Phe has likely allowed a more specialized ability to cope with multiple developmental and physiological situations that demand the biosynthesis of Phe- derived compounds. Taken together, the above results suggest that conifers may possess a greater and specialized flexibility for the biosynthesis of Phe that facilitates adaptation to different physiological and environmental conditions. Consistent with the above findings, gymnosperms, compared to angiosperms, exhibit a more phylogenetically diverse set of PAL enzymes that catalyze the first and committed step in the phenylpropanoid pathway (Bagal et al., 2012). The recent release of several conifer genomes will aid in the identification of other putative specificities in the biosynthesis of Phe in these plants and reveal the conservation of these pathways throughout the evolution of land plants.

Phe-derived compounds/phenylpropanoids

Phenylpropanoids are an important group of phenolic plant secondary metabolites derived from Phe although in certain plants can be also derived from Tyr. Phenylpropanoids are compounds that contain a phenyl ring with

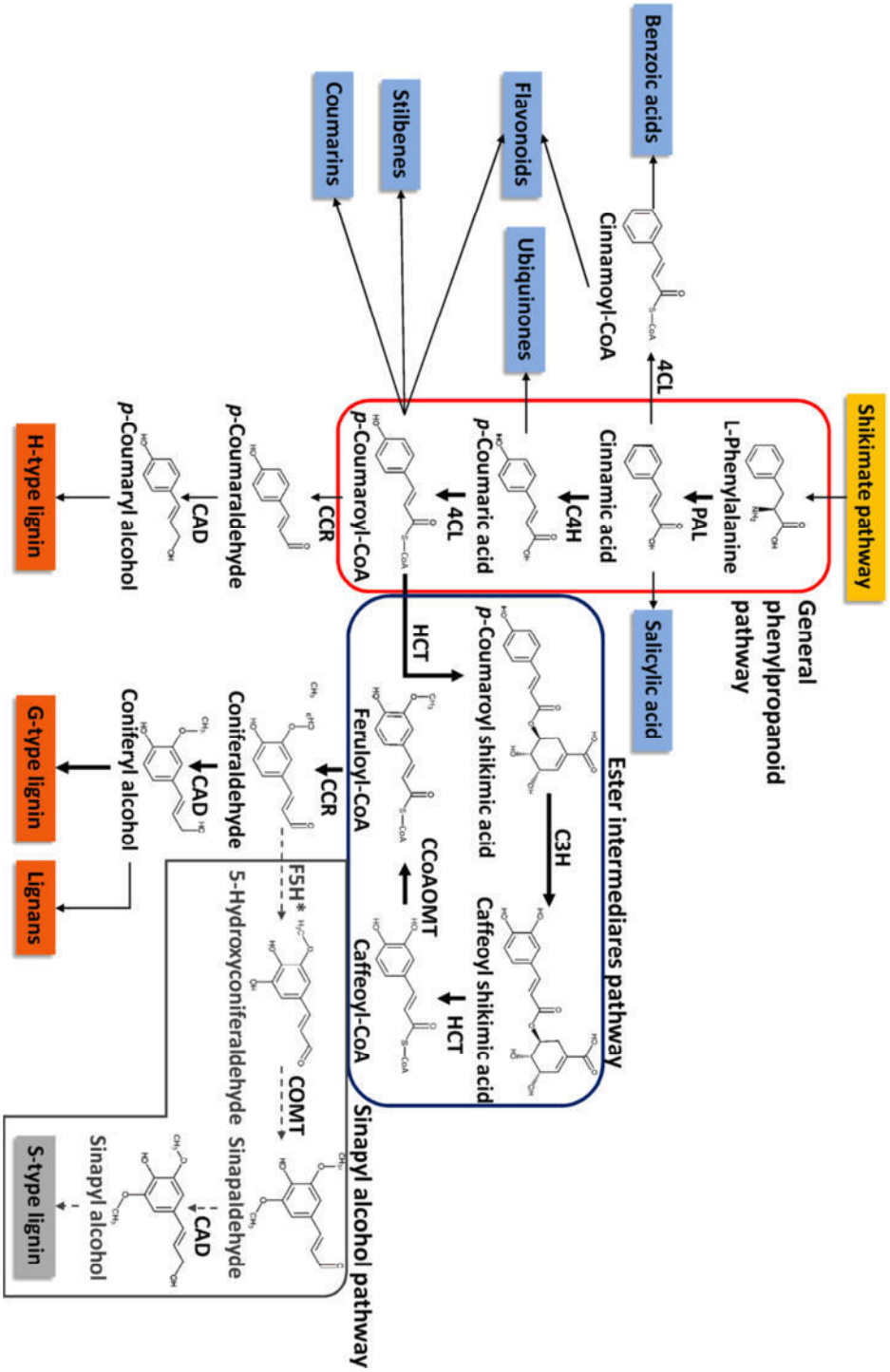
a C3 side chain (Heldt and Piechulla, 2011). This structure is conserved only in monolignols and their metabolic intermediaries, but the term "phenylpropanoids" also include metabolites derived from these compounds and intermediaries of Phe synthesis, such as flavonoids, stilbenes, coumarins, salicylate and gallate, which exhibit very divergent structures. Phenylpropanoids and derivatives can polymerize producing compounds such as lignans, lignin, condensed and hydrolysable tannins, and may be part of suberin and cutin (Heldt and Piechulla, 2011). All of these compounds have important biological activities and serve as cell wall constituents, antibiotics, light protectants and flower pigments and participate in the formation of impermeable layers.

Phenylpropanoids involved in wood formation in conifers

The monolignol synthesis pathway has received much attention because these compounds are the precursors of lignin, which is the most abundant organic compound on Earth, after cellulose. Lignin comprises 30% of the plant biomass, and it is an essential component of wood (Boerjan et al., 2003). Therefore, the synthesis of these metabolites is as powerful sink that uses a considerable percentage of the carbon and energy derived from photosynthesis. Lignin composition and characteristics possess a huge economic importance because they determine wood properties (Boerjan et al., 2003). The metabolic pathway leading to the biosynthesis of lignin was extensively studied in several plant species, including conifers. In contrast, other topics, such as its demonstrated plasticity (Cass et al., 2015) or the precise transcriptional network that regulates these pathways, are much less studied. However these topics are fundamental to our understanding of the multiple mechanisms that are critical for plant development and environmental adaptation. Conifers possess enzymes similar to those of

angiosperms for the synthesis of *p*-coumaryl M1H and coniferyl M1G monolignols but lack the enzymatic machinery needed to synthesize sinapyl M1S subunits (**Figure 3**; see Wagner et al., 2012 for acomprehensive description of the biosynthesis of monolignols in conifers).Recent publications of the genome sequences of *Picea glauca* (Birol et al., 2013; Warren et al., 2015), *P. abies* (Nystedt et al., 2013) and *Pinus taeda* (Neale et al., 2014; Zimin et al., 2014) enabled the identification of full-length gene models for all enzymes in the monolignol biosynthetic pathway, except ferulic acid 5-hydroxylase/coniferaldehyde5-hydroxylase (F5H). This result is consistent with the lack of F5H activity in conifers, which prevents the existence of S units in the lignin of these species (Wagner et al., 2015). The lack in S units yields a more condensed lignin, which is more difficult to remove from the polysaccharide components of the cell wall matrix during industrial processing (Ralph, 2010). This is an important issue because conifer forests dominate large extensions of the Northern hemisphere and provide approximately the 60% of the wood used for industrial purposes (Farjon, 2010). Therefore, the elucidation and manipulation of the phenylpropanoid pathway in conifers is of great interest.

Figure 3 (next page). Simplified scheme of monolignol synthesis pathway in conifers. Thick arrows highlight the route channeling the higher amount of carbon in conifers, the synthesis of coniferyl alcohol. The gray and discontinued arrows highlight the sinapyl alcohol synthesis that does occur in conifers because of the lack a gene encoding the F5H enzyme. Asterisk highlights the lack of F5H enzyme in conifers. Phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate:CoA ligase (4CL), cinnamoyl-CoA reductase (CCR), cinnamyl-alcohol dehydrogenase (CAD), shikimate O-hydroxycinnamoyltransferase (HCT), p-coumarate 3-hydroxylase (C3H), caffeoyl-CoA O-methyltransferase (CCoAOMT), ferulic acid 5-hydroxylase/coniferaldehyde 5-hydroxylase (F5H), and caffeic acid 3-O-methyltransferase (COMT).


Shikimate pathway
General phenylpropanoid pathway
L-Phenylalanine
PAL
Cinnamic acid
C4H
p-Coumaric acid
4CL
p-Coumaroyl-CoA
CCR
p-Coumaraldehyde
CAD
p-Coumaryl alcohol
H-type lignin
HCT
Feruloyl-CoA
CCR
Coniferaldehyde
CAD
Coniferyl alcohol
G-type lignin
Lignans
Ester intermediares pathway
p-Coumaroyl shikimic acid
C3H
Caffeoyl shikimic acid
HCT
Caffeoyl-CoA
CCoAOMT
Sinapyl alcohol pathway
F5H*
5-Hydroxyconiferaldehyde
COMT
Sinapaldehyde
CAD
Sinapyl alcohol
S-type lignin
Salicylic acid
Benzoic acids
Cinnamoyl-CoA
4CL
Flavonoids
Ubiquinones
Stilbenes
Coumarins

Investigation of phenylpropanoid metabolism using genetic manipulation in conifers

The discovery of mutants and the generation of transgenic lines in conifers is extremely difficult compared to herbaceous plants, such as *Arabidopsis*, but monolignol synthesis and lignification in conifers was investigated using transgenic lines and transformed cell cultures induced to differentiate tracheary element (Möller et al., 2003, 2005; Wagner et al., 2005, 2007, 2009, 2011, 2013, 2015; Wadenbäck et al., 2008; Bedon et al., 2009; Tront in et al., 2014).

The first enzyme in the synthesis of monolignols, 4-coumarate-CoA ligase (4CL), was investigated using genetic manipulation (Wagner et al., 2009). The activity of 4CL is essential for secondary metabolism because it is responsible for *p*-coumaroyl-CoA production, which is the precursor of several types of secondary metabolites, such as monolignols, flavonoids, stilbenes, and coumarins (**Figure 3**) (Cheynier et al., 2013). The role of 4CL was examined using the generation of RNAi transgenic lines in *P. radiata* (Wagner et al., 2009). Suppression of 4CL gene expression exerted the biggest impact on lignin production of all of the genetic manipulations of phenylpropanoid related genes in conifers (Wagner et al., 2012). The transgenic plants exhibited a reduced lignin content of approximately 50%. This enzyme acts at the beginning of the metabolic pathway, but the accumulation of certain levels of lignin is explained by the existence of a small 4CL gene family in conifers (Wei and Wang, 2004; Koutaniemi et al., 2007). Notably, the decrease in lignin content in the 4CL RNAi lines was accompanied by an increase in flavonoids and bark (tissue rich in tannins, a flavonoid polymer) and alterations in carbohydrate metabolism (Wagner et al., 2009). These findings support different roles for each 4CL paralog

and, possibly, the ability to redistribute the C surplus caused by a decrease in monolignol synthesis. Therefore, gene co- expression analyses in conifers highlighted the existence of co- regulatory mechanisms between different pathways of secondary metabolism (e.g., synthesis of Phe, monolignols, flavonoids, and terpenoids), which are likely mediated by Myb transcription factors (Cañas et al., 2015; Raherison et al., 2015). The formation of membrane-associated enzyme complexes for the channeling and regulation of metabolite synthesis was suggested (Mast et al., 2010). This may explain why another member of the gene family does not completely replace the role of the 4CL gene after suppression.

Other enzymes that were investigated using genetic manipulation include shikimate *O*-hydroxycinnamoyl transferase (HCT), caffeoyl-CoA *O*-methyltransferase (CCoAOMT), the cinnamoyl-CoA reductase (CCR; Wagner et al., 2007, 2011, 2013; Wadenbäck et al., 2008) and the most widely studied, cinnamyl-alcohol dehydrogenase (CAD; Möller et al., 2003, 2005; Wagner et al., 2005; Bedon et al., 2009; Trontin et al., 2014). Suppression of gene expression generally decreased lignin content, and this decrease was more pronounced when the modifications included enzymes that are involved in the first steps of the pathway. This was accompanied by alterations in the metabolic profiles and composition of lignin. For example, these metabolic changes in the stems of *P. pinaster* CAD RNAi lines affected C metabolism and N metabolism (Trontin et al., 2014). Taken together, the data obtained from all of these cell lines support a metabolic plasticity in lignin biosynthesis and the incorporation of non-traditional compounds to lignin (Wagner et al., 2012). This flexibility may be explained by the existence of gene families that code the enzymes of the phenylpropanoid metabolism and their promiscuity in substrate use. The production of syringyl lignin was achieved recently in conifers using

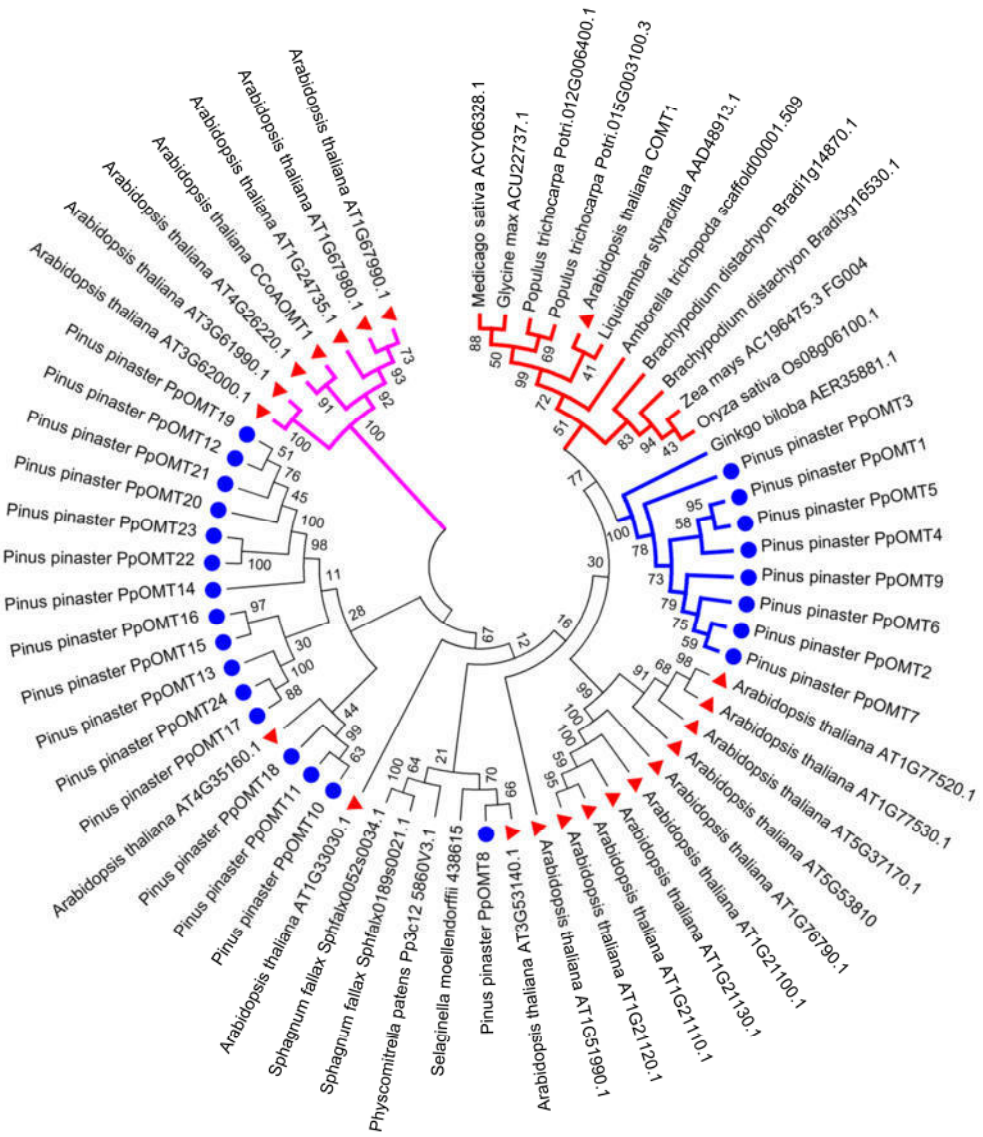
bioengineering approaches (Wagner et al., 2015). The authors of this work manipulated conifer cells for the first time to obtain tracheary elements that synthesized and incorporated sinapyl M1S alcohols into lignin. The synthesis of M1S alcohols in conifers is prevented by the lack of a gene encoding the ferulate 5-hydroxylase/coniferaldehyde 5-hydroxylase (F5H or CAld5H) enzyme. This enzyme transfers a hydroxyl group to the phenolic ring of the coniferaldehyde to produce 5-hydroxylconiferaldehyde, which is methylated by caffeic acid 3-*O*-methyltransferase (COMT) to synthesize sinapaldehyde, which is finally reduced by CAD to generate the sinapyl alcohol (M1S). These authors generated *P. radiata* tracheary elements transformed with F5H or F5H/COMT genes from the angiosperm plant *Liquidambar styraciflua* and produced lignin with S-type units. Notably, some authors considered that conifers lack genes encoding COMT enzymes, but the tracheary elements transformed with an angiosperm F5H gene produced synapyl alcohol at a detectable level (Wagner et al., 2015). This result suggests the existence of COMT activity in conifers, but with lower catalytic capacity, which was shown in the same work. Co-transformation with F5H and COMT resulted in huge amounts of synapyl alcohols compared to single transformants. The lack of the F5H gene may induce a different selection pressure on the *O*-methyltransferases, which previously exhibited COMT activity, and alter the affinity for substrates or the preferential use of new substrates such as flavonoids. A phylogenetic analysis of conifer *O*-methyltransferases shows a group of enzymes that are highly related to the COMT of angiosperms but with an amino acid identity near 55% (**Figure 4**).

Genetic manipulation studies in conifers remain limited despite the above–mentioned reports, and obtaining mutants is extremely complicated because of the long life cycles of conifer species. Only one mutant for a

gene in monolignol metabolism was isolated and characterized, *cad-n1* in *P. taeda* (MacKay et al., 1997). However, the implementation of new genome editing techniques may greatly impact conifer biotechnology. Genome editing in model plant species has undergone extraordinary development in the last few years because of CRISPR/Cas9 technology (Schaeffer and Nakata, 2015). The specificity of this technique for the selection of genome regions to be edited permits the isolation of mutants for specific conifer genes and the substitution of genes with different functions. Mutant lines for determined phenotypes may be generated with truncated or eliminated genes using this technology. Genome editing of the 4-coumarate:CoA ligase (4CL) gene family was performed in *Populus* recently, which extended the use of this powerful CRISPR/Cas9 technology from model plant species to woody perennials (Tsai and Xue, 2015).

Phenylpropanoids involved in biotic interactions in conifers

Most gymnosperms have long generation times, and some species are among the longest living plant species. These trees must cope with a wide range of stresses during their long life cycles, particularly different types of biotic stresses, such as fungi, insects, or herbivorous animals, which generate significant forest damage annually. The bark of gymnosperms functions as the first line of defense to delay or stop the establishment of pathogens. These plants also evolved complex chemical defense mechanisms that are associated with the production of different secondary metabolites including multiple terpenoids and phenolic metabolites (Warren et al., 2015).



◄ Figure 4. Phylogenetic tree of the deduced protein sequences of plant genes encoding caffeic acid 3-O-methyltransferase (COMT). In the tree, the protein sequences correspond to virtually all the sequences of O-methyltransferases from Arabidopsis thaliana, all the sequences of O-methyltransferases, type COMT, found in the Pinus pinaster SustainPineDB and the most similar sequence to Ath-COMT1 from the rest of the shown species. The CLUSTALW program was used for sequence alignments (Thompson et al., 1994). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987) with 1000 bootstrap replications. The optimal tree with the sum of branch length = 12.56422456 is shown. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved 60 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 156 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). The sequences used for the alignments and phylogenetic trees were obtained in Phytozome database (http://phytozome.jgi.doe.gov) and GenBank and, for P. pinaster, in SustainPineDB (http://www.scbi.uma.es/sustainpinedb/). The alignment, tree and accession numbers are available in TreeBase (http://purl.org/phylo/treebase/phylows/study/TB2:S18814). Red triangles correspond to A. thaliana sequences and the blue circles to P. pinaster sequences. Purple branches correspond to the Ath-CCoAOMT group of sequences, red branches to the angiosperm COMT group and blue branches to the gymnosperm COMT-like group.

Conifers possess a set of constitutive defenses against a range of potential attacking organisms. Phenylpropanoids play an essential role in these processes at different levels: (i) deposition of lignin and suberin on tissues enhances resistance to the penetration of small organisms; and (ii) chemical defenses, including phenolics, terpenoids, and alkaloids, are released throughout bark tissues and function as antifeedants and toxins. The cortical parenchyma of conifers contains large amounts of phenolics. Similarly, polyphenolic parenchyma cells are specialized for the synthesis and accumulation of phenolic compounds (Hudgins and Franceschi, 2004)

with antifeedant and antifungal activities (Beckman, 2000). These cells store starch and lipids (Krekling et al., 2000), which makes them attractive bait for pathogens. Lignification of the secondary wall of the sclerenchyma cells can also functions as a mechanical defense against bark beetles. The formation of callus tissue in response to wounding also relies on the production of phenylpropanoids because this tissue becomes lignified, suberized and impregnated with phenolic compounds (Franceschi et al., 2005). Hammerbacher et al. (2014) recently demonstrated that other Phe-derived compounds, proanthocyanidins and flavanols, are produced by *P. abies* in response to fungal infection, and these compounds must be considered chemical defense compounds. Pathogen-induced early lignification of fibers also functions as a structural defense in *Pinaceae* (Hudgins et al., 2003).

The activation of inducible chemical defenses in conifers involves a very efficient broad range system that is highly dependent on the production of phenolics and terpenoids. Different studies demonstrated a correlation among wound- and pathogen-induced damage, polyphenolic parenchyma cell expansion and the accumulation of phenolic compounds (Franceschi et al., 2000; Kusumoto and Suzuki, 2003). This accumulation was also accompanied by a variation in the type of phenolics (Brignolas et al., 1995; Lieutier et al., 1996), which suggests that more specific and toxic compounds for pathogens are produced. These phenolics changes were accompanied by the transcriptional activation of the flavonoid and stilbene pathways (Brignolas et al., 1995). Krekling et al. (2000) demonstrated that seasonal changes also influenced the phenolic content within these cells. Similarly, Danielsson et al. (2011) reported a correlation between the *P. abies* resistance to the fungus *Heterobasidion annosum*, the production of piceasides and flavonoids and the induction of genes in the flavonoid and

proanthocyanidin pathways. The existence of other flavonoids that confer resistance against other fungal pathogens in conifers was also described (Song et al., 2011; Oliva et al., 2015).

A complete understanding of conifers defenses is essential to manage future pests that may affect forests. The roles of different types of Phe-derived compounds in the defense of these plants is becoming more clearly defined, but it still requires of a deeper analysis of the metabolic pathways involved and the technical capacity to characterize and identify these compounds. The recent characterization of various conifer genomes is a fundamental and promising resource (Birol et al., 2013; Nystedt et al., 2013; Neale et al., 2014; Warren et al., 2015).

Phenylpropanoids involved in ultraviolet protection in conifers

Solar ultraviolet radiation potentially generates multiple negative effects in plants, such as damage to proteins, lipids, or DNA; growth reduction; or the inhibition of photosynthesis (Jordan, 2002; Frohnmeyer and Staiger, 2003; Barnes et al., 2016). Plants have developed various protection mechanisms to face these threats, including the biosynthesis of flavonoids and related phenylpropanoids that function as “UV sunscreens” and antioxidants (Caldwell, 1971; Yamaguchi et al., 2009; Agati et al., 2013). These compounds accumulate in vacuoles of the subepidermal cells of leaves protecting the inner cell layers from UV-B damage (Weissenböck et al., 1986; Landry et al., 1995; Kaffarnik et al., 2006). The presence of hydroxycinnamic acids and flavonol glycosides conjugated to the cell walls of multiple conifer species was also demonstrated (Strack et al., 1988).

Conifers are capable of growing in very different latitudes and heights, and these trees tolerate extremely different doses of UV radiation. The needles of these plants are rich in UV-B-absorbing soluble and cell wall-bounded phenolic compounds, such as lignans, coumarins, flavonoids, stilbenes, and hydroxycinnamic acids (Strack et al., 1988; Schnitzler et al., 1996; Fischbach et al., 1999). Notably, these conifer needles absorb short wavelength radiation, unlike herbaceous plants (DeLucia et al., 1992). In *P. sylvestris* and *P. abies*, the flavonol 3-*O*-glycoside acylated by with p-coumaric acid at position 3″ and p-coumaric or ferulic acid at position 6″ of the molecule were described as the main UV-B screening pigments (Schnitzler et al., 1996; Fischbach et al., 1999). These compounds are synthesized following UV-B irradiation and accumulate in epidermal cells, where they exhibit their effective protective function (Schnitzler et al., 1996; Kaffarnik et al., 2006). The flavonol glycosides kaempferol, isorhamnetin, and quercetin were also detected in epidermal tissue isolated from *P. sylvestris* needles. Remarkably, the biosynthesis of these glycosides directly depends on the activity of two critical enzymes in the synthesis of flavonoids and phenylpropanoids, chalcone synthase and PAL (Schnitzler et al., 1996), which demonstrate their *de novo* synthesis.

Identification of the exact phenylpropanoids that are involved in each developmental stage or physiological scenario in conifers is highly important for the future biochemical engineering of these plants toward the generation of new and enhanced forest-derived products.

Transcriptional regulation of gene expression

Regulation of the biosynthesis and utilization of Phe is a complex process that involves the coordinated expression of genes encoding enzymes located in different subcellular compartments and cellular types. An understanding of the transcription regulatory network associated with phenylpropanoid and lignin biosynthesis in conifers is crucial for future applications for tree improvement and sustainable forest management.

Transcriptome analyses indicated that several genes coding for TFs are preferentially expressed during wood formation in plants (Demura and Fukuda, 2007; Du and Groover, 2010). The best characterized are a set of transcription factors in the MYB and NAC families, which control lignin biosynthesis during wood formation.

The subfamily of R2R3-MYB factors is one of the largest transcription factor families in plants, with an estimated number of greater than 100 members in each species (Martin and Paz-Ares, 1997). These TFs bind AC elements in the promoter regions of phenylpropanoid and lignin biosynthesis genes to activate transcription. In trees, EgMYB2 from *Eucalyptus grandis* and PtMYB1 and PtMYB4 from *P. taeda* bind AC elements in the promoter of their target genes, which are expressed in developing wood and cause secondary wall thickening. Constitutive overexpression of these MYB transcription factors in tobacco (EgMYB2 and PtMYB4), *Arabidopsis* (PtMYB4) or white spruce (PtMYB1 and PtMYB8) increased secondary-wall thickening or led to ectopic lignin deposition (Patzlaff et al., 2003a,b; Goicoechea et al., 2005; Bomal et al., 2008).

Phenylalanine metabolism is finely regulated in conifers, primarily through transcriptional control. High transcript levels for three key genes *GS1b, PAL*, and *PAT* were detected in the stem and roots of young trees, and the compression wood of adult maritime pine trees (Craven-Bartle et al., 2013). These genes contain AC elements in their promoter regions, and the consensus sequence CCAACCAC/A functions as a *cis*-regulatory element involved in the transcriptional activation that is mediated by the PpMYB8 transcription factor (Craven-Bartle et al., 2013). Phe is the precursor for phenylpropanoids, and the regulation of Phe biosynthesis and metabolic utilization should occur in a coordinate manner, possibly in the same cell types. The overexpression of *PtMYB8* in white spruce clearly revealed the involvement of this transcription factor in secondary cell wall biogenesis and in lignin deposition during compression wood formation (Bomal et al., 2008). The *Myb8* gene exhibited strong similarity with *AtMYB61*, which is expressed in xylem tissues of *Arabidopsis*, and it plays an important role in the regulation of lignification (Newman et al., 2004). The co-expression of *PAT, PAL, GS1b*, and *MYB8* transcripts in the same cellular types further supports a coordinated transcriptional regulation of the pathway. The accumulation of MYB8 transcripts in compression wood of white spruce and maritime pine (Bedon et al., 2007; Craven-Bartle et al., 2013) is consistent with the ectopic lignification that resulted from the constitutive overexpression of *AtMYB61* in *Arabidopsis* (Newman et al., 2004). Other transcription factors in other families, such as NtLIM1 and ACBF, also bind to AC elements. Antisense inhibition of *NtLIM1* expression in transgenic tobacco plants reduced the expression levels of several phenylpropanoid genes, including *PAL* and *CAD*, and the total lignin content of stems, which indicated that it was required for normal lignin biosynthesis (Kawaoka et al., 2000). The relationship between the ACBF protein and its possible

involvement in lignin biosynthesis has not been established (Seguin et al., 1997).

Other well-characterized TFs that are involved in the regulation of wood formation belong to the NAC family. These proteins are plant-specific proteins with a highly conserved N-terminal NAC domain, which has been implicated in nuclear localization, DNA binding, and homo- and/or heterodimer formation with other NAC domain proteins (Olsen et al., 2005). Similar to the MYB family, the NAC transcription factors comprise a large gene family with more than 100 members in *Arabidopsis thaliana* (Ooka et al., 2003), *Oryza sativa* (Ooka et al., 2003; Nuruzzaman et al., 2010), *Glycine max* (Le et al., 2011), *Populus trichocarpa* (Hu et al., 2010), and *E. grandis* (Hussey et al., 2015). At least 37 genes encoding NAC proteins were identified in the *P. pinaster* genome (Pascual et al., 2015). The NAC proteins participate in many developmental processes, including secondary cell wall formation (Yamaguchi and Demura, 2010). A group of poplar wood-associated NAC domain proteins (PtrWNDs; Zhong and Ye, 2007; Zhong et al., 2010b) and an *Eucalyptus* wood-associated NAC (EgWND1; Zhong et al., 2010a) are functional orthologs of the *Arabidopsis* secondary wall NACs including SND1, NST1/2 and VND6/7 (Zhong and Ye, 2007). The overexpression of these TFs in *Arabidopsis* increased the expression of secondary wall biosynthetic genes and caused concomitant ectopic deposition of lignin in secondary walls, which suggests that these TFs play a role in the control of secondary wall biosynthesis. The NAC protein subfamily, including VND, NST, SMB, and BRN of *Arabidopsis* is termed the VNS family (Ohtani et al., 2011; Xu et al., 2014). The number of VNS genes varies between plant species and does not appear to correlate with genome size or the presence of woody tissues (Zhu et al., 2012; Nakano et al., 2015). In woody angiosperms, *P.*

trichocarpa has 16 (Zhong et al., 2010b; Ohtani et al., 2011; Li et al., 2012) and *E. grandis* has six VNS genes (Myburg et al., 2014; Hussey et al., 2015). In conifers, *P. abies* has four VNS genes (Nystedt et al., 2013), *P. glauca* has two VNS genes (Duval et al., 2014), and *P. pinaster* has three VNS genes (Pascual et al., 2015). Recent studies have demonstrated that the xylem transcriptomes of vascular plants are highly conserved (Li et al., 2010) and that this VNS family is conserved across a wide range of plant species, including non-vascular land plant such as bryophytes. The moss *Physcomitrella patens* contains eight VNS genes (Zhu et al., 2012; Xu et al., 2014).

Previous work in *Arabidopsis* demonstrated the existence of a complex transcriptional network involved in secondary cell wall biosynthesis (Zhong and Ye, 2007; Zhong et al., 2007, 2008, 2010a). The MYB regulators, *MYB46* and *MYB83*, are downstream of the NACs (SND1, NST1, and VND6/7) in *Arabidopsis*. The transcription factors R2R3-MYB are key in this transcriptional network, and *MYB46* is a key regulator of the biosynthesis of all three major secondary cell wall components, including cellulose, hemicellulose, and lignin (Zhong et al., 2007). They are targets of SND1 and NST1 (McCarthy et al., 2009) and situated in the second step of the transcriptional cascade. Overexpression of MYB46 or its homolog MYB83 increased the expression of other transcription factors, including MYB20, MYB42, MYB43, and MYB85, which are directly involved in the synthesis of secondary cell wall components (Zhong et al., 2008). Other transcription factors, including NACs, MYBs, and KNATs, downstream of the NAC and MYB master regulators form a complex regulatory network in secondary cell wall biosynthesis. A similar transcriptional network was identified in poplar (Zhong et al., 2010b, 2011, 2013; Lin et al., 2013), a woody angiosperm that is phylogenetically close to *Arabidopsis*. These

different studies support the hypothesis that the transcriptional regulatory network governing secondary cell wall biosynthesis is largely conserved in plant species.

The transcriptional network in conifers that regulates the secondary cell wall synthesis has not been studied in detail. Bomal et al. (2008) revealed a possible relationship between MYB-like transcription factors in *P. glauca*, which is similar to the network proposed by Zhong and Ye (2007) and Zhong et al. (2007). Two NAC genes were recently identified in *P. glauca* that may act as upstream regulators of MYBs (Duval et al., 2014). However, the complexity of the regulatory gene networks controlling wood formation in *Populus* indicates that intensive research is necessary to fully define the transcriptional regulatory hierarchy in conifers.

Conclusion

This review updates recent research advances regarding the biosynthesis and the metabolic fate of Phe in conifers. Two alternative routes have been identified: the ancient phenylpyruvate pathway that is present in many microorganisms, and the arogenate pathway that possibly evolved later during plant evolution. Besides, an efficient nitrogen recycling mechanism is required to maintain sustained growth during xylem formation. Further research efforts are necessary to clearly determine how these two pathways are regulated during vegetative development and in response to environmental changes. In regard to Phe utilization, the genetic manipulation of genes encoding key enzymes in lignin biosynthesis leads to new knowledge of considerable interest for the biochemical engineering of conifers and the generation of improved and novel forest-derived products. Progress has been made in transcriptional regulation but further research is

required to elucidate the complex regulatory networks involved in the control of lignin production during wood formation. Furthermore, the occurrence of a variety of additional Phe-demanding pathways in conifers makes these plants attractive candidates for more thorough analyses aimed to clarify the metabolic utilization of this important amino acid.

Acknowledgment

Research work in the author's laboratory was supported by grants from the Spanish Ministerio de Economía y Competitividad (BIO2015-69285-R) and Junta de Andalucía (BIO-474).

References

Agati G, Brunetti C, Di Ferdinando M, Ferrini F, Pollastri S and Tattini M (2013) *Functional roles of flavonoids in photoprotection: new evidence, lessons from the past*. Plant Physiol Biochem, 72:35-45.

Bagal UR, Leebens-Mack JH, Lorenz WW and Dean JF (2012) *The phenylalanine ammonia lyase (PAL) gene family shows a gymnosperm-specific lineage*. BMC Genomics, 13:S1.

Barnes PW, Flint SD, Tobler MA and Ryel RJ (2016) *Diurnal adjustment in ultraviolet sunscreen protection is widespread among higher plants*. Oecologia, 181:55-63.

Beckman CH (2000) *Phenolic-storing cells: keys to programmed cell death and periderm formation in wilt disease resistance and in general defence responses in plants?* Physiol Mol Plant Pathol, 57:101-110.

Bedon F, Grima-Pettenati J and MacKay J (2007) *Conifer R2R3-MYB transcription factors: sequence analysis and gene expression in wood forming tissues of white spruce* (Picea glauca). BMC Plant Biol. 7:17.

Bedon F, Levasseur C, Grima-Pettenati J, Séguin A and MacKay J (2009) *Sequence analysis and functional characterization of the promoter of the* Picea glauca *cinnamyl alcohol dehydrogenase gene in transgenic white spruce plants*. Plant Cell Rep, 28: 787-800.

Birol I, Raymond A, Jackman SD, Pleasance S, Coope R, Taylor GA, et al. (2013) *Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data*. Bioinformatics, 29:1492-1497.

Boerjan W, Ralph J and Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol, 54:519-546.

Bomal C, Bedon F, Caron S, Mansfield SD, Levasseur C, Cooke JEK, et al. (2008) *Involvement of* Pinus taeda *MYB1 and Myb8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis*. *J Exp Bot,* 59:3925-3939.

Bonner C and Jensen R (1987a) *Prephenate aminotransferase*. Methods Enzymol, 142: 479-487.

Bonner C and Jensen R (1987b) *Arogenate dehydrogenase*. Methods Enzymol, 142:488-494.

Brignolas F, Lacroix B, Lieutier F, Sauvard D, Drouet A, Claudot AC, et al. (1995) *Induced responses in phenolic metabolism in two Norway spruce clones after wounding and inoculation with* Ophiostoma polonicum, *a bark beetle-associated fungus*. *Plant Physiol,* 109:821-827.

Caldwell MM (1971) *Solar UV irradiation and the growth and development of higher plants*. In: Photophysiology, Vol. 6, ed. A. C. Giese (New York, NY: Academic Press), pp. 131-177.

Cañas RA, Canales J, Muñoz-Hernández C, Granados JM, Ávila C, García-Martín ML, et al. (2015) *Understanding developmental and adaptive cues in pine through metabolite profiling and co-expression network analysis*. J Exp Bot, 66:3113-3127.

Cánovas FM, Avila C, Cantón FR, Cañas RA and de la Torre F (2007) *Ammonium assimilation and amino acid metabolism in conifers*. J Exp Bot, 58:2307-2318.

Cass CL, Peraldi A, Dowd PF, Mottiar Y, Santoro N, Karlen SD, et al. (2015) *Effects of PHENYLALANINE AMMONIA LYASE (PAL) knockdown on cell wall composition, biomass digestibility, and biotic and abiotic stress responses in* Brachypodium. J Exp Bot, 66:4317-4335.

Cheynier V, Comte G, Davies KM, Lattanzio V and Martens S (2013) *Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology*. *Plant Physiol Biochem*, 72:1-20.

Claros MG, Aguilar M and Cánovas FM (2010) *Evidence for an operative glutamine translocator in chloroplasts from maritime pine* (Pinus pinaster Ait.) *cotyledons*. Plant Biol, 12:717-723.

Craven-Bartle B, Pascual MB, Cánovas FM and Ávila C (2013) *A Myb transcription factor regulates genes of the phenylalaline pathway in maritime pine*. Plant J, 74:755-766.

Danielsson M, Lundén K, Elfstrand M, Hu J, Zhao T, Arnerup J, et al. (2011) *Chemical and transcripcional responses ofr Norway spruce genotypes with different susceptibility to* Heterobasidion spp. *infection.* BMC Plant Biol, 11:154.

de la Torre F, Cañas RA, Pascual MB, Ávila C and Cánovas FM (2014) *Plastidic aspartate aminotransferases and the biosynthesis of essential amino acids*. J Exp Bot, 65:5527-5534.

DeLucia EH, Day TA and Vogelman TC (1992) *Ultraviolet-B and visible light penetration into needles of two species of subalpine conifers during foliar development*. Plant Cell Environ, 15:921-929.

Demura T and Fukuda H (2007) *Transcriptional regulation in wood formation*. Trends Plant Sci, 12:64-70.

Du J and Groover A (2010) *Transcriptional regulation of secondary growth and wood formation*. J Integr Plant Biol, 52:17-27.

Duval I, Lachance D, Giguère I, Bomal C, Morency MJ, Pelletier G, et al. (2014) *Large-scale screening of transcription factor–promoter interactions in spruce reveals a transcriptional network involved in vascular development*. J Exp Bot, 65:2319-2333.

El-Azaz J, de la Torre FN, Avila C and Cánovas FM (2016) *Identification of a small protein domain present in all plant lineages that confers high prephenate dehydratase activity*. Plant J, doi: 10.1111/tpj.13195 [Epub ahead of print].

Farjon A (2010) *A Handbook of the World's Conifers*. Leiden-Boston: Brill.

Fischbach RJ, Kossmann B, Panten H, Steinbrecher R, Heller W, Seidlitz HK, et al. (1999) *Seasonal accumulation of ultraviolet-B screening pigments in needles of Norway spruce*. Plant Cell Environ, 22:27-37.

Fischer R and Jensen R (1987a) *Prephenate dehydratase (monofunctional)*. Methods Enzymol, 142:507-512.

Fischer R and Jensen R (1987b) *Prephenate dehydrogenase (monofunctional)*. Methods Enzymol, 142:503-507.

Franceschi VR, Krokene P, Christiansen E and Krekling T (2005) *Anatomical and chemical defenses of conifer bark against bark beetles and other pests*. New Phytol, 167:353-375.

Franceschi VR, Krokene P, Krekling T and Christiansen E (2000) *Phloem parenchyma cells are involved in local and distant defense responses to fungal inoculation or bark-beetle attack in* Norway spruce *(*Pinaceae*)*. Am J Bot, 87:314-326.

Frohnmeyer H and Staiger D (2003) *Ultraviolet-B radiation-mediated responses in plants. Balancing damage and protection*. Plant Physiol, 133:1420-1428.

Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, et al. (2005) *EgMYB2, a new transcriptional activator from* Eucalyptus xylem*, regulates secondary cell wall formation and lignin biosynthesis*. Plant J, 43:553-567.

Hammerbacher A, Paetz C, Wright LP, Fischer TC, Bohlmann J, Davis AJ, et al. (2014) *Flavan-3-ols in Norway spruce: biosynthesis, accumulation, and function in response to attack by the bark beetle-associated fungus* Ceratocystis polonica. Plant Physiol, 164:2107-2122.

Heldt HW and Piechulla B (2011) *Plant Biochemistry*. San Diego, CA: Elsevier Academic Press.

Hu R, Qi G, Kong Y, Kong D, Gao Q and Zhou G (2010) *Comprehensive analysis of NAC domain transcription factor gene family in* Populus trichocarpa. BMC Plant Biol, 10:145.

Hudgins JW, Christiansen E and Franceschi VR (2003) *Methyl jasmonate induces changes mimicking anatomical defenses in diverse members of the* Pinaceae. Tree Physiol, 23:361-371.

Hudgins JW and Franceschi VR (2004) *Methyl jasmonate-induced ethylene production is responsible for conifer phloem defense responses and reprogramming of stem cambial zone for traumatic resin duct formation*. Plant Physiol, 135:2134–2149.

Hussey SG, Saïdi MN, Hefer CA, Myburg AA and Grima-Pettenati J (2015) *Structural, evolutionary and functional analysis of the NAC domain protein family in* Eucalyptus. New Phytol, 206:1337-1350.

Jordan BR (2002) *Molecular response of plant cells to UV-B stress*. Funct Plant Biol, 29:909-916.

Kaffarnik F, Seidlitz HK, Obermaier J, Sandermann H Jr and Heller W (2006) *Environmental and developmental effects on the biosynthesis of UV-B screening pigments in Scots pine (*Pinus sylvestris L.*) needles*. Plant Cell Environ, 29:1484-1491.

Kawaoka A, Kaothien P, Yoshida, K., Endo S, Yamada K and Ebinuma H (2000) *Functional analysis of tobacco LIM protein Ntlim1 involved in lignin biosynthesis*. Plant J, 22:289-301.

Koutaniemi S, Warinowski T, Kärkönen A, Alatalo E, Fossdal CG, Saranpää P, et al. (2007) *Expression profiling of the lignin biosynthetic pathway in Norway spruce using EST sequencing and real-time RT-PCR*. Plant Mol Biol, 65:311-328.

Krekling T, Franceschi VR, Berryman AA and Christiansen E (2000) *The structure and development of polyphenolic parenchyma cells in Norway spruce* (Picea abies) *bark*. Flora, 195:354-369.

Kusumoto D and Suzuki K (2003) *Spatial distribution and time-course of polyphenol accumulation as a defense response induced by wounding in the phloem of* Chamaecyparis obtusa. New Phytol, 159:167-173.

Landry LG, Chapple CC and Last RL (1995) Arabidopsis *mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage*. Plant Physiol, 109:1159-1166.

Le DT, Nishiyama R, Watanabe Y, Mochida K, Yamaguchi-Shinozaki K, Shinozaki K, et al. (2011) *Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress*. DNA Res, 18:263-276.

Li Q, Lin YC, Sun YH, Song J, Chen H, Zhang XH, et al. (2012) *Splice variant of the SND1 transcription factor is a dominant negative of SND1 members and their regulation in* Populus trichocarpa. Proc Natl Acad Sci USA, 109:14699–14704.

Li X, Wu HX and Southerton SG (2010) *Comparative genomics reveals conservative evolution of the xylem transcriptome in vascular plants*. BMC Evol Biol, 10:190.

Lieutier F, Sauvard D, Brignolas F, Picron V, Yart A, Bastien C, et al. (1996) *Changes in phenolic metabolites of Scots pine phloem induced by* Ophiostoma brunneo-ciliatum, *a bark beetle-associated fungus*. Eur J For Pathol 26:145-158.

Lin YC, Li W, Sun YH, Kumari S, Wei H, Li Q, et al. (2013) *SND1 transcription factor–directed quantitative functional hierarchical genetic regulatory network in wood formation in Populus trichocarpa*. Plant Cell, 25:4324-4341.

MacKay JJ, O'Malley DM, Presnell T, Booker FL, Campbell MM, Whetten RW, et al. (1997) *Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase*. *Proc Natl Acad Sci USA*, 94:8255-8260.

Maeda H and Dudareva N (2012) *The shikimate pathway and aromatic amino acid biosynthesis in plants*. Annu Rev Plant Biol, 63:73-105.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, Cooper BR, et al. (2010) *RNAi suppression of Arogenate Dehydratase1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22:832-849.

Martin C and Paz-Ares J (1997) *MYB transcription factors in plants*. Trends Plant Sci, 13: 67-73.

Mast S, Peng L, Jordan TW, Flint H, Phillips L, Donaldson L, et al. (2010) *Proteomic analysis of membrane preparations from developing* Pinus radiata *compression wood*. Tree Physiol, 30:1456-1468.

McCarthy RL, Zhong R and Ye ZH (2009) *MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in* Arabidopsis. Plant Cell Physiol, 50:1950-1964.

Möller R, McDonald AG, Walter C and Harris PJ (2003) *Cell differentiation, secondary cell-wall formation and transformation of callus tissue of* Pinus radiata *D. Don*. Planta, 217:736-747.

Möller R, Steward D, Phillips L, Heather F and Wagner A (2005). *Gene silencing of cinnamyl alcohol dehydrogenase in* Pinus radiata *callus cultures*. Plant Physiol Biochem, 43: 1061-1066.

Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, et al. (2014) *The genome of* Eucalyptus grandis. Nature, 510:356-362.

Nakano Y, Yamaguchi M, Endo H, Rejab NA and Ohtani M (2015) *NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants*. Front Plant Sci, 6:288.

Neale DB, Wegrzyn JL, Stevens KA, Zimin AV, Puiu D, Crepeau MW, et al. (2014) *Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies*. Genome Biol, 15:R59.

Newman LJ, Perazza DE, Juda L and Campbell MM (2004). *Involvement of the R2R3-MYB, At MYB61, in the ectopic lignication and dark photomorphogenic components of the det3 mutant phenothype*. Plant J, 37:239-250.

Nuruzzaman M, Manimekalai R, Sharoni AM, Satoh K, Kondoh H, Ooka H, et al. (2010) *Genome-wide analysis of NAC transcription factor family in rice*. Gene, 465:30-44.

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, et al. (2013) *The Norway spruce genome sequence and conifer genome evolution*. Nature, 497:579-584.

Ohtani M, Nishikubo N, Xu B, Yamaguchi M, Mitsuda N, Goue N, et al. (2011) *A NAC domain protein family contributing to the regulation of wood formation in poplar*. Plant J, 67:499-512.

Oliva J, Rommel S, Fossdal CG, Hietala AM, Nemesio-Gorriz M, Solheim H, et al. (2015) *Transcriptional responses of Norway spruce* (Picea abies) *inner sapwood against* Heterobasidion parviporum. Tree Physiol, 35:1007-1015.

Olsen AN, Ernst HA, Leggio LL and Skriver K (2005) *NAC transcription factors: structurally distinct, functionally diverse*. Trends Plant Sci, 10:79-87.

Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, Murakami K, et al. (2003) *Comprehensive analysis of NAC family genes in* Oryza sativa *and* Arabidopsis thaliana. DNA Res, 10: 239-247.

Pascual MB, Cánovas FM and Ávila C (2015) *The NAC transcription factor family in maritime pine (*Pinus pinaster*): molecular regulation of two genes involved in stress responses*. BMC Plant Biol, 15:254.

Patzlaff A, McInnis S, Courtenay A, Surman C, Newman LJ, Smith C, et al. (2003a) *Characterisation of a pine MYB that regulates lignification*. Plant J, 36:743-754.

Patzlaff A, Newman LJ, Dubos C, Whetten RW, Smith C, McInnis S, et al. (2003b) *Characterization of PtMYB1, an R2R3-MYB from pine xylem*. Plant Mol Biol, 53:597-608.

Pribat A, Noiriel A, Morse AM, Davis JM, Fouquet R, Loizeau K, et al. (2010) *Nonflowering plants possess a unique folate-dependent phenylalanine hydroxylase that is localized in chloroplasts*. Plant Cell, 22:3410-3422.

Raherison ES, Giguère I, Caron S, Lamara M and MacKay JJ (2015) *Modular organization of the white spruce (*Picea glauca*) transcriptome reveals functional organization and evolutionary signatures*. New Phytol, 207:172-187.

Ralph J (2010) *Hydroxycinnamates in lignification*. Phytochemistry, 9:65-83.

Saitou N and Nei M (1987) *The neighbor-joining method: a new method for reconstructing phylogenetic trees*. Mol Biol Evol, 4:406-425.

Schaeffer SM and Nakata PA (2015) *CRISPR/Cas9-mediated genome editing and gene replacement in plants: transitioning from lab to field*. Plant Sci, 240:130-142.

Schnitzler JP, Jungblut TP, Heller W, Köfferlein M, Hutzler P, Heinzmann U, et al. (1996) *Tissue localization of u.v.-B-screening pigments and of chalcone synthase mRNA in needles of Scots pine seedlings*. New Phytol, 132:247-258.

Seguin A, Laible G, Leyva A, Dixon RA and Lamb CJ (1997). *Characterization of a gene encoding a DNA-binding protein that interacts in vitro with vascular specific cis elements of the phenylalanine ammonia-lyase promoter.* Plant Mol Biol 35:281-291.

Song Z, Chen W, Du X, Zhang H, Lin L and Xu H (2011) *Chemical constituents of* Picea neoveitchii. Phytochemistry, 72:490-494.

Strack D, Heilemann J, Mömken M and Wray V (1988) *Cell wall-conjugated phenolics from coniferous leaves*. Phytochemistry, 27: 3517-3521.

Tamura K, Stecher G, Peterson D, Filipski A and Kumar S (2013) *MEGA6: molecular evolutionary genetics analysis version 6.0*. Mol Biol Evol 30:2725-2729.

Thompson JD, Higgins DG and Gibson TJ (1994) *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice*. Nucleic Acid Res, 22:4673-4680.

Trontin JF, Teyssier C, Avila C, Debille S, Le Metté C, Lesage-Descauses MC, et al. (2014) *Molecular phenotyping of Maritime pine somatic plants transformed with an RNAi construct targeting cinnamyl alcohol dehydrogenase (CAD)*. In: *Proceedings of the Third International Conference of the IUFRO unit 2.09.02: Somatic Embryogenesis and Other Vegetative Propagation Technologies, Woody Plant Production Integrating Genetic and Vegetative Propagation Technologies*, Vitoria-Gasteiz, 140.

Tsai CJ and Xue LJ (2015). *CRISPRing into the Woods*. GM Crops Food, 6:206-215.

Tzin V and Galili G (2010) *New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants*. Mol Plant, 3:956-972.

Wadenbäck J, von Arnold S, Egertsdotter U, Walter MH, Grima-Pettenati J, Goffner D, et al. (2008) *Lignin biosynthesis in transgenic Norway spruce plants harbouring an antisense construct for cinnamoyl CoA reductase (CCR)*. Transgenic Res, 17:379-392.

Wagner A, Donaldson L, Kim H, Phillips L, Flint H, Steward D, et al. (2009) *Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm* Pinus radiata. Plant Physiol, 149:370-383.

Wagner A, Donaldson L and Ralph J (2012) *Lignification and lignin manipulations in conifers*. In: *Lignins: Biosynthesis, Biodegradation and Bioengineering. Collection: Advances in Botanical Research*, eds L. Jouanin and C. Lapierre (San Diego, CA: Elsevier Academic Press), pp. 37-76.

Wagner A, Phillips L, Narayan RD, Moody JM and Geddes B (2005) *Gene silencing studies in the gymnosperm species* Pinus radiata. Plant Cell Rep, 24:95-102.

Wagner A, Ralph J, Akiyama T, Flint H, Phillips L, Torr K, et al. (2007) *Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in* Pinus radiata. Proc Natl Acad Sci USA, 104:11856-11861.

Wagner A, Tobimatsu Y, Goeminne G, Phillips L, Flint H, Steward D, et al. (2013) *Suppression of CCR impacts metabolite profile and cell wall composition in* Pinus radiata *tracheary elements*. Plant Mol Biol, 81:105-117.

Wagner A, Tobimatsu Y, Phillips L, Flint H, Geddes B, Lu F, et al. (2015) *Syringyl lignin production in conifers: proof of concept in a Pine tracheary element system*. Proc Natl Acad Sci USA, 112:6218-6223.

Wagner A, Tobimatsu Y, Phillips L, Flint H, Torr K, Donaldson L, et al. (2011) *CCoAOMT suppression modifies lignin composition in* Pinus radiata. Plant J, 67:119-129.

Warren RL, Keeling CI, Yuen MMS, Raymond A, Taylor GA, Vandervalk BP, et al. (2015) *Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism*. Plant J, 3:189-212.

Wei XX and Wang XQ (2004) *Evolution of 4-coumarate:coenzyme A ligase (4CL) gene and divergence of* Larix (Pinaceae). Mol Phylogenet Evol, 31:542-553.

Weissenböck G, Hedrich R and Sachs G (1986) *Secondary phenolic products in isolated guard cell, epidermal cell and mesophyll cell protoplasts – distribution and determination*. Protoplasma, 134:141-148.

Xu B, Ohtani M, Yamaguchi M, Toyooka K, Wakazaki M, Sato M, et al. (2014) *Contribution of NAC transcription factors of plant adaptation to land*. Science, 343:1505-1508.

Yamaguchi LF, Kato MJ and Di Mascio P (2009) *Biflavonoids from* Araucaria angustifolia *protect against DNA UV-induced damage*. Phytochemistry, 70:615-620.

Yamaguchi M and Demura T (2010) *Transcriptional regulation of secondary wall formation controlled by NAC domain proteins*. Plant Biotechnol, 27:237-242.

Zhong R, Lee C and Ye ZH (2010a) *Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis*. Trends Plant Sci, 15:625-631.

Zhong R, Lee C and Ye ZH (2010b) *Functional characterization of poplar wood-associated NAC domain transcription factors*. Plant Physiol, 152:1044-1055.

Zhong R, Lee C, Zhou J, McCarthy RL and Ye ZH (2008) *A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in* Arabidopsis. Plant Cell, 20:2763-2782.

Zhong R, McCarthy RL, Haghighat M and Ye ZH (2013). *The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation*. PLoS ONE, 8:e69219.

Zhong R, McCarthy RL, Lee C and Ye ZH (2011) *Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar*. Plant Physiol, 157:1452-1468.

Zhong R, Richardson EA and Ye ZH (2007) *The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in* Arabidopsis. Plant Cell, 19: 2776-2792.

Zhong R and Ye ZH (2007) *Regulation of cell wall biosynthesis*. Curr Opin Plant Biol, 10:564-572.

Zhu T, Nevo E, Sun D and Peng J (2012) *Phylogenetic analyses unravel the evolutionary of NAC proteins in plants*. Evolution, 66:1833-1848.

Zimin A, Stevens KA, Crepeau MW, Holtz-Morris A, Koriabine M, Marçais G, et al. (2014) *Sequencing and assembly of the 22-gb loblolly pine genome*. Genetics, 196:875-890.

Zuckerkandl E and Pauling L (1965) *Evolutionary divergence and convergence in proteins*. In: *Evolving Genes and Proteins*, eds V. Bryson and H. J. Vogel (New York, NY: Academic Press), pp. 97-166.

Paper 6

A member of the arogenate deshydratase gene family *(ADT2)* is essential for female gametophyte development in *Arabidopsis*

Jorge El-Azaz et al.

Research paper draft

Summary

Phenylalanine (Phe) biosynthesis in plants is an exceptionally important process due the doubly essential function of Phe as building block for proteins and as the main precursor of phenylpropanoid biosynthesis. The prephenate pathway, located in plant plastids, connects chorismate, the final product of the shikimate pathway, with the biosynthesis of the aromatic amino acids Phe and Tyr. Two alternative routes of Phe biosynthesis have been reported: one depending of arogenate, and the other on phenylpyruvate. While the arogenate pathway is considered as the main route for the biosynthesis of Phe in plants, the role of the phenylpyruvate pathway still remains unclear. Here, we report a female semi-sterility phenotype caused by the *adt2* T-DNA insertional mutation, demonstrating this gene as essential for an adequate female gametophyte development. This result makes a clear distinction between ADT2 and the 5 remaining ADTs from *Arabidopsis*, which had been reported to have a partially overlapping function. Moreover, we have found that *PHA2* gene from *Saccharomyces cerevisiae*, which encodes a PDT enzyme without detectable ADT activity, can rescue the observed *adt2* mutant phenotype when is expressed under control of the 5'-upstream region of *ADT2*. Furthermore, the heterologous expression of *ADT3*, that encodes a

monofunctional ADT, is also able to complement the *adt2* semi-sterility phenotype, suggesting that in this organ Phe can be synthesized from both arogenate and phenylpyruvate simultaneously. These findings highlight the importance of Phe metabolism for an adequate embryo development, either for protein and/or phenolic biosynthesis, providing new insights into the biosynthesis of Phe from different pathways.

INTRODUCTION

Phenylalanine (Phe) biosynthesis in plants is an exceptionally important process due the doubly essential function of Phe as building block for proteins and as the main precursor of phenylpropanoid biosynthesis. Phenylpropanoids comprise a wide range of aromatic compounds that play a key role in plant growth, development and stress responses. Land colonization by the first terrestrial plants, one of the most important evolutionary steps in the history of life, would not have been possible without the emergence of specialized metabolic pathways that are required for the biosynthesis of these secondary metabolites.

The prephenate pathway, located in plant plastids, connects chorismate, the final product of the shikimate pathway, with the biosynthesis of the aromatic amino acids Phe and Tyr (Maeda and Dudareva, 2012) (**Figure 1**). In the prephenate pathway, two alternative routes for the biosynthesis of Phe have been reported: i) in the arogenate pathway, prephenate is transaminated by prephenate-aminotransferase (PAT) to generate arogenate, which is decarboxylated and dehydrated by arogenate dehydratase (ADT) to give Phe (Bonner and Jensen, 1987a); ii) in the phenylpyruvate pathway, present in most microorganisms, the decarboxylation and dehydratation of prephenate results in the generation

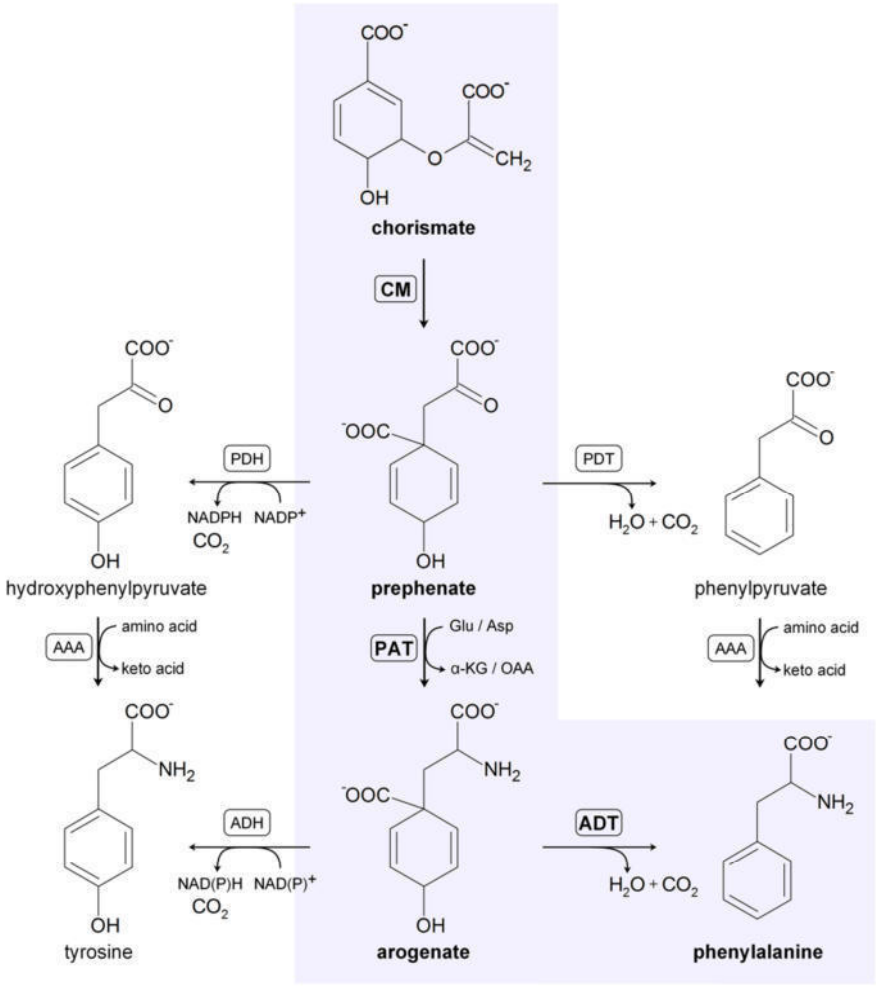


Figure 1. The prephenate pathway in plants. The arogenate branch, marked in pale blue, is considered to be the major source of phenylalanine in plants. Enzyme abbreviations: CM, chorismate mutase; PDH, prephenate dehydrogenase; PDT, prephenate dehydratase; PAT, prephenate amino transferase; AAA, aromatic amino acid amino transferase; ADH, arogenate dehydrogenase; ADT, arogenate dehydratase.

of phenylpyruvate in a reaction catalyzed by prephenate dehydratase (PDT). Subsequently, phenylpyruvate is transaminated to Phe through a phenylpyruvate-aminotransferase (Fischer and Jensen, 1987a). Prephenate and arogenate are also precursors for the biosynthesis of Tyr through the activity of prephenate dehydrogenase and arogenate dehydrogenase, respectively (**Figure** 1; Fischer and Jensen, 1987b; Bonner and Jensen, 1987b).

The arogenate pathway has been proposed as the main route of Phe biosynthesis in plants (Maeda et al*.,* 2010; 2011). Nevertheless, under certain conditions, the occurrence of PDT activity in plants has been reported in the last years (Warpeha et al., 2006; Yoo et al*.,* 2013). The molecular and kinetic characterization of the ADT family from *Arabidopsis thaliana* (Cho et al., 2007) and, very recently, *Pinus pinaster* (El-Azaz et al., 2016) has also shown that some ADT isoenzymes can also work as PDT enzymes *in vitro*, although under lower catalytic efficiencies compared to arogenate. In addition, these ADT/PDT bifunctional proteins can rescue the Phe auxotrophic phenotype caused by the deletion of the endogenous PDT activity in the yeast mutant *pha2* (Bross *et al*., 2011; El-Azaz et al., 2016). Further evidences about the existence of a functional PDT pathway in plants are given by the accumulation, in response to exogenous shikimate supply, of phenylpyruvate and the restoration of the Phe levels in *Petunia hybrida* petals silenced for ADT activity (Maeda et al., 2010). In the same line, Yoo et al. (2013) have described a cytosolic phenylpyruvate-aminotransferase in *P. hybrida* that catalyzes the efficient interconversion of phenylpyruvate and Tyr to Phe and 4-hydroxyphenylpyruvate, respectively. Through feeding experiments with labelled phenylpyruvate, as well as silencing and overexpression analysis, these authors proposed the existence of a functional phenylpyruvate pathway that has an enhanced

flux when blocking the arogenate pathway. Consistently, the virus-induced gene silencing of PAT in *Nicotiana benthamiana* resulted in the transcriptional activation of two genes encoding for ADT/PDT enzymes (de la Torre et al., 2014). Putting it all together, there are strong evidences suggesting the existence of a functional phenylpyruvate pathway in plants. However, none of these studies have reported evidences about what is the specific role of the phenylpyruvate pathway in plant development.

Previous reports (Cho et al., 2007; Rippert et al., 2009) suggested that all six *Arabidopsis ADT* genes are expressed more or less ubiquitously in multiple tissues with differential intensity. These previous reports have shown the possibility to generate heterozygous knock-out mutants for all the *Arabidopsis ADTs,* with the sole exception of *ADT2*, and their different combinations (Corea et al., 2012), evidencing a partially redundant role in plant growth and development. Here, we report female semi-sterility phenotype caused by the *ADT2* T-DNA insertional mutation, demonstrating this gene as essential for an adequate female gametophyte development, and thus to complete the plant life cycle. This result makes a clear distinction between ADT2 and the 5 remaining ADTs from *Arabidopsis*, which had been reported to have a partially overlapping function. Moreover, we have found that *PHA2* gene from *Saccharomyces cerevisiae*, which encodes a monofunctional PDT enzyme without detectable ADT activity (Bross et al., 2011), can rescue the observed *adt2* mutant phenotype. These findings highlight the importance of Phe metabolism for an adequate embryo development, either for protein and/or phenolic biosynthesis. In addition, this study provides new evidences that support the existence of a functional phenylpyruvate pathway in plants.

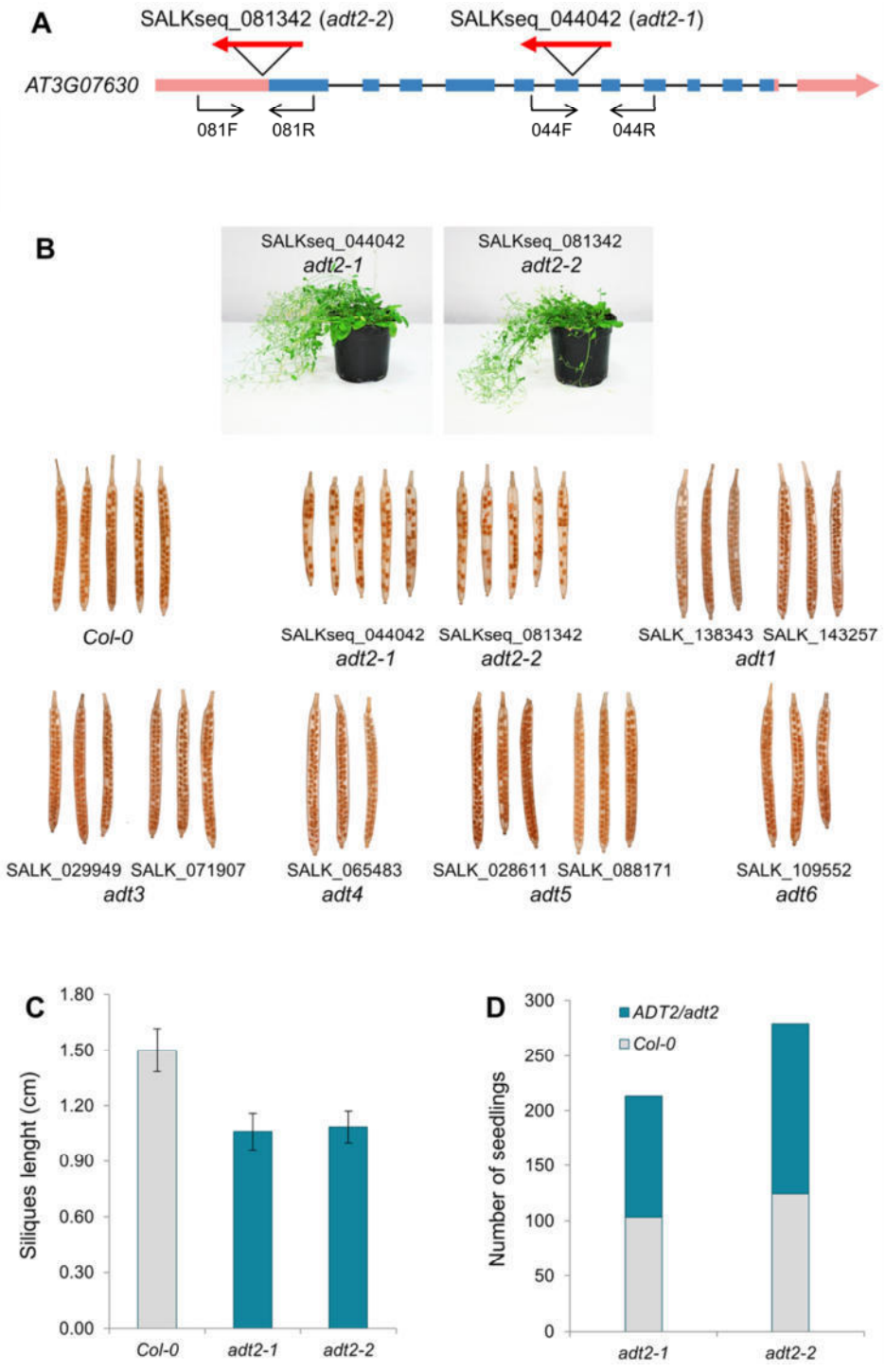
RESULTS

The *adt2* insertional mutants are female semi-sterile

Previous investigations, carried out with T-DNA insertional lines of *A. thaliana*, were unable to obtain an homozygous T-DNA line of *ADT2* (AT3G07630), with its mRNA expression completely abolished (Corea et al., 2012). Homozygous knock-out lines were confirmed for members of the *ADT* gene family, from *ADT1* to *ADT6*, with the sole exception of *ADT2*. This previous result suggested *ADT2* as an essential gene in *Arabidopsis* life cycle.

In order to further investigate *ADT2* as an essential gene, we analyzed the phenotypes associated with a T-DNA insertional mutation for the six ADT isoenzymes from *Arabidopsis* (**Supplemental table 1**). T-DNA mutants were requested from the Arabidopsis Biological Resource Center (ABRC).

Figure 2. (next page) Molecular and phenotypic analysis of *adt2* insertional mutants. **(A)** Scheme indicating the position of the T-DNA insertions in the *ADT2* (AT3G07630) mutant lines requested, named SALKseq_044042 (*adt2-1*) and SALKseq_081342 (*adt2-2*). Upon arrival, both lines were confirmed by PCR as heterozygous. **(B)** Above, prostrate phenotype of the *adt2-1* and *adt2-2* mutant plants. Below, mature siliques from the different ADT gene family T-DNA mutants, showing the existence of an abnormal phenotype in the siliques from the *adt2* T-DNA lines analyzed. Pictures were taken with the same magnification. **(C)** Average siliques length (cm) measured for Columbia wild-type plants (*Col-0*) and both *adt-2* T-DNA heterozygous lines. **(D)** Genotypes rate determined in the *adt-2* T-DNA progeny. 2-weeks old plants were screened by PCR for the T-DNA insertion, resulting in an adjusted 1:1 segregation rate, with no homozygous T-DNA plants detected.



Two independent SALK T-DNA heterozygous lines were ordered for *ADT2*: SALKseq_044042 (*adt2-1*, insert position at exon number 6) and SALKseq_081342 (*adt2-2*, insert position at -34 pb from the ORF starting codon) (**Figure 2A**). Insertion sites for each ADT knock-out line were located in the SIGnAL database (http://signal.salk.edu/cgi-bin/tdnaexpress) and T-DNA homozygous lines were confirmed by PCR (Supplemental figure 1).

Heterozygous *adt2-1* and *adt2-2* plants had a prostrate phenotype which has been previously observed in other ADT mutants (Corea et al., 2012; **Figure 2B**). More interestingly, heterozygous *adt2* plants produce shorter siliques (**Figures 2B, 2C**) with a clearly reduced seed set when compared to WT plants, indicating ovule abortion in about a 50% rate. This phenotype was not observed for any of the other *adt1* and *adt3-6* mutants (**Figure 2B**). As previously reported by Corea and collaborators (2012), we were unable to obtain a homozygous *adt2* KO line. A total number of 213 two-week-old seedlings for the *adt2-1* line and 279 for the *adt2-2* line, coming from independent parental heterozygous plants, were analyzed for their respective T-DNA insertions in the *ADT2* gene. Segregation rates of the T-DNA insertion for the heterozygous *adt2-1* and *adt2-2* progenies showed a distorted segregation ratio of 1:1, characteristic of a female semi-sterility phenotype (**Figure 2D**) (Page and Grossniklaus, 2002). Hence, from 213 seedlings analyzed corresponding to *adt2-1*, we found that 103 had not a detectable T-DNA insertion in *ADT2*, whereas 110 were heterozygous *adt2-1* mutants. Similarly, of 279 seddlings from the *adt2-2* mutant line, we found a total number of 124 wild-type plants and 155 heterozygous *adt2-2* plants. No homozygous *adt2-1* or *adt2-2* mutants were detected in this screening, strongly suggesting that the mutant alelles *adt2-1* and *adt2-2* can not be fully transmitted through the gamethophytes and/or caused seed lethality.

Similar results were repeatedly obtained after several generations of growth. Crosses between WT pistils and heterozygous T-DNA pollen gave heterozygous *adt2* descendant plants with the same described phenotype, while the reciprocal cross between heterozygous *adt2* pistils and WT pollen led to a fully WT progeny, confirming that the semi-sterility phenotype observed in heterozygous *adt2* mutant plants is attributable to the female reproductive organs.

Tissue specific expression of *ADT2*

An *in silico* analysis of the data available from the mature *Arabidopsis* female gametophyte transcriptome, which were previously published by Wuest et al. (2010), revealed a clear isoform-dependent expression pattern in the female gametophyte for the *ADT* gene family (Supplemental table 2). While *ADT1* and *ADT2* expression levels were detectable in the female gametophyte cells, *ADT3-6* mRNAs were not found in this tissue. Interestingly, the expression patterns of *ADT1* and *ADT2* in the female gametophyte differed; in particular, *ADT1* mRNA was detected in all female gametophyte cell types (synergid cells, egg cell and central cell) whereas *ADT2* was found to be expressed in the egg cell and central cell, but not in the synergids.

To further corroborate *ADT2* expression at the tissue level, the 5'-flanking regulatory region of *ADT2* was cloned into pGWB3 vector, upstream of the *GUS* reporter gene. The resulting gene construct, named P_{ADT2}*:GUS*, was used to transform *Arabidopsis* WT plants. The histochemical GUS localization of P_{ADT2}*:GUS* plants (**Figure 3**) indicates that P_{ADT2}*:GUS* is mainly expressed in nascent leaves (**Figures 3A, 3B, 3D**), during flower development (**Figures 3E-3I**) and in the formation of female gametophyte

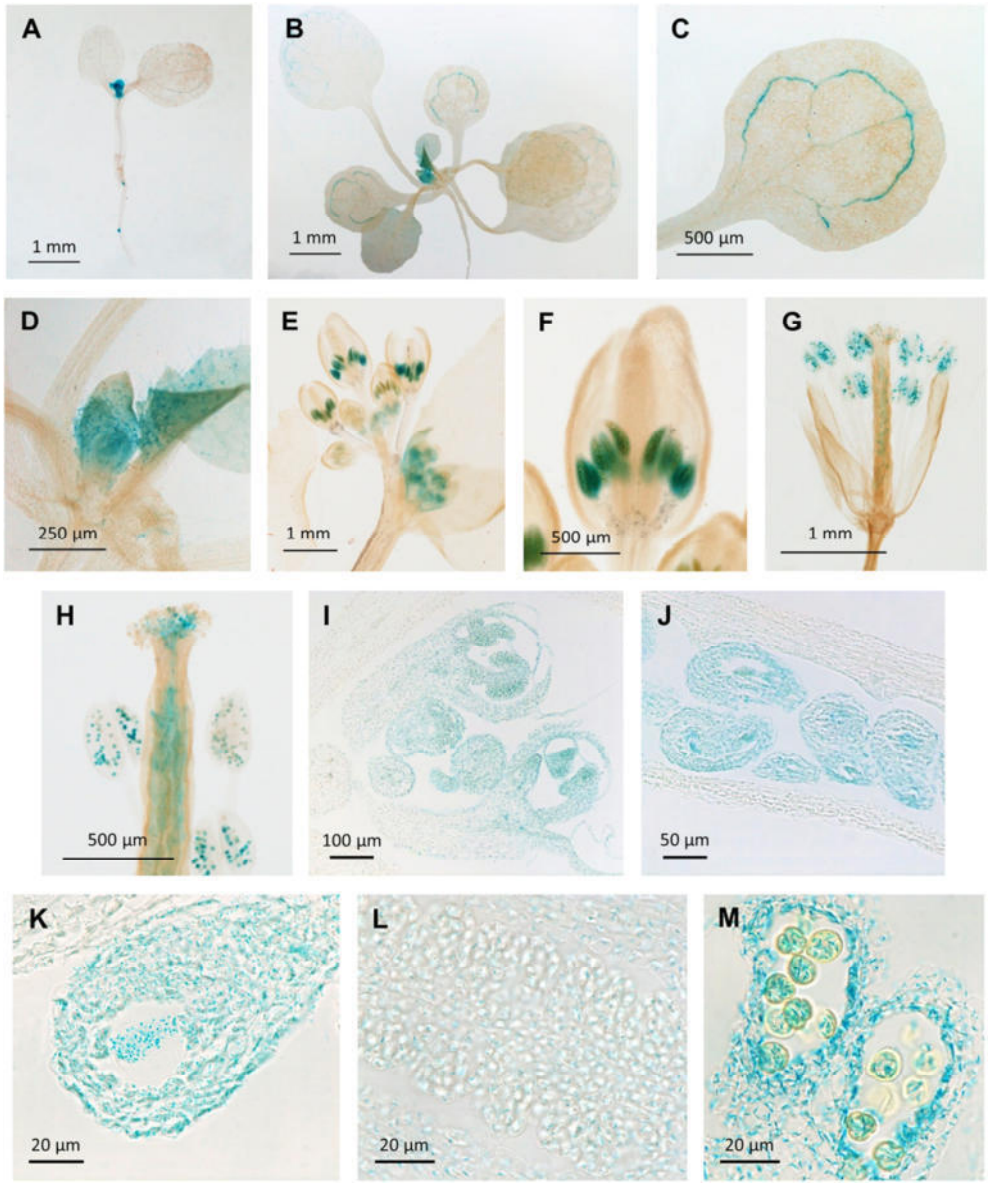


Figure 3. Histochemical localization of GUS in transgenic *Arabidopsis thaliana* plants contanaing the P_{ADT2}:*GUS* construct. **(A)** One-week-old seedling. **(B-D)** Two-week-old seedlings. **(E-F)** Emerging flowers from an adult plant. **(G)** Mature flower. **(H)** Detail from the pistil and polen anters from a mature flower. **(I)** Floral primordial. **(F)** Young siliqua with early stage seeds. **(G)** Early stage seed and embryo **(H)** Ovaries from a flower in stage 8-9. **(M)** Pollen anthers from a flowe in stage 8-9.

(**Figure 3J, 3L**), pollen grain (**Figure 3M**) and seeds (**Figure 3K**). The GUS staining detected during the development of the floral organs and the seed may indicate that ADT2 is playing a role during sexual organ development and plant reproduction.

Complementation of T-DNA phenotype by *ADT2*

To functionally test if the observed female semi-sterility phenotype is the result of a deficiency in the expression of *ADT2*, we transformed heterozygous *adt2-1* mutants with a plant expression construct containing the *ADT2* full length cDNA under control of the same *ADT2* 5'-upstream regulatory region used for the P_{ADT2}*:GUS* construct. The pGWB610-P_{ADT2}*:ADT2* (C-terminal FLAG tag) construct was done for this purpose. Transgenic lines were generated and selected for BASTA resistance. T2 plants from 3 independent transgenic lines for the P_{ADT2}*:ADT2* construct were analyzed by PCR, showing the existence of homozygous T-DNA plants among the T2 progeny. T3 progeny from the T2 homozygous plants were grown and analyzed by PCR, confirming the homozygous *adt2* T-DNA genotype of their respective parental plants (**Figure 4A**). No homozygous *adt2* T-DNA mutants were found among the T2 and T3 control plants transformed with the empty vector. Moreover, T3 plants transformed with the P_{ADT2}*:ADT2* construct showed siliques with a restored wild-type-like length and seed set, lacking the characteristic female semi-sterility phenotype associated with the T-DNA insertion (**Figures 4C, 4D**), demonstrating the ability of P_{ADT2}*:ADT2* to recover the knock-out phenotype. P_{ADT2}*:ADT2* expression was confirmed by western blot for the FLAG tagged recombinant protein (**Figure 4B**).

PHA2 and ADT3 heterologous expression rescue the *adt2* mutant phenotype

Although the arogenate pathway has been proposed as the main route of Phe biosynthesis in plants, PDT activity in plants has been reported in the last years and under certain conditions (Warpeha et al., 2006; Yoo et al., 2013). Since *Arabidopsis ADT2* encodes a bifunctional ADT/PDT enzyme, we were interested in verifying if the *adt2* T-DNA phenotype could be rescued by overexpressing monofunctional PDT or ADT enzymes.

To test if *adt2* T-DNA plants can be complemented by a monofunctional PDT enzyme, *Saccharomyces cerevisiae PHA2* cDNA, which encodes a PDT without any detectable ADT activity (Bross et al., 2011), was used to transform heterozygous *adt2-1* T-DNA plants. In order to target heterologous yeast PHA2 into *Arabidopsis* plastids, where Phe biosynthesis takes place (Rippert et al., 2009), a chimeric *PHA2* gene was made fusing a 243-bp-length fragment from *Arabidopsis ADT2* putative chloroplats transit peptide into the 5'-flank of *PHA2* cDNA. This chimeric ORF was cloned into pGWB5 expression vector (CaM *35S* promoter; C-terminal GFP fusion) and the pGWB5-*PHA2-GFP* construct was transiently expressed in *Nicotiana benthamiana* leaves via agroinfiltration. The chimeric *PHA2* gene was successfully expressed and plastidial localization of PHA2 protein confirmed by confocal microscopy (Supplemental figure 2A). To further confirm that heterologous PHA2 was effectively targeted into *Arabidopsis* plastids, we performed a stable *Arabidopsis* transformation using the same pGWB5-*PHA2-GFP* chimeric construct, which confirmed plastidial localization for the PHA2-GFP fusion protein (Supplemental figure 2B).

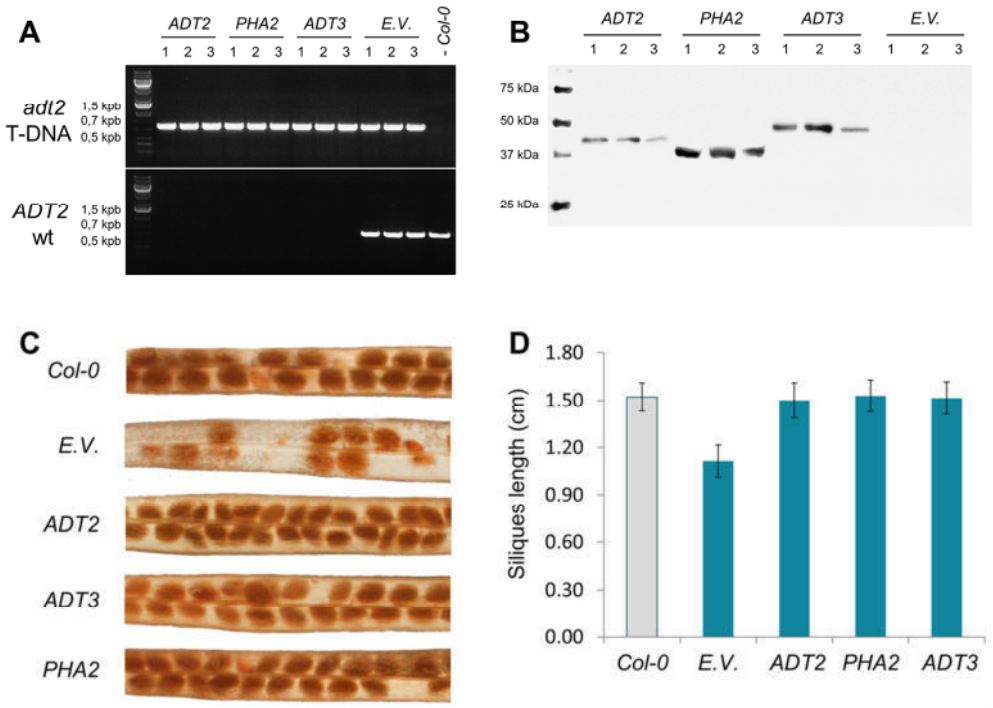


Figure 4. **Complementation of the *adt2* mutant phenotype**. **(A)** PCR test of T3 adult plants, confirming the existence of homozygous *adt2-1* T-DNA insertion mutants lines when transformed with *ADT2, PHA2 or ADT3* under control of ADT2 promoter (P_{ADT2}). Numbers ranging from 1 to 3 are indicating the 3 independent transgenic lines selected for each construct. **(B)** Anti-FLAG western blot of total protein extracts from 3 independent T3 transgenic lines (1 to 3) expressing *ADT2*, *ADT3* or *PHA2* under control of P_{ADT2}. T3 lines transformed with the empty vector were used as negative control. The predicted masses for the recombinant protenins were: 42.1 kDa (ADT2), 38.2 kDa (PHA2) and 46.1 kDa (ADT3). **(C)** Mature siliques from wild-type plants (*Col-0*) compared to T3 T-DNA mutant plants transformed with the empty vector (*E.V.*), *ADT2*, *PHA2* or *ADT3* expression constructs. **(D)** Average siliques length observed in T3 transformant plants compared to wild-type (*Col-0*), when transformed with the empty vector (*E.V.*), *ADT2, PHA2* or *ADT3* under control of the *ADT2* promoter.

Heterozygous *adt2-1* T-DNA mutant plants were transformed using the construct pGWB610-P_{ADT2}:*PHA2*, carriyin the chimeric PHA2 ORF previously described. 3 independent P_{ADT2}:*PHA2* transgenic lines were selected for their analysis. Similarly as described for the P_{ADT2}:*ADT2* transgenic plants, we were able to confirm the presence of T-DNA homozygous plants among the T2 P_{ADT2}:*PHA2* lines. Their respective T3 progenies confirmed the homozygous *adt2-1* genotype and also showed siliquas with a wild type phenotype (**Figures 4A-4D**), suggesting that PDT can operate in an alternative pathway for Phe biosynthesis in the female gametophyte.

In a similar way, heterozygous *adt2-1* T-DNA plants were also transformed with a pGWB610-P_{ADT2}:*ADT3* construct, containing the full-length cDNA of the *Arabidopsis ADT3* gene, that as been shown to work *in vitro* as a monofunctional ADT without any detectable PDT activity (Cho et al., 2007). As mentioned before for the P_{ADT2}:*ADT2* and P_{ADT2}:*PHA2* transgenic plants, the PCR and phenotype analysis of the T2 and T3 progenies from different independent transgenic lines confirmed the functional complementation of gametophytic *adt2* defficiency by *ADT3*, pointing that the expression of a monofunctional ADT enzyme under control of P_{ADT2} can also rescue the observed mutant phenotype (**Figures 4A-4D**).

DISCUSSION

Plants life cycle is constituted by the alternation of two strongly differentiated generations: the sporophyte, a diploid multicellular stage which is developed from the zygote, and the gametophyte, a haploid stage developed from spores and responsible for the gamete formation and

sexual reproduction (Yang et al., 2010). Most flowering plants presents a characteristic seven-celled female gametophyte, also called embryo sac, which represents one of the major defining characteristics of this group of plants (Sundaresan and Alandete-Saz, 2010). Flowering plant's female gametophytes possess two gametes, the egg cell and the central cell, that undergo double-fertilization. Whereas the egg cell will develop the embryo after being fertilized, the central cell will develop into the seed endosperm, which will supply nutrients to the embryo and nascent seedling in the early stages of germination (Tekleyohans et al., 2016).

Genetic analysis tools have revealed mutants defective in almost all stages of female gametophyte development, and their analysis during the last decade have started to reveal novel features of the female gametophyte developmental program. Nervertheless, we lack almost any information about the relevance of Phe metabolism during plant reproduction. In a comprehensively identification of genes required for female gametophyte development carried out by Pagnussat et al. (2005), it was shown the existence of a certain number of genes involved in secondary metabolism that are essential for female gametophyte development in *Arabidopsis*. From a total number of 130 mutants lines that were defective in female gametophyte development and function, around a 10% were found to be mutants for genes related to secondary metabolism. Very interestingly, a mutant defective in the *At2g22250* locus, which encodes the single-copy *PAT* gene from *Arabidopsis* (Maeda et al., 2011), was described as Maternal Effect Embryo Arrest 17 (MEE17) presenting embryonic arrest at the one-cell zygotic stage (Pagnussat et al. 2005). Here, we have demonstrated that *adt2-1* and *adt2-2* mutant alelles cause a defect in the female gametophyte development in *Arabidopsis*, resulting in ovule lethality and hence being unable to obtain homozygous *adt2* plants (**Figure 2**).

Given the fact that both genes, *PAT* and *ADT2,* are essential for *Arabidopsis* reproduction, it highlights the importance of Phe metabolism for an adequate embryo development, either for protein and/or phenolic biosynthesis.

Recent investigations into aromatic amino acid biosynthesis have revealed that plants, in contrast to most microbes, predominantly synthesize phenylalanine via the arogenate pathway (Maeda et al., 2010; Yoo et al., 2013). Nevertheless, it has been suggested that plants can also utilize a microbial-like phenylpyruvate pathway to produce phenylalanine. De la Torre et al., (2014) reported that *N. benthamiana PDT1* and *PDT2* are clearly overexpressed in silenced tobacco leaves when *PAT* is RNAi silenced. Moreover, *NbPPAT*, homolog of the *Arabidopsis* gene *At5g36160*, which encodes an aromatic amino acid aminotransferase that is able to convert phenylpyruvate into Phe (Prabhu and Hudson, 2010), was approximately 2- to 3-fold overexpressed in *PAT* RNAi *Nicotiana* leaves, suggesting that phenylpyruvate pathway is induced when the arogenate pathway becomes a limitant step in Phe biosynthesis. These results are according to Yoo et al. (2013) who reported that flux through the phenylpyruvate pathway is increased when the entry point to the arogenate pathway is blocked in petunia petals. These authors linked the plant phenylpyruvate pathway to a cytosolic tyrosine aminotransferase (*At5g53970*), which could suggest the existence of a coordinated regulation between tyrosine catabolism and the hypothetical extra-plastidial Phe metabolism.

The results presented in this paper show that *S. cerevisiae PHA2* gene, which encodes a PDT with strict specificity for prephenate, can rescue the *adt2* mutant phenotype when *PHA2* is expressed under control of the *ADT2*

5'-upstream regulatory region in the heterozygous *adt2-1* mutant background (**Figure 4**). Moreover, yeast PHA2 protein has been targeted within plants plastids, fusing PHA2 cDNA with ADT2 putative plastid transit peptide, allowing PHA2 to be adequately located in this subcellular compartment. Very interestingly, as far as ADT2 can operate either as ADT or PDT *in vitro* and *in vivo* (Cho et al., 2007; Bross et al., 2011), the *adt2* complementation by PHA2 strongly suggest that phenylpyruvate pathway could be taking part in the biosynthesis of Phe within the female gametophyte of *Arabidopsis*. In addition, the biosynthesis of Phe via PDT would require the co-expression of a phenylpyruvate-aminotransferase enzyme (Fischer and Jensen, 1987a). In the model proposed by Yoo and collaborators (2013) for *Petunia hybrida* petals, phenylpyruvate rendered by PDT in the plastids would be exported to the cytosol, were it is converted into Phe by PhPPY-AT. The authors suggested *At5g53970* as the putative *Arabidopsis* ortholog of PhPPY-AT, which was originally characterized as a tyrosine aminotransferase involved in tocopherol biosynthesis in the cytosol (Riewe et al., 2012), hence connecting phenylpyruvate-mediated Phe biosynthesis and Tyr catabolism. However, the data available from the *Arabidopsis* female gametophyte transcriptome (Wuest et al., 2010) reveal that *At5g53970* expression is not detected in this organ. According to previous reports based on the Arabidopsis Information Resource (TAIR; http://www.arabidopsis.org/) there are six genes in *Arabidopsis* which could encode, based on computational prediction, enzymes with tyrosine aminotransferase activity: *At5g36160*, *At5g53970*, *At4g23590*, *At4g28410*, *At4g28420*, *At2g24850* (Riewe et al., 2012). From these six genes, only *At5g36160* and *At5g53970* have been experimentally confirmed as tyrosine aminotransferase conding genes (Lopukhina et al., 2001; Prabhu and Hudson, 2010). According to the data published by Wuest and collaborators (2010), *At2g24850* remains as the single tyrosine

aminotransferase family member of *Arabidopsis* with detectable expression levels in the female gametophyte of this plant, and more specifically in the egg cell, but not in the central cell and the synergids, providing evidence for additional tyrosine aminotransferases isoforms involved in phenylpyruvate pathway.

As mentioned before, ADT2 can operate either as ADT or PDT *in vitro* and *in vivo* (Cho et al., 2007; Bross et al., 2011). Given the fact that PHA2 heterologous expression under control of P_{ADT2} can rescue the *adt2* mutant phenotype, we were interested in investigating if the female semi-sterility phenotype caused by the deficiency in ADT2 could be specifically related to the phenylalanine biosynthesis from phenylpyruvate. For this purpose, heterozygous *adt2* T-DNA plants were transformed with a P_{ADT2}*:ADT3* expression construct, containing the full-length cDNA of the *Arabidopsis ADT3* gene, which as been shown to work *in vitro* as a monofunctional ADT without any detectable PDT activity (Cho et al., 2007). Our results (**Figure 4**) demonstrate that *ADT3*, similarly to *S. cerevisiae PHA2*, is able to fully recover the *adt2* female semi-sterility phenotype, pointing that the expression of a monofunctional ADT enzyme under control of P_{ADT2} can rescue the observed mutant phenotype. Together with P_{ADT2}*:PHA2* complementation of *adt2* mutants, these results lay out that the biosynthesis of Phe in the female gametophyte of *Arabidopsis* can occur either from arogenate or phenylpyruvate. Thus, the female semi-sterility phenotype observed when there is a deficiency in ADT2 in the female gametophyte of *Arabidopsis* may be mainly attributed to the *ADT2* gene expression regulation during the development of this floral organ, rather than ADT2 particular enzymatic or kinetic properties, more particularly its ability to alternatively catalyze the dehydration and decarboxylation of both prephenate or arogenate. Additional efforts will be required to gain more

insight into the expression regulatory networks that lie behind Phe biosynthesis during the formation of the female gametophyte.

On the other hand, *PAT* was previously reported as an essential gene for the female gametophyte development in *Arabidopsis* (Pagnussat et al., 2005). Since PAT is encoded by a single-copy gene in most of the plants, included *Arabidopsis*, it could suggest that the biosynthesis of phenylalanine occurs mainly through the arogenate pathway in the female gametophyte, although phenylpyruvate pathway could also be operative, as we report here. Nevertheless, we do not know whether *pat* and *adt2* female gametophyte lethality are the result of the same development/reproductive defect or, by contrast, are caused at distinct stages of female gametophyte development and fecundation, or different cell types within the embryo sac. Furthermore, it is important to note that PAT is also an essential enzymatic activity for Tyr biosynthesis, since a prephenate dehydrogenase coding gene (PDH; **Figure 1**) has not been described to date in *Arabidopsis* (Schenck et al., 2015). Yet, *pat* lethality could be more closely linked to a deficiency in the biosynthesis of Tyr, as there is not an alternative pathway for the biosynthesis of this amino acid in *Arabidopsis*, whereas without a supply or arogenate Phe could be produced from phenylpyruvate in a metabolic context where ADT2, a functional PDT, is expressed.

In conclusion, our results provide new evidence about the aromatic amino acid biosynthesis in plants, advancing the putative co-existence of cytosolic and plastidial Phe biosynthetic routes in the female gametophyte of *Arabidopsis*, which could differ from one cellular type to another. In adition, our results set out new questions about gene regulation in the biosynthesis of this amino acid and its derived phenolic compounds, which will require further investigation to be adequately addressed.

EXPERIMENTAL PROCEDURES

Arabidopsis plants manipulation and genotyping

T-DNA insertion lines for all six *ADT* genes in *Arabidopsis* (supplemental Table S1) were requested from the Salk Institute Genomic Analysis Laboratory. *Arabidopsis* plants were grown in soil in a climate controlled room under long-day photoperiod (16h light /8h dark) at 22º C.

Segregation ratios corresponding to *adt2* T-DNA insertional mutants were screened by PCR using genomic DNA extracted from two-weeks-old seedlings grown in MS medium. Genomic DNA was extracted as described by Lu (2011), with the exception that the entire plant/rosette was used for the extraction. The T-DNA direction was confirmed by PCR for both lines. T-DNA insertion in the lines SALK_044042 and SALK_081342 was detected using the primers 044Fwd and LBTDNA or 081Fwd and LBTDNA, respectively. The *ADT2* non-mutant allele was detected using the primers 044Fwd and 044Rvs, for the line SALK_044042, or 081Fwd and 081Rvs, for the line SALK_081342. PCRs were done with GoTaq® Flexi DNA Polymerase (Promega).

For the T-DNA lines of the remaining 5 *ADTs*, DNA was extracted from leaves of individual plants, and these samples were then used as a template for two PCRs with different primer sets. The following primers were paired with the primer LBTDNA to detect the corresponding T-DNA insertion, from ADT1 to 6 (except ADT2): TADT1Rvs, TADT3Fwd, TADT4Fwd, TADT5Fwd, and TADT6Rvs. The corresponding wild-type alleles were detected in a separate PCR with the following primers combinations: TADT1Fwd / TADT1Rvs, TADT3Fwd / TADT3Rvs,

TADT4Fwd / TADT4Rvs, TADT5Fwd / TADT5Rvs, and TADT6Fwd / TADT6Rvs.

cDNAs cloning

A. thaliana ADT2 and *ADT3* ORFs were cloned from cDNA obtained from adult rosette leaves using the PCR primers ADT2Fwd/ADT2Rvs and ADT3Fwd/ADT3Rvs, respectively. The 5'- upstream region of *ADT2* (P_{ADT2}) was cloned from genomic DNA of *Arabidopsis* extracted from mature rosette leaves. The primers PADT2Fwd and PADT2Rvs were used for this PCR. *PHA2* was amplified from genomic DNA of *S. cerevisiae* strain BY4741 using the PCR primers ctpPHA2Fwd and PHA2Rvs whereas the putative plastid transit peptide from *ADT2* was amplified using the primers ADT2Fwd and ctpPHA2Rvs. Both PCR products were purified and fused together by PCR using the primers ADT2Fwd and PHA2Rvs. The resulting PCR products were purified from gel and cloned into pJET1.2/blunt (Thermo Scientific) and confirmed by sequencing.

Plasmid DNA constructs

A. thaliana ADT2 was amplified from *pJET1.2:ADT2* construct using the primers PADT2ADT2Fwd and ADT2AttB2. P_{ADT2} was amplified from the *pJET1.2:*P_{ADT2} construct using the primers PADT2AttB1 and PADT2ADT2Rvs. Both PCR products were purified and fused using the primers PADT2AttB1 and ADT2AttB2. The resulting PCR product was re-amplified using the primers AttB1 and AttB2, purified from gel, and cloned into Gateway® pDONR207™ (Thermo Fisher Scientific) donor vector using BP Clonase® II enzyme mix. Subsequently, it was recombined into the pGWB610 destination vector using the *Gateway® LR* Clonase® II mix.

To generate the *PHA2* expression construct, *PHA2* ORF was amplified from the *pJET1.2:PHA2* construct using the oligonucleotides pADT2ADT2Fwd and PHA2Rvs, and purified from gel. In a second PCR reaction, the primers PADT2AttB1 and PHA2AttB2 were used to fuse the *PHA2* PCR product to P_{ADT2}, which was previously amplified from the *pJET1.2:*P_{ADT2} construct using the primers PADT2Fwd and PADT2ADT2Rvs. The resulting PCR product was purified from gel and re-amplified using the primers AttB1 and AttB2. This product was finally purified from gel and cloned first into pDONR207™, and then into the pGWB5 and pGWB610 destination vectors, as previously described.

A. thaliana ADT3 was amplified from *pJET1.2:ADT3* construct using the primers PADT2ADT3Fwd and ADT3AttB2. P_{ADT2} was amplified from the *pJET1.2:*P_{ADT2} construct using the primers PADT2AttB1 and PADT2ADT3Rvs. Both PCR products were purified and fused using the primers PADT2AttB1 and ADT3AttB2. The resulting PCR product was re-amplified using the primers AttB1 and AttB2, purified from gel, and cloned into Gateway® pDONR207™ (Thermo Fisher Scientific) donor vector using BP Clonase® II enzyme mix. Subsequently, it was recombined into pGWB610 destination vector, as described for the pGWB610:*ADT2*.

P_{ADT2} was fused to a *GUS* reporter gene (P_{ADT2}:*GUS*) amplifying P_{ADT2} from the *pJET1.2:*P_{ADT2} construct with the primers PADT2AttB1 and PADT2AttB2. This PCR product was re-amplified using the primers AttB1 and AttB2, purified from gel and cloned into pDONR207™. Later, it was recombined into pGWB3 destination vector.

All PCR reactions were done using the iProof™ High-Fidelity Master Mix (Bio-Rad Laboratories). *E. coli* DH5α strain was used for all the transformations required.

adt2 plants complementation

Electrocompetent cells of *Agrobacterium tumefaciens* C58C1 strain were transformed with the pGWBs expression constructs and selected for the corresponding antibiotic resistance. Heterozygous *adt2* plants (T0) from the SALK_044042 line (*adt2-1*) were transformed using the floral dip method (Clough and Bent, 1998). Control plants were transformed with the corresponding empty vectors.

T0 plants were grown and T1 seeds collected from them and selected for DL-phosphinothricin (Duchefa Biochemie) resistance in 0.5X MS medium. Resistant plants were transferred into soil. The transgenic construct and the *adt2-1* T-DNA insertion were confirmed by PCR using the primers AttB1 / AttB2 and LBTDNA / 044Fwd, respectively. 3 transgenic lines were selected for each construct and control empty vectors, and their T2 seeds were selected for DL-phosphinothricin resistance in 0.5X MS plates. 16 plants for each transgenic line were transferred into soil. The presence of homozygous *adt2-1* plants was analyzed by PCR in these plants. T3 plants descending from homozygous *adt2-1* T2 plants were selected for DL-phosphinothricin resistance; the homozygous *adt2-1* genotype was subsequently confirmed in the T3 progeny, demonstrating the homozygous *adt2-1* genotype of the corresponding T2 parentals and therefore the functional complementation of the T-DNA knock-out phenotype.

Protein extraction and western blot

Total proteins from young inflorescences of P_{ADT2}:*ADT2*, P_{ADT2}:*ADT3* and P_{ADT2}:*PHA2* trasgenic plants were extracted in 1% sodiumd dodecyl sulfate (SDS). Protein concentration was determined by the Bradford dye-binding method (Bradford, 1976), after removing SDS from the samples (Zaman and Verwilghen, 1979). SDS-PAGE and western blots were carried out as described previously (Cánovas et al., 1984). The following antibodies were used: 1:1000 anti-FLAG (SantaCruz Biotechnology) and 1:10,000 anti-mouse HRP-conjugate (SantaCruz Biotechnology).

GUS staining

T1 seeds from Col-0 plants transformed with the construct *pGWB3:*P_{ADT2}*:GUS* were selected in 0.5X MS media for Hygromycin B (Duchefa Biochemie; 49 µL of antibiotic per liter of medium) resistance. Resistant seedlings were transferred into soil and the P_{ADT2}:*GUS* insertion confirmed by PCR using the AttB1 / AttB2 pair of primers. Samples from the transgenic seedlings were taken from different organs and developmental stages, as indicated in the corresponding figure. Collected tissues were immediately infiltrated with the GUS staining solution using a vacuum pump. Infiltrated samples were incubated overnight at 37º C in the dark. After GUS incubation, samples were incubated in pure ethanol until get fully decolorized, and then transferred into glycerol for imaging and conservation. GUS staining solution was prepared as described by Kirchsteiger et al., 2012.

PHA2-GFP detection

N. benthamiana leaves from around 1-month-plants, grown under long-dat conditions at 25ºC, were agroinfiltrated with *A. tumefaciens* C58C1 carrying the pGWB5:PHA2 construct. After two days, green fluorescence was visualized in fresh leaves under a Leica TCS SP5 confocal microscope (Leica Microsystems).

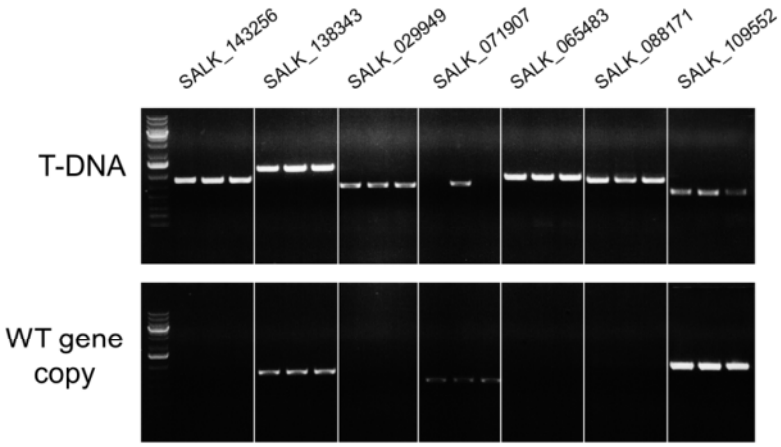
Acknowledgments

We would like to give thanks to Ana Álvarez for her helping on *Arabidopsis* seedlings genomic DNA extraction and genotyping. We are also grateful to Doctora Maria Belén Pascual Moreno, for her advice on GUS histochemical techniques. Research work in the author's laboratory was supported by grants from the Spanish Ministerio de Economía y Competitividad (BIO2015-69285-R) and Junta de Andalucía (BIO-474). JEA is the recipient of a Formación Profesorado Universitario fellowship (Ministerio Educación, Spain).

Suplemental data

Supplemental table 1. SALK T-DNA insertion lines of *Arabidopsis* ADTs.

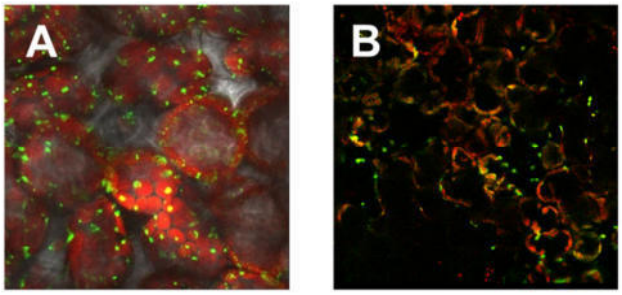
gene name	*locus*	SALK line	Homozyg.	Insert position
ADT1	*AT1G11790*	SALK_138343	Y	5' UTR
		SALK_143257	Y	Exon 2
ADT2	*AT3G07630*	SALKseq_081342	N	5' UTR
		SALKseq_044042	N	Exon 6
ADT3	*AT2G27820*	SALK_029949	Y	Exon 1
		SALK_071907	Y	Exon 1
ADT4	*AT3G44720*	SALK_065483	Y	Exon 1
ADT5	*AT5G22630*	SALK_028611	N	Exon 1
		SALK_088171	Y	Exon 1
ADT6	*AT1G08250*	SALK_109552	Y	Exon 1



Supplemental figure 1. PCR analysis of SALK T-DNA insertion lines of *Arabidopsis* ADTs. Above, PCR detection of the T-DNA insertion using specific primers; below, PCR detecting the wild-type allele of the corresponding *ADT*. PCR analysis for the line SALK_028611 (*ADT5*) failed. Ladder: GeneRuler 1 kb plus (Thermo Fischer Scientific).

Supplemental table 2. Expression pattern of *Arabidopsis ADTs* in the female gametophyte. + indicates significant positive mRNA detection by microarray; - indicates no expression. Raw data extracted from Wuest et al., 2010.

gene name	**cell type**		
	synergid	egg cell	central cell
ADT1	+	+	+
ADT2	-	+	+
ADT3	-	-	-
ADT4	-	-	-
ADT5	-	-	-
ADT6	-	-	-



Supplemental figure 2. Subcellular localization of plastid-targeted chimeric PHA2-GFP fusion protein with the putative plastid transit peptide from ADT2 fused to its N-terminus. (A) Transient expression in *Nicothiana benthamiana* leaves under control of CaM P*35S* promoter. (B) Same chimeric PHA2-GFP fusion protein expressed in leaves of stable transformant *Arabidopsis*.

Suplemental table 3. List of PCR primers used in this paper. Sequences are indicated from 5'-.

Primer name	Primer sequence
044Fwd	GATACTGCTGGCGCTGCGAAGG
044Rvs	ACTCCAGGCCCTTCCTCTAACG
081Fwd	ACTCCTGGTTTTATGGTCTGCC
081Rvs	CTTCACTTCCATAGCTCTCACGG
LBTDNA	GCGTGGACCGCTTGCTGCAACT
TADT1Fwd	GTCGTATAGTCTCATAAATGG
TADT1Rvs	CTGAGGATGGCTTAGAACAC
TADT3Fwd	ATTCGAAGTTGCGTTTCAAG
TADT3Rvs	CAAACTAATGTTCCTAAACGC
TADT4Fwd	GACTATTACTGATCTATCTCC
TADT4Rvs	ACATTAGGAAACGAGTGACG
TADT5Fwd	GTGATCAATTCGACGTTGCG
TADT5Rvs	CCTTTATGTTCTTGAGCTGCG
TADT6Fwd	ATGAAAGCTCTATCATCTTC
TADT6Rvs	TGCGATCCAAAGCTCCACC

ADT2Fwd	ATGGCAATGCACACTGTTCG
ADT2Rvs	GAGCATTGTAGTGTCCACTGG
PHA2Fwd	ATGGCCAGCAAGACTTTGAGGG
PHA2Rvs	TTATTTGTGATAATATCTCTC
ADT3Fwd	ATGAGAACTCTCTTACCTTCC
ADT3Rvs	CAATGAAAATGTTGATGACG
ctpRvs	AGGAAGCAGAGGCGAGTC
ctpPHA2Fwd	GACTCGCCTCTGCTTCCTATGGCCAGCAAGACTTTGAGGG
ctpPHA2Rvs	CCCTCAAAGTCTTGCTGGCCATAGGAAGCAGAGGCGAGTC
ADT2AttB1	AAAAAGCAGGCTTAATGGCAATGCACACTGTTCG
ADT2AttB2	AGAAAGCTGGGTTGAGCATTGTAGTGTCCACTGG
PHA2AttB2	AGAAAGCTGGGTTTTATTTGTGATAATATCTCTC
ADT3AttB2	AGAAAGCTGGGTTCAATGAAAATGTTGATGACG
pADT2Fwd	ACCTTTTCCATTCTAATTCC
pADT2Rvs	TGATGTTGTTTTGACGGCTG
pADT2AttB1	AAAAAGCAGGCTTAACCTTTTCCATTCTAATTCC
pADT2AttB2	AGAAAGCTGGGTTTGATGTTGTTTTGACGGCTG
pADT2ADT2Fwd	CAGCCGTCAAAACAACATCAATGGCAATGCACACTGTTCG
pADT2ADT2Rvs	CGAACAGTGTGCATTGCCATTGATGTTGTTTTGACGGCTG
pADT2ADT3Fwd	CAGCCGTCAAAACAACATCAATGAGAACTCTCTTACCTTCC
pADT2ADT3Rvs	GGAAGGTAAGAGAGTTCTCATTGATGTTGTTTTGACGGCTG
AttB1	GGGGACAAGTTTGTACAAAAAAGCAGGCT
AttB2	GGGGACCACTTTGTACAAGAAAGCTGGGT

References

Bonner C and Jensen R (1987) *Prephenate aminotransferase*. Methods Enzymol, 142:479-487.

Bonner C and Jensen R (1987) *Arogenate dehydrogenase*. Methods Enzymol, 142:488-894.

Bradford MM (1976) *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 72:248-254.

Bross CD, Corea OR, Kaldis A, Menassa R, Bernards MA and Kohalmi SE (2011) *Complementation of the* pha2 *yeast mutant suggests functional differences for*

arogenate dehydratases from Arabidopsis thaliana. Plant Physiol Biochem, 49(8):882-890.

Cánovas F, Valpuesta V and de Castro IN (1984) *Characterization of tomato leaf glutamine-synthetase*. Plant Science Letters, 37:79-85.

Cho MH, Corea OR, Yang H, Bedgar DL, Laskar DD, et al. (2007) *Phenylalanine biosynthesis in* Arabidopsis thaliana. *Identification and characterization of arogenate dehydratases*. J Biol Chem. 282(42):30827-30835.

Clough SJ and Bent AF (1998) *Floral dip: a simplified method for* Agrobacterium-*mediated transformation of* Arabidopsis thaliana. Plant J, 16(6):735-743.

Corea OR, Ki C, Cardenas CL, Kim SJ, Brewer SE, et al. (2012) *Arogenate dehydratase isoenzymes profoundly and differentially modulate carbon flux into lignins*. J Biol Chem. 287(14):11446-11459.

de la Torre F, El-Azaz J, Avila C and Cánovas FM (2014a) *Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism*. Plant Physiol, 164:92-104.

El-Azaz J, de la Torre F, Ávila C and Cánovas FM (2016) *Identification of a small protein domain present in all plant lineages that confers high prephenate dehydratase activity*. Plant J, 87(2):215-229.

Fischer R. and Jensen R (1987a) *Prephenate dehydratase (monofunctional)*. Methods Enzymol, 142:507-512.

Fischer R. and Jensen R (1987b) *Prephenate dehydrogenase (monofunctional)*. Methods Enzymol, 142:503-507.

Lopukhina A, Dettenberg M, Weiler EW and Hollander-Czytko H (2001) *Cloning and characterization of a coronatine-regulated tyrosine aminotransferase from* Arabidopsis. Plant Physiol, 126:1678-1687.

Lu Y (2011) *Extract Genomic DNA from Arabidopsis Leaves (Can be Used for Other Tissues as Well)*. Bio-protocol. Bio101: e90. DOI: 10.21769/BioProtoc.90.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, et al. (2010) *RNAi suppression of Arogenate Dehydratase 1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22(3):832-849.

Maeda H, Yoo H and Dudareva N (2011) *Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate*. Nat Chem Biol, 7(1):19-21.

Maeda H and Dudareva N (2012) *The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants*. Annu Rev Plant Biol, 63:73-105.

Page DR and Grossniklaus U (2002) *The art and design of genetic screens:* Arabidopsis thaliana. Nat Rev Genet, 3:124-136

Pagnussat GC, Yu HJ, Ngo QA, Rajani S, Mayalagu S, Johnson CS, Capron A, Xie L-F, Ye D, Sundaresan V (2005) *Genetic and molecular identification of genes required for*

female gametophyte development and function in Arabidopsis. Development, 132: 603-614

Prabhu PR and Hudson AO (2010) *Identification and Partial Characterization of an L-Tyrosine Aminotransferase (TAT) from* Arabidopsis thaliana. Biochem Res Int, 2010:549572.

Riewe D, Koohi M, Lisec J, Pfeiffer M, Lippmann R, et al. (2012) *A tyrosine aminotransferase involved in tocopherol synthesis in* Arabidopsis. Plant J, 71(5):850-859.

Rippert P, Puyaubert J, Grisollet D, Derrier L and Matringe M (2009) *Tyrosine and phenylalanine are synthesized within the plastids in* Arabidopsis. Plant Physiol, 149:1251-1260.

Schenck CA, Chen S, Siehl DL and Maeda H (2015) *Non-plastidic, tyrosine-insensitive prephenate dehydrogenases from legumes*. Nat Chem Biol, 11(1):52-57.

Sundaresan V and Alandete-Saez M (2010) *Pattern formation in miniature: the female gametophyte of flowering plants*. Development, 137:179-189.

Tekleyohans DG, Nakel T and Gross-Hardt R (2016) *Patterning the female gametophyte of flowering plants*. Plant Physiol, Preview. Published on December 5, 2016, as DOI:10.1104/pp.16.01472

Warpeha KM, Lateef SS, Lapik Y, Anderson M, Lee B-S and Kaufman LS (2006) *G-protein-coupled receptor 1, G-protein Gα-subunit 1, and prephenate dehydratase 1 are required for blue light-induced production of phenylalanine in etiolated* Arabidopsis. Plant Physiol, 140:844-855.

Wuest SE, Vijverberg K, Schmidt A, Weiss M, Gheyselinck J, et al. (2010) Arabidopsis *female gametophyte gene expression map reveals similarities between plant and animal gametes*. Curr Biol, 20(6):506-512.

Yang W-C, Shi D-Q and Chen Y-H (2010) *Female Gametophyte Development in Flowering Plants*. Annu Rev Plant Biol, 61:89-108.

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, et al.(2013) *An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase*. Nat Commun, 4:2833.

Zaman Z and Verwilghen RL (1979) *Quantitation of proteins solubilized in sodium dodecyl sulfate-mercaptoethanol-Tris electrophoresis buffer*. Anal Biochem, 100(1):64-69.

V. GENERAL DISCUSSION

1. The study of Phe biosynthesis reveals a complex history about the origin of plastids

The phylogenetic history of land plants is still full of open questions with difficult interpretation. Already on the origin of photosynthetic eukaryotes, the acquisition of plastids, most likely through endosymbiotic events with photosynthetically competent organisms, is still a matter of discussion. While the various algae phyla have plastids from multiple independent origins, it seems to be clear that inside the viridiplantae lineage (composed by modern green algae and land plants or embryophytes) plastids where acquired through an endosymbiotic event with a cyanobacteria. This adquisition, in which a photosynthetic cyanobacterium was captured by a heterotrophic protist and eventually transformed into a double-membrane-bound organelle, is termed as “primary endosymbiosis” (Chan et al., 2011). This key event in eukaryote evolution is suggested to occur approximately 1.5 billion years ago, resulting in a chimeric eukaryote that gave rise to the three major photosynthetic lineages: the red, green (including land plants), and glaucophyte algae. These three lineages form the clade *Archaeplastida*. The basic idea that a single evolutionary event resulted in algae and plant plastid is supported by many molecular evidences, especially by phylogenetic analysis that show the high degree of similarity between plastid and cyanobacteria genes (Baum, 2013).

Although there is no additional origins of primary plastids, at least seven additional eukaryotic lineages, as euglenoids, brown algae and dinoflagellates, have established endosymbiotic relationships with unicellular algae as plastids (Keeling, 2010). This shows that, during the evolution of photosynthetic eukaryotes, it has been easier to integrate a photosynthetic eukaryotic cell rather than a photosynthetic prokaryote. Ball

et al. (2013) argued that *Archaeplastida* were able to accommodate a prokaryotic guest because they were subjected to coinfections by chlamydial endoparasites, which could facilitate establishing a metabolic dialogue between all the parts involved. In this infected host, the presence of certain enzymes secreted from the parasitic chlamydia to the host cytoplasm could allow a mutualistic association between the eukaryotic cell and an incipient plastid. Otherwise, this association had been impossible because of the inability of the host to synthesize and use storage carbohydrates from ADP-Glucose, most likely the carbohydrates exported by the engulfed cyanobacteria (Facchinelly et al., 2013). The fact that phylogenomic analyses have identified between 21 and 55 specific genes that appear to have been acquired by plants from *Chlamydiae*, as a result of horizontal gene transfer, supports this hypothesis (Baum, 2013). In the light of this analysis, it becomes more easy to understand how unlikely was the primary endosymbiosis and why it has remained as a single event in life's history, even though of its undeniable evolutionary advantages.

When the nuclear localization of genes encoding proteins of plastidial localization was demonstrated, it became clear that plastids are genetically dependent of their eukaryotic host and had relinquished many of their original genes during evolution. As a result of this, plastidial genomes suffered a strong decrease in size and many genes were transferred from the endosymbiont to the nuclear genome. Whereas the average genome size for modern cyanobacteria is up to 3-3.5 Mpb (Kaneko and Tabata, 1997), plants plastids genomes range from 0.109 (*Ephedra equisetina*) to 0.217 Mbp (*Pelargonium sp.*), with an average size of 0.120-0.160 Mpb (Daniell et al., 2016). Similarly, the number of genes decreased from more than 3,000 putative genes in *Synechocystis* (Kaneko et al., 1996) to around typically 120-130 genes in the plastids of most plants (Daniell et al., 2016)

and about 200 in red algae (McFadden, 2001). In coherence with this loss of genes in the plastid genome, it has been estimated in *Arabidopsis* that more than 2,000 nuclear genes encodes proteins with plastid functions (Abdallah et al., 2000).

Aromatic amino acid biosynthesis genes are an example of genes encoded in the nuclear genome but whose products are located in the plastids of higher plants. Several experimental and *in silico* evidences allow to conclude that the plastids contain a full set of biosynthetic enzymes for the production of Phe, Tyr and Trp from phospho*enol*pyruvate and erytrose-4-phosphate (see section I.4 and I.5) (Maeda and Dudareva, 2012). In contrast with original assumptions, modern genomic studies, supported by the possibility of getting access to progressively larger and better genomic resources, have revealed unexpectedly complex histories of cyanobacterial and plastid genes, including those transferred from the plastidial to the nuclear genome. Recent reports suggest that the evolutionary origin of some plants genes of the prephenate pathway, more particularly PAT and ADTs, but not ADHs, is more closely related to homologs from *Bacteroidetes/Chlorobi* bacteria, rather than cyanobacteria (Dornfeld et al., 2013).

Dornfeld et al. (2013) conducted a detailed biochemical characterization of PAT and ADT/PDT enzymes from *Chlorobium tepidum* (*Chlorobi*) and *Synechocystis* (cyanobacterium). *PAT* gene from plants was characterized for the first time in our laboratory as a prokaryotic-type aspartate aminotransferase from *Pinus pinaster* belonging to subfamily Ib (de la Torre et al., 2006). Interestingly, *C. tepidum* aspartate aminotransferase Ib exhibits biochemical properties that are more similar to the plant counterparts, as an strong catalytic efficiency towards prephenate,

than to the *Synechocystis* homolog, which exhibits an almost undetectable PAT activity (Graindorge et al., 2014) Furthermore, the lack of an efficient PAT activity in the *Synechocystis* aspartate aminotransferase Ib enzyme is consistent with the existence of BCAT-type (from branched chain aminotransferases) enzymes with strong PAT activity and broad substrate specificity in cyanobacteria (Graindorge et al., 2014). According to Dornfeld et al. (2013), *C. tepidum* ADT/PDT enzyme showed 10-fold higher catalytic efficiency for arogenate than for prephenate, despite the fact that it was previously described as a PDT enzyme (Tan et al., 2008). In contrast with *Bacteroidetes/Chlorobi* ADT/PDTs, *Synechocystis* and the other extant cyanobacteria lack ADT activity (Stenmark et al., 1974; Hall et al., 1982) and possess the bifunctional P-protein encoded by the *PheA* gene, which presents both CM and PDT activity in the same catalytic domain. P-proteins are a common feature among the majority of bacterias, included *Escherichia coli* (Zhang et al., 2000), but seems to be absent in the *Bacteroidetes/Chlorobi* lineage. In conclusion, Phe pathway genes found in the eukaryotic descendants of primary endosymbiosis are more closely related, on the basis of either phylogenetic and biochemical analysis, with those found in the *Bacteroidetes/Chlorobi*.

Another example of plant prokaryotic-type genes, without cyanobacterial origin, are members of the menaquinone/phylloquinone biosynthesis gene cluster, the *men* cluster. This cluster is widely conserved between prokaryotes and land plants and encodes the necessary enzymes for the biosynthesis of vitamin K (Heimann and Katsaros, 2013). In plants, *men* genes are located in the nuclear genome and encode plastid-locatlized proteins. Many of the *men* genes of algae and plants are most closely related to sequences from non-cyanobacterial prokaryotes. While some members of the cluster, like *MenA* and *MenG*, are cyanobacterial-type,

MenE is of deltaproteobacterial origin (Gross et al., 2008), showing a pattern characteristic of evolutionary chimerism. Hence, Gross et al. (2008) suggested that bacterial ancestor of primary plastids could have been a chimeric organism. Similar observations could be inferred from a comprehensive phylogenetic analysis of the KsgA/Dim1 protein family of rRNA adenine dimethyltransferases (rAD), another kind of well conserved gene family with multiple functions. It has been reported that plastid-targeted rADs are more closely related to *Chlamydiae* in green algae and plants, while red algal sequences cluster better with the *Bacteroidetes/Chlorobi* sequences (Park et al., 2009).

In order to explain the phylogenetic relationships found in plants Phe biosynthesis genes, Dornfeld et al. (2013) proposed that these genes were acquired by the eukaryotic ancestor from a bacterial lineage that gave rise to a sister phyla *Bacteroidetes/Chlorobi*. On the other hand, Gross et al. (2008) have proposed that plastids were acquired from a chimeric ancestral cyanobacterium. This cyanobacterium could have possess a set of biosynthesis cluster genes closely related to those found in another bacterial groups, as *Bacteroidetes/Chlorobi*. After primary endosymbiosis, while the ancestral genes were retained by the plastid progenitor, modern-day cyanobacterial lineages acquired gene clusters from distantly related organisms. This could explain the separation between plastids and cyanobacteria when certain modern gene families are used for phylogenetic analysis. Unfortunately, it is difficult to discriminate between both alternatives as far as we have a really limited knowledge about the kind of eukaryotic cell and cyanobacteria that participated in the primary endosymbiosis. Anyway, these unusual phylogenies point to a complex evolutionary history with more bacterial groups involved than previously assumed for the “green lineage”.

2. Land adaptation and evolution of Phe metabolism

In the early steps of plant land colonization, the acquisition of the novel ability to synthesize progressively more abundant and diverse phenolic compounds from the aromatic amino acids played an essential role in their adaptation to new environments. After starting to leave their original aquatic habitat, the ancestors of first land plants got to deal with the underlying requirements of a solid ground environment: hydric stress, ultraviolet radiation, different microbial communities and forces of *gravity.* This transition from aquatic to terrestrial habitats promoted the emergence of key adaptations, as specialized metabolic pathways, like plants secondary metabolism.

Plants phenolic metabolism is one of the major innovations that probably occurred in the early stages of adaptation to land environment, as shown by its existence in the whole plant kingdom. The acquisition of the ability to deaminate the amino acid Phe by the enzyme PAL and to hydroxylate its aromatic ring probably led to the accumulation of the first phenylpropanoids, which probably resulted quite useful as UV-protectant compounds (Weng and Chapple, 2010). Extant bryophytes and liverworts, which are among the most primitive land plants, are rich in phenolic compounds (Umezawa, 2003; Yu et al., 2014), mainly solubles, supporting this hypothesis about the origin of phenolic metabolism in land plants.

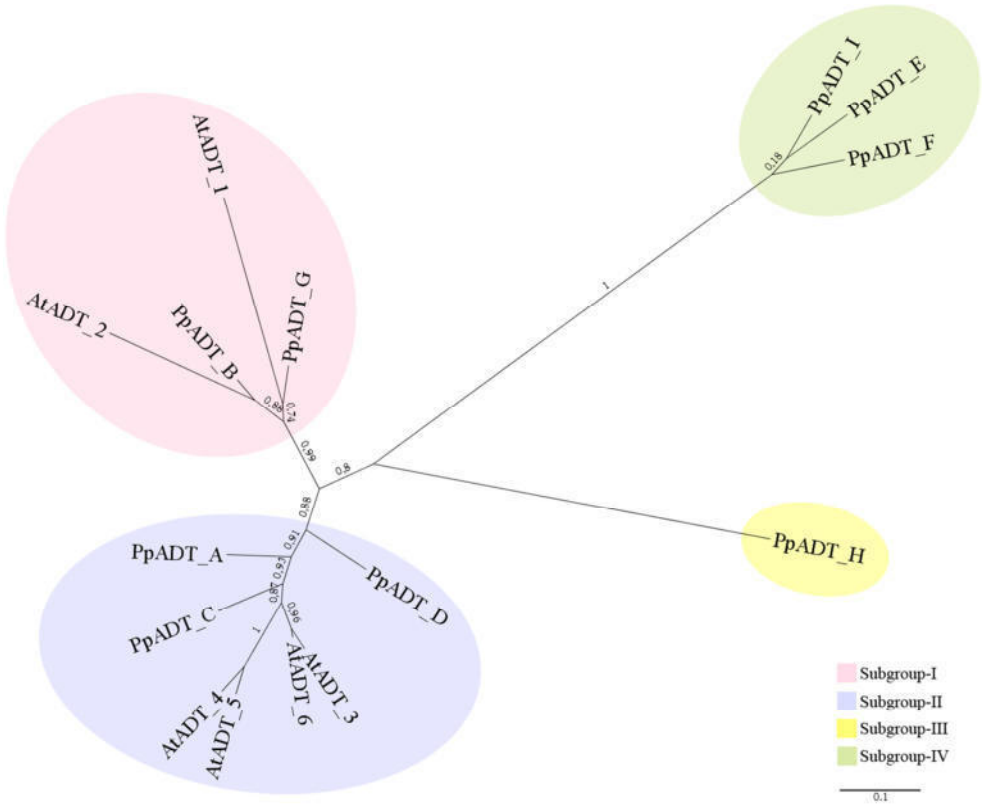
Thus, as a direct consequence of the emergence of new metabolic pathways adaptively advantageous, the metabolic demand of tyrosine and particularly Phe probably rise dramatically at the same time that first land plants were becoming more independent from their original aquatic medium. Phylogenetic analysis of prephenate pathway genes reveals that,

while PAT seems to have been preserved as a single copy-gene from algae to plants, the family of ADT/PDT genes have expanded very significantly in the land plants lineage. This growing number of ADT/PDT genes can be already observed in the most primitive land plants, as shown in the maximum-likelihood phylogeny in the figure 3 from paper 2. As can be seen in this figure, while the variety of analyzed algae exclusively possess a single-copy ADT/PDT gene, *Marchantia polymorpha* (liverwort), *Physcomytrella patens* (moss), and *Sphagnum phallax* (moss) have 2, 4 and 4 copies, respectively. In contrast with algae, these primitive plants already possess PAL activity and are able to synthesize phenolic compounds from Phe, suggesting that phenolic metabolism probably co-occurred with an expansion of ADT/PDT gene family. This interesting observation can be addressed as the adaptive response to the emergence of a new metabolic fate for this aromatic amino acid. Whereas algae still use Phe only for protein biosynthesis, the emergence of PAL and the accumulation of the first simple phenols put selective pressure on plants capability of synthesizing larger amounts of this amino acid. Duplication of genes coding for Phe pathway enzymes could be the first response to this requirement, and the resulting enhancement in the accumulation of UV-protectant compounds probably became a key factor for survival.

A comparative analysis performed in petunia petals showed that PAT activity is more than 3 orders of magnitude higher than ADT activity (Maeda and Dudareva, 2012), suggesting that ADT catalyzes a rate-limiting step in Phe biosynthesis. Moreover, given the pivotal branch point position of ADT and its potential ability to regulate the flux through the pathway (for instance, in response to Phe concentration by the ACT regulatory domain: see section I.7), it is not surprising that there had been an increase in the number of ADT/PDT genes in response to a growing Phe demand.

Moreover, the emergence of vascular plants and lignin metabolism surely led to an even larger demand of this amino acid. In vascular plants, lignin biosynthesis involves a massive carbon flux comprising more than 30% of photosynthetically fixed CO_2 (Boerjan et al., 2003). In this way, the number of ADT genes trends to be larger in vascular plants such as *Selaginella moellendorffii* (4 copies), *Arabidopsis thaliana* (6 copies), *Glycine max* (5 copies), *Brachipodium distachyon* (6 copies) or *Zea mays* (5 copies), when compared to non-vascular plants (mosses and liverworts). Successive rounds of ADT gene duplication were necessary to synthesize large amounts of Phe, especially when considering that Phe demand raises dramatically in certain developmental stages and tissues, more particularly during the vascular bundles formation. At the same time, the necessary substrate for an increasing isoform-related specialization and regulation was provided, making possible to coordinate Phe biosynthesis for protein formation and phenylpropanoids biosynthesis. This particular issue will be further discussed in the following section: 5. Different ADT isoforms for different phenylpropanoids.

As demonstrated in this thesis manuscript, ADT gene family duplication is particularly relevant in conifers. As woody plants, conifers divert large amounts of carbon into lignin biosynthesis during their long life cycles. The high number of ADT isoforms found in *Pinus pinaster*, with at least 9 unique transcripts, but also in other gimnosperms such as *Pinus taeda* and *Picea glauca*, remarks the importance of phenolic metabolism for this group of plants. We hypothesize that a complex secondary metabolism, a feature of conifer trees, is co-rrelated to a larger and more diverse ADT/PDT family (**Figure V.1**). This particular issue has been addressed in paper 5: *Biosynthesis and Metabolic Fate of Phenylalanine in Conifers*. Conifer's larger ADT gene families are also consistent with the high frequency of



V.1. Maximum likelihood phylogeny of *Pinus pinaster* and *Arabidopsis thaliana* ADT proteins. Larger and more diverse *ADTs* families are a common feature in conifers. *A. thaliana* (At) presents 6 ADTs distributed among 2 clusters, termed as Subgroup-I (red; bifunctional ADT/PDT proteins) and Subgroup-II (blue; monofunctional ADT), whereas *P. pinaster* (Pp) has at least 9 ADTs, distributed among 4 different clusters, with Subgroup-III (yellow) and Subgroup-IV (green) as characteristic of conifers. This could be corelated to the upstream regulation of a larger and more diverse secondary metabolism. *A. thaliana* ADT sequences were obtained from TAIR. *P. pinaster* sequences were inferred by BLASTP in SustainPine DB. Sequences were aligned using MUSCLE and the resulting alignment was curated by Gblocks. The tree was determined using PhyML with maximum likelihood set as optimality criterion, and subsequently edited as unrooted tree using FigTree v1.4.2. Bootstrapping confidence values were calculated with 1000 replicates and are expressed as a decimal.

gene and pseudogene duplications in their genomes, particularly for genes of secondary/specialized metabolism. This conclusion is in agreement with the findings of Warren et al. (2015): a comprehensive *Picea glauca* genome annotation of genes putatively related to terpenoids and phenolics metabolism uncovered an exceptional diversity of genes involved in such pathways. This characteristic could be probably associated with their undeniable adaptive success, both over different geological eras and different environments, still dominating the present-day forest biomes in temperate climates (Farjon, 2014).

3. The "phenylpyruvate branch" could operate as an alternative route for phenylalanine biosynthesis

The arogenate pathway is considered the major route for biosynthesis of Phe in plants (Maeda *et al*., 2010, 2011). Nevertheless, PDT activity in plants was previously reported under particular conditions (Warpeha *et al*., 2008). The molecular and kinetic characterization of the ADT family from *Arabidopsis thaliana* showed that ADT1, ADT2 and ADT6 can also work as PDT *in vitro*, although with lower catalytic efficiencies (Cho et al., 2007). In addition, *ADT1* and *ADT2* can rescue the Phe auxotrophic phenotype caused by the deletion of the endogenous PDT activity in the yeast mutant *pha2*, suggesting that these enzymes exhibit PDT activity in this heterologous system (Bross *et al*., 2011).

The potential existence of a functional phenylpyruvate pathway in plants was genetically observed for the first time in *Petunia hybrida* petals suppressed for *ADT* expression (Maeda et al., 2010). In this model, *P. hybrida* ADT1 RNAi-suppression significantly reduced ADT activity, levels of Phe, and its derived phenylpropanoid and benzenoid volatiles emission. Interestingly, transgenic petals supplied with exogenous shikimate showed prephenate and arogenate accumulation, along with higher levels of phenylpyruvate and a partial restoration of the reduced Phe pool, suggesting that Phe biosynthesis via phenylpyruvate route can operate *in planta*.

Paper 1, *Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism*, present new evidences about the potential existence of a functional PDT pathway in plants. *Nicotiana benthamiana* plants silenced for *PAT* showed significant

increase in the levels of Phe of approximately 2-fold compared with control plants. Phe accumulation in these plants rises even dramatically when *PAT* is co-repressed along with *Asp5*, reaching an accumulation of 8-fold compared with control. This finding confirms the previous work of Maeda et al. (2011), in which a ~70% reduction of PAT activity in the petals of *P. hybrida* RNAi plants led to a relative slight reduction of around 20% in Phe levels. However, the authors reported that they were not able to detect significant levels of phenylpyruvate or PDT activity, while at the same time expression levels of *P. hybrida ADT2* and *ADT3*, which can also convert prephenate into phenylpyruvate, were not altered in silenced petals. In contrast, the results presented in this thesis manuscript show that *N. benthamiana NbPDT1* and *NbPDT2* are clearly overexpressed in silenced tobacco leaves, either in *PAT* or *PAT/Asp5* silenced plants. We have also found that *NbPPAT*, the homolog of the *Arabidopsis* gene *At5g36160*, which encodes an aromatic amino acid aminotransferase that is able to convert phenylpyruvate into Phe (Prabhu and Hudson, 2010), is around 2- to 3-fold overexpressed in *PAT* and *PAT/Asp5* RNAi plants. Moreover, the downregulation of *P. hybrida PAT* in the *ADT1* RNAi background also rescues Phe and Phe-derived volatile levels (Yoo et al., 2013).

The subcellular localization of the “phenylpyruvate branch” enzymes deserves special attention, as far as a major evidence points to a cytosolic localization for at least some of the phenylpyruvate pathway candidate enzymes. The phenylpyruvate branch model proposed by Yoo *et al*. (2013) involves a cytosolic phenylpyruvate-aminotransferase catalyzing the interconversion of phenylpyruvate and Tyr to Phe and 4-hydroxyphenylpyruvate, respectively, whereas PDT activity is considered to take place within the plant plastid. Hence, phenylpyruvate should be exported from the plastids towards cytosol, where the conversion into Phe

occurs (**Figure IV.2**). Consistent with this model, the results reported in paper 1 demonstrate the existence in *Nicotiana benthamiana* of a cytosolic aromatic amino acid aminotransferase which is overexpressed as part of the genetic re-arrangement induced by *PAT* silencing.

Previous subcellular localization studies in *Arabidopsis thaliana* had placed all 6 ADT within the chloroplast of protoplasts and light-grown rosette leaves (Rippert et al., 2009). Nevertheless, it has been reported the presence of ADT3 in the cytosolic fraction when young etiolated seedlings

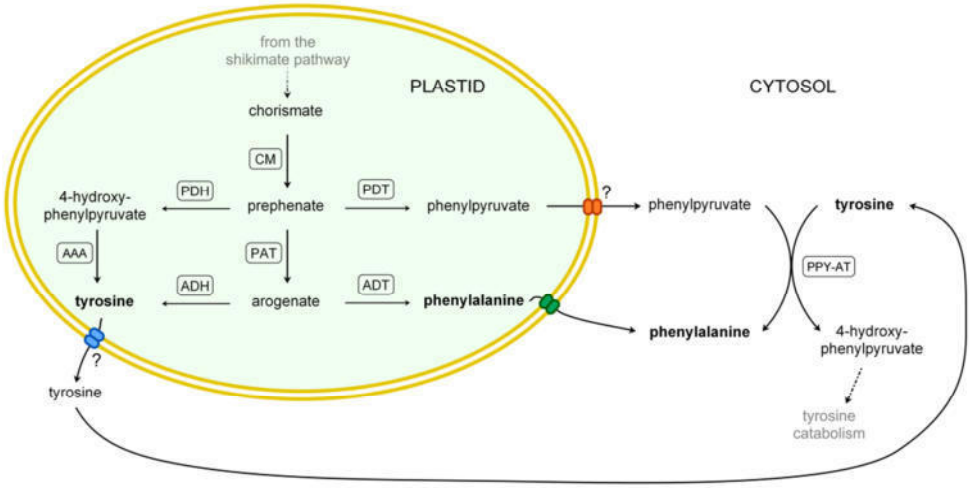


Figure V.2. A new model for phenylalanine biosynthesis in plants. As proposed by Yoo et al. (2013), this includes a cytosolic step by phenylpyruvate aminotransferase activity (PPY-AT), which is able to synthesize phenylalanine from phenylpyruvate, using tyrosine as amino donor. 4-hydroxyphenylpyruvate produced by PPY-AT reaction has been suggested to be linked to tyrosine catabolism. Question marks indicate transporters which still remain unidentified. Enzymes abbreviation: CM, chorismate mutase; PDT, prephenate dehydratase; PAT, prephenate aminotransferase; ADT, arogenate dehydratase; PDH, prephenate dehydrogenase; ADH, arogenate dehydrogenase; AAA, aromatic amino acid aminotransferase.

were analyzed for this purpose (Warpeha et al., 2008). Very recently, Para et al. (2016) have reported further evidences about the cytosolic localization of ADT3 in etiolated *Arabidopsis* seedlings. This novel report of a cytosolic localization for an ADT enzyme is accompanied by the presence in the cytosol of a chorismate mutase isoform, which acts at the first committed step in the prephenate pathway, suggesting the possibility of a full set of Phe biosynthetic enzymes, from the phenylpyruvate branch, with cytosolic localization (Para et al., 2016). Nevertheless, the subcellular localization reported for ADT3 in etiolated seedlings of *Arabidopsis* should be confirmed for the other ADT isoforms, more particularly ADT1, ADT2 or ADT6, which can act as PDT, whereas ADT3 does not have that function. If the conversion of chorismate into Phe via PDT could happen entirely or not in the cytosol will require further investigation.

4. The structure of the ACT domain is related to PDT activity.

As explained in the previous section, the redirection of carbon flux from the "arogenate branch" of prephenate pathway towards the formation of phenylpyruvate could make possible to withstand the inability to synthesize Phe when PAT and/or ADT are mainly blocked. Furthermore, it lay out the interesting searching for plant genes participating in the microbial-like phenylpyruvate branch.

Paper 2, Identification of a small protein domain present in all plant lineages that confers high prephenate dehydratase activity, site directed mutagenesis, functional complementation in yeast and *in vitro* characterization of purified proteins have allowed to identify a region of 22 amino acids within the C-terminal ACT regulatory domain that seems to be characteristic of bifunctional ADT/PDT enzymes. The name of PAC, from prephenate dehydratase activity conferring, has been proposed for this region. The presence of this 22-amino-acid-length motif confers comparatively high affinity towards prephenate and a high ratio K_{cat}/K_m for the PDT reaction. The importance of the PAC domain for an efficient PDT activity was confirmed by its ability to reduce the apparent K_m towards the substrate when was introduced in a protein with low PDT activity, such as PpADT-C. Within the PAC domain a single residue, Ala-314, is essential for conferring a low apparent K_m for prephenate, and for enhancing V_{max} as well. The position of the residue Ala-314 and the PAC domain itself, located between two conserved motifs previously described as involved in the allosteric binding of Phe, suggest a link between the Phe-mediated allosteric regulation (a characteristic of the ACT domains) and PDT activity. In fact, these results demonstrate that PDT activity is determined mainly in the regulatory domain, and not in the catalytic domain of the enzyme.

However, the mechanism whereby the ACT regulatory domain could determine PDT activity is unclear. Although deeper structural and crystallographic analysis will be required to adequately address this question, in paper 3 we have proposed the following preliminary hypothesis: prephenate, in addition to be substrate of the enzymatic reaction, could operate as an allosteric PDT-specific activator by itself. This activation would be mediated by the recognition of prephenate in the PAC domain followed by a conformational change in the enzyme, resembling the allosteric recognition of Phe in the same site (although with different effects). This hypothesis could explain why PAC-containing proteins can work efficiently as PDT *in vitro* with the sole presence of the substrate of the reaction, whereas when the PAC domain is absent or removed, they cannot, making possible to explain how the same catalytic site can be apparently monofunctional or bifunctional, as observed for PpADT-C.

The presence of the PAC domain along the entire Archaeplastida domain suggests that the existence of bifunctional ADT/PDT enzymes, far from being restricted to Arabidopsis and other flowering plants, is a common feature of all land plants and their more closely related ancestors. Based on this observation, we believe that at the origin and early evolution of land plants the PAC domain and therefore the PDT-competence was retained by ADTs as a primitive character. Yet, monofunctional ADTs would be a derived character or synapomorphy distinguishing emerging vascular plants, surely related to the increasing need for gene specialization in secondary metabolism. On the other hand, the preservation of PDT activity along all land plants leads the even more exciting question of why PDT activity and phenylpyruvate branch, most likely a primitive character, has been preserved for millions of years of plants evolution. Taken together this suggests an essential role for PDT throughout the life cycle of plants.

5. Different ADT isoforms for different phenylpropanoids

How vascular plants switch carbon flux differentially for protein biosynthesis or phenylpropanoid metabolism is not certainly well understood. Due to its branching position in the prephenate pathway and their putative role as rate-limiting enzymes in Phe biosynthesis, ADTs could be good candidates for a differential regulation of Phe use into primary or secondary metabolism (Corea et al., 2012a; 2012b). Previous reports (Cho et al., 2007; Rippert et al., 2009) suggested that all six *Arabidopsis* ADTs are expressed more or less ubiquitously in various tissues with differential intensity, evidencing a partially redundant role in plant growth and development. Nevertheless, in the last few years it has been demonstrated that it is also possible to observe certain degree of ADT isoform specialization, corelated to the biosynthesis of particular secondary metabolites and/or developmental and adaptive responses.

For example, very recently it has been reported a differential contribution of *Arabidopsis* ADT isoforms towards sucrose-induced anthocyanin biosynthesis (Chen et al., 2016). These authors have found that all six ADT isoforms function redundantly but differentially in regulating anthocyanin biosynthesis, pointing out that ADT2 contributes mostly to anthocyanin biosynthesis among all ADT isoforms. Then, the down-regulation of *ADT2* expression by artificial microRNA interference in *Arabidopsis* transgenic plants led to a reduction in *ADT2* mRNA levels of about 90% and a reduction of around 65% of WT in anthocyanin content. At the same time, ectopic expression of *ADT4* and *ADT5*, but not others ADTs, led to an over-accumulation of anthocyanin with levels over 60% to 120% higher than WT. As proposed by these authors, this may result from ADT4 and ADT5 lack of sensitivity to feedback-inhibition by Phe. Nevertheless, the results reported

supports that anthocyanin biosynthesis is rate-limited by Phe availability in the model studied and, at the same time, certain ADT isoforms (ADT2) have a major role in anthocyanin biosynthesis in response to sucrose treatment.

On the other hand, ADT3 has been suggested as essential for UV-protectant pigments biosynthesis during early stages of seedling development (Warpeha et al., 2008). Phe production in etiolated *Arabidopsis* seedlings lacking a functional ADT3 is strongly affected and, in consequence, phenylpropanoid production, not being able to accumulate UV-protectant compounds in response to UV treatment. Interestingly, Phe addition to the growth media restores WT phenotype in the *adt3* mutants, confirming the essential role of ADT3 for Phe biosynthesis. It the last few months, ADT3 has been also linked to cell's ability to balance and modulate the accumulation of reactive oxygen species (ROS), via phenolic compounds that take part in non-enzymatic mechanisms to scavenge ROS under several stress conditions (Para et al., 2016).

With respect to secondary cell wall formation, it has been reported in *Arabidopsis* that gene silencing of individual ADT isoenzymes and/or their combination differentially affect cell wall formation (Corea et al., 2012a). A significant reduction in stem lengths to a range of 67-85% (compared to WT) can be observed when *ADT5* expression is abolished in combination with *ADT4*, *ADT1/4*, *ADT3/4* and *ADT3/4/6*. The *adt4/5* mutant and its different combinations also displayed a small and dwarf phenotype. Similarly, the *adt5* KO line showed the largest reductions, of any single ADT mutant analyzed, in stem lignin contents and thioacydolitic monomer release, relative to WT. These results are consistent with previous findings showing that *ADT4* and *ADT5* had the highest expression levels in WT

stems, whereas the remaining *ADT* genes are expressed comparatively at lower levels in this tissue (Rippert et al., 2009). Taken together, these data suggest that both ADT4 and ADT5 could have a dominant role during stem lignification in *Arabidopsis*. Moreover, it has been found that ADT mRNA silencing led to reduced Phe pools only in lignin-deficient lines, and in tissues and at time points where lignin biosynthesis use to be constitutively highly active (Corea et al., 2012b). Interestingly, this effect is not evident across all ADT KO lines. These results indicate that ADT isoforms are functional redundancy, despite ADT4 and ADT5 play a major role in *Arabidopsis* growth.

Besides *Arabidopsis* as model organism, the biological role of ADT has been also studied in *Petunia hybrida*, in association with floral volatiles formation (Maeda et al., 2010). Through an RNAi-approach, *ADT1* from *P. hybrida* was specifically silenced in petals, as this gene was found to be the most expressed in the studied organ. Consequently, the ADT1-RNAi petunia petals showed a drastic decrease in Phe accumulation by 75% to 82% from control flowers levels, and up to a 70% reduction in total emission of phenylpropanoid and benzenoid compounds relative to controls. Similarly to what can be inferred from similar reports, the resulting phenotype can be partially rescued by feeding transgenic flowers with exogenous Phe. Since the activity levels of various enzymes of scents biosynthesis were also measured, including PAL, with no alterations detected in the silenced petals, it is plausible that Phe level is one of the major factors determining the volatile emissions in petunia flowers.

The differential function of ADTs has received certain degree of attention in short-living herbaceous plants particularly *Arabidopsis* and *Petunia*, however, we lack any information on this subject in woody plants,

notwithstanding that they undergo much stronger lignification processes, with larger amounts of carbon diverted into lignin biosynthesis. Woody plants form reaction wood, a specialized secondary xylem, in response a non-vertical orientation, helping the stem to recover the correct growth direction (Du et al., 2007). In gymnosperms, reaction wood is called compression wood and it develops on the underside of branches and inclined stems. Tracheids in compression wood present thickered cell walls, with a higher lignin/cellulose ratio than normal wood. Therefore, compression wood constitutes an exceptional model for the study of metabolism and gene regulatory networks that underlay xylogenesis in woody plants (Villalobos et al., 2012). As shown in paper 4, *PpADT-A* and *PpADT-D* are overexpressed in compression wood when compared to opposite wood, which is known to have lower lignin content and is more similar to normal wood (Yeh et al., 2005). It also has been demonstrated that *PpADT-A* and *PpADT-D* expression is under control of PpMYB-8, a R2R3-MYB transcription factor. Together with previous reports (Craven-Bartle et al., 2013) the presented results reveal the important role of PpMYB-8 in controlling Phe biosynthesis during lignification. Therefore, the established link between PpADT-A/PpADT-D and compression wood formation becomes the first contribution on the study of ADTs and growth in a woody plant, providing new evidences about how Phe availability could act as a limiting factor for plant growth.

6. Phe biosynthesis and female gametophyte development.

Plant life cycle is constituted by the alternation of two strongly differentiated generations: the sporophyte, a diploid multicellular stage which is developed from the zygote, and the gametophyte, an haploid stage developed from spores and responsible for the gametes formation and sexual reproduction (Yang et al., 2010). Whereas animal gametes are formed directly after meiosis, plant gametes are produced only after growth of the multicellular haploid gametophyte. During the evolution of plants, this second phase of the life cycle has progressively suffered a reduction in size, which is particularly extreme in flowering plants. In these plants, the gametophytic generation is short-lived and its development takes place within the parental (sporophyte) tissues.

Most flowering plants presents a characteristic seven-celled female gametophyte, also called embryo sac, which represents one of the major defining characteristics of this group of plants. Flowering plant's female gametophytes are small enough to be packaged within an ovary, and generate two gametes that undergo double-fertilization (Sundaresan and Alandete-Saz, 2010). These two female gametes are called egg cell and central cell, and get fertilized by one sperm cell each one, which are transferred from the pollen grain. Whereas the fertilized egg cell will develop the embryo, the adjacent central cell proliferates to generate the endosperm, the seed's nutrient reservoir (Tekleyohans et al., 2016).

Despite previous reports have shown the possibility to generate double knock-out *Arabidopsis* mutants for the remaining 5 ADTs and their different combinations (Corea et al., 2012a), in paper 6 we have demonstrated that it is not possible to obtain an homozygous *adt2* knockout mutant, presumably

due to its essential role in female gametophyte development. The female semi-sterility phenotype caused by the *adt2* mutation demonstrates an essential physiological role of this ADT isoform during the life cycle of plant. Very interestingly, the locus *At2g22250*, which corresponds with the single *PAT* gene in *Arabidopsis*, was described in a previous report as Maternal Effect Embryo Arrest 17 (MEE17), presenting embryonic arrest at the one-cell zygotic stage (Pagnussat et al., 2005). Given the fact that both genes (*PAT* and *ADT2*) are essential for *Arabidopsis* reproduction, it highlights the importance of Phe metabolism for embryo development, either for protein or phenolics biosynthesis. The complementation assays carried out in the heterozygous *adt2* plants have shown that either PHA2, the monofunctional PDT from *Saccharomyces cerevisiae*, as ADT3, one of the monofunctional ADT from *Arabidopsis*, can rescue the *adt2* female semi-sterility phenotype. It provides an strong evidence about that biosynthesis of Phe can occur either from arogenate or phenylpyruvate during the development of the female gametophyte of *Arabidopsis*, opening new issues about gene expression regulation and the cross-talk between Tyr and Phe metabolism in this organ.

Because of its inaccessibility and small size, little is known about the molecular basis of cell specification and differentiation in the female gametophyte. Nevertheless, genetic analysis tools have revealed mutants defective in almost all stages of female gametophyte development, and their analysis during the last decade have started to reveal novel features of the female gametophyte developmental program. Whereas previous investigations have established links between phenylpropanoid metabolism and male gametophyte development or pollen germination (Yang et al., 2001; Fallenberg and Vogt, 2015), our current knowledge about its female counterpart is comparatively very poor. However, a comprehensively

identification of genes required for female gametophyte development showed the existence of a certain number of genes implicated in secondary metabolism that are involved both in the nuclear division and nuclear fusion phases of the female gametogenesis (Pagnussat et al., 2005).

On the other hand, our understanding of how the sporophyte signals are controlling the developing gametophyte is still poor, although a variety of reports suggest that the sporophytic tissues have a key role in modulating the female gametophyte development (Figueiredo al Köhler, 2016). As mentioned before, plant phenolics also have been implicated in direct interactions with transport and signal transduction pathways, playing an important regulatory role in multiple physiological and developmental processes. Further efforts will be necessary to shed light on phenolics metabolism and its possible implication with female gametogenesis and the establishment of a molecular dialogue between the female gametophyte and the mother plant.

7. References

Abdallah a F, Salaminib F and Leisterb D (2000) *A prediction of the size and evolutionary origin of the proteome of chloroplasts of* Arabidopsis. Trends Plant Sci, 5(4):141-142.

Ball SG, Subtil A, Bhattacharya D, Moustafa A, Weber APM, et al. (2013) *Metabolic effectors secreted by bacterial pathogens; essential facilitators of plastid endosymbiosis?* Plant Cell, 25:7-21.

Baum D (2013) *The origin of primary plastids: a pas de deux or a ménage à trois?* Plant Cell, 25(1):4-6.

Boerjan W, Ralph J and Baucher M (2003) *Lignin biosynthesis*. Annu Rev Plant Biol, 54:519-546.

Bross CD, Corea OR, Kaldis A, Menassa R, Bernards MA and Kohalmi SE (2011) *Complementation of the* pha2 *yeast mutant suggests functional differences for arogenate dehydratases from* Arabidopsis thaliana. Plant Physiol Biochem, 49(8):882-890.

Chan CX, Gross J, Yoon HS and Bhattacharya D (2011) *Plastid origin and evolution: new models provide insights into old problems*. Plant Physiol, 155(4):1552-1560.

Chen Q, Man C, Li D, Tan H, Xie Y and Huang J (2016) *Arogenate Dehydratase Isoforms Differentially Regulate Anthocyanin Biosynthesis in* Arabidopsis thaliana. Mol Plant, 9(12):1609-1619.

Cho MH, Corea OR, Yang H, Bedgar DL, Laskar DD, et al. (2007) *Phenylalanine biosynthesis in* Arabidopsis thaliana. *Identification and characterization of arogenate dehydratases*. J Biol Chem. 282(42):30827-30835.

Corea OR, Ki C, Cardenas CL, Kim SJ, Brewer SE, et al. (2012a) *Arogenate dehydratase isoenzymes profoundly and differentially modulate carbon flux into lignins*. J Biol Chem, 287(14):11446-11459.

Corea OR, Bedgar DL, Davin LB and Lewis NG (2012b) *The arogenate dehydratase gene family: towards understanding differential regulation of carbon flux through phenylalanine into primary versus secondary metabolic pathways*. Phytochemistry, 82:22-37.

Craven-Bartle B, Pascual MB, Cánovas FM and Avila C (2013) *A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine*. Plant J, 74(5):755-766.

Daniell H, Lin C-S, Yu M and Chang W-J (2016) *Chloroplast genomes: diversity, evolution, and applications in genetic engineering*. Genome Biol, 17:134.

de la Torre F, De Santis L, Suárez MF, Crespillo R and Cánovas FM (2006) *Identification and functional analysis of a prokaryotic-type aspartate aminotransferase: implications for plant amino acid metabolism*. Plant J, 46(3):414-425.

Dornfeld C, Weisberg AJ, K C R, Dudareva N, Jelesko JG, Maeda HA (2014) *Phylobiochemical characterization of class-Ib aspartate/prephenate aminotransferases reveals evolution of the plant arogenate Phe pathway*. Plant Cell, 26(7):3101-3114.

Du S and Yamamoto F (2007) *An overview of the biology of reaction wood formation*. J Integr Plant Biol, 49:131-143.

Facchinelli F, Colleoni C, Ball SG and Weber AP (2013) *Chlamydia, cyanobiont, or host: who was on top in the ménage à trois?* Trends Plant Sci, 18(12):673-679.

Farjon, A (2014) *Conifers of the world: resources for conifer research. On-line database. Royal Botanical Gardens Kew, UK*. Available from: http://herbaria.plants.ox.ac.uk/bol/conifers

Fellenberg C and Vogt T (2015) *Evolutionarily conserved phenylpropanoid pattern on angiosperm pollen*. Trends Plant Sci, 20(4):212-218

Figueiredo DD and Köhler C (2016) *Bridging the generation gap: communication between maternal sporophyte, female gametophyte and fertilization products*. Curr Opin Plant Biol, 29:16-20.

Geoffrey I and McFadden (2001) *Chloroplast Origin and Integration.* Plant Physiology, 125(1):50-53.

Graindorge M, Giustini C, Kraut A, Moyet L, Curien G and Matringe M (2014) *Three different classes of aminotransferases evolved prephenate aminotransferase functionality in arogenate-competent microorganisms*. J Biol Chem, 289(6):3198-3208.

Gross J, Meurer J and Bhattacharya D (2008) *Evidence of a chimeric genome in the cyanobacterial ancestor of plastids*. BMC Evol Biol, 8:117.

Hall GC, Flick MB, Gherna RL and Jensen RA (1982). *Biochemical diversity for biosynthesis of aromatic amino acids among the cyanobacteria*. J Bacteriol, 149:65–78.

Heimann K and Katsaros C (2013) *Advances in Algal Cell Biology*. Series: Marine and Freshwater Botany. De Gruyter Ed., Berlin.

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, et al. (1996) *Sequence analysis of the genome of the unicellular cyanobacterium* Synechocystis sp. *strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions*. DNA Res, 3(3):109-36.

Kaneko T and Tabata S (1997) *Complete genome structure of the unicellular cyanobacterium* Synechocystis sp. PCC6803. Plant Cell Physiol, 38(11):1171-1176.

Keeling PJ (2010) *The endosymbiotic origin, diversification and fate of plastids.* Phil Trans R Soc B, 365:729-748.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, et al. (2010) *RNAi suppression of Arogenate Dehydratase 1 reveals that Phe is synthesized predominantly via the arogenate pathway in petunia petals*. Plant Cell, 22(3):832-849.

Maeda H, Yoo H and Dudareva N (2011) *Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate*. Nat Chem Biol, 7(1):19-21.

Maeda H and Dudareva N (2012) *The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants*. Annu Rev Plant Biol, 63:73-105.

McFadden GI (2001) *Chloroplast origin and integration*. Plant Physiol, 125(1):50-53.

Pagnussat GC, Yu H-J, Ngo QA, Sarojam Rajani S, Sevugan Mayalagu S, et al. (2005) *Genetic and molecular identification of genes required for female gametophyte development and function in* Arabidopsis. Development, 132:603-614.

Para A, Muhammad DS, Orozco-Nunnelly DA, Memishi R, Alvarez S, Naldrett MJ and Warpeha KM (2016) *The Dehydratase ADT3 Affects ROS Homeostasis and Cotyledon Development*. *Plant Physiol,* 172(2):1045-1060.

Park AK, Kim H and Jin HJ (2009) *Comprehensive phylogenetic analysis of evolutionarily conserved rRNA adenine dimethyltransferase suggests diverse bacterial contributions to the nucleus-encoded plastid proteome*. Mol Phylogenet Evol, 50(2):282-289.

Prabhu PR and Hudson AO (2010) *Identification and Partial Characterization of an L-Tyrosine Aminotransferase (TAT) from* Arabidopsis thaliana. Biochem Res Int, 2010:549572.

Rippert P, Puyaubert J, Grisollet D, Derrier L and Matringe M (2009) *Tyrosine and Phe are synthesized within the plastids in* Arabidopsis. Plant Physiol, 149:1251–1260.

Stenmark SL, Pierson DL, Jensen RA and Glover GI (1974) *Blue-green bacteria synthesize L-tyrosine by the pretyrosine pathway*. Nature, 247:290–292.

Sundaresan V and Alandete-Saez M (2010) *Pattern formation in miniature: the female gametophyte of flowering plants*. Development, 137:179-189

Tan K, Li H, Zhang R, Gu M, Clancy ST and Joachimiak A (2008) *Structures of open (R) and close (T) states of prephenate dehydratase (PDT)—implication of allosteric regulation by L- Phe*. J Struct Biol, 162:94–107.

Tekleyohans DG, Nakel T and Gross-Hardt R (2016) Patterning the female gametophyte of flowering plants. Plant Physiol, Preview. Published on December 5, 2016, as DOI:10.1104/pp.16.01472

Umezawa T (2003) *Diversity in lignin biosynthesis*. Phytochem Rev, 2(3):371-390.

Villalobos DP, Díaz-Moreno SM, El-Sayed S, Cañas RA, Osuna D, et al. (2012) *Reprogramming of gene expression during compression wood formation in pine: Coordinated modulation of S-adenosylmethionine, lignin and lignan related genes*. BMC Plant Biol, 12:100.

Warpeha KM, Gibbons J, Carol A, Slusser J, Tree R, Durham W and Kaufman LS (2008) *Adequate Phe synthesis mediated by G protein is critical for protection from UV radiation damage in young etiolated* Arabidopsis thaliana *seedlings*. Plant Cell Environ, 31(12):1756-1770.

Warren RL, Keeling CI, Yuen MM, Raymond A, Taylor GA, et al. (2015) *Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism*. Plant J, 83(2):189-212.

Weng JK and Chapple C (2010) *The origin and evolution of lignin biosynthesis*. New Phytol, 187(2):273-285.

Yang S, Sweetman JP, Amirsadeghi S, Barghchi M, Huttly AK, Chung WI and Twell D (2001) *Novel anther-specific myb genes from tobacco as putative regulators of Phe ammonia-lyase expression*. Plant Physiol, 126(4):1738-1753.

Yang W-C, Shi D-Q and Chen Y-H (2010) *Female Gametophyte Development in Flowering Plants*. Annu Rev Plant Biol, 61:89-108

Yeh T-F, Goldfarb B, Chang H-M, Peszlen I, Braun J and Kadla J (2005) *Comparison of morphological and chemical properties between juvenile wood and compression wood of loblolly pine*. Holzforschung, 59:669-674.

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, et al. (2013) *An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase*. Nat Commun, 4:2833.

Yu HN, Liu XY, Gao S, Han X-J, Cheng A-X and Lou H-X (2014) *Molecular cloning and functional characterization of a Phe ammonia-lyase from liverwort* Plagiochasma appendiculatum. Plant Cell Tiss Organ Cult, 117:265.

Zhang S, Wilson DB, Ganem B (2000) *Probing the catalytic mechanism of prephenate dehydratase by site-directed mutagenesis of the* Escherichia coli *P-protein dehydratase domain*. Biochemistry, 39:4722–4728.

VI. CONCLUSIONS

1. Bifunctional aspartate aminotransferase / prephenate aminotransferase is, for its critical role in amino acid homeostasis in the plastids, essential for the growth and development of plants.

2. A small protein domain, named PAC domain, confers high prephenate dehydratase activity to certain arogenate dehydratase enzymes from plants. This domain has been found to be preserved among all terrestrial plants and their algal ancestors, suggesting that PAC domain and prephenate dehydratase activity are primitive features preserved during the evolution of plants. The localization of the PAC domain in the regulatory region of the enzymes suggests a direct connection between allosteric regulation and prephenate dehydratase activity.

3. The silencing of prephenate aminotransferase in *N. benthamiana* activates the expression of genes for prephenate dehydratases and aromatic aminotransferases, suggesting that phenylalanine can be synthesized from phenylpyruvate in response to a deficiency in the arogenate-mediated biosynthesis of this amino acid.

4. The disruption of a member of the arogenate dehydratase gene family (*ADT2*) in *Arabidopsis thaliana* leads to a female semi-sterility phenotype. This is the single arogenate dehydratase that has been demonstrated to be essential in the life cycle of plants.

5. A yeast monofunctional prephenate dehydratase (PHA2) can rescue the female semi-sterility phenotype caused by ADT2 disruption, demonstrating that the biosynthesis of phenylalanine can occur via phenylpyruvate in this organ, and strongly suggest that the product of ADT2, a bifunctional arogenate/prephenate dehydratase enzyme is acting as prephenate dehydratase in this particular context.

6. Two members of the arogenate dehydratase gene family in *Pinus pinaster*, *PpADT-A* and *PpADT-D*, have been linked to compression wood formation and lignification and are transcriptionally regulated by PpMYB8, an activator of multiple genes through the lignin pathway. These results suggest that both enzymes are specialized isoforms involved in the supply of phenylalanine for lignin biosynthesis in maritime pine.