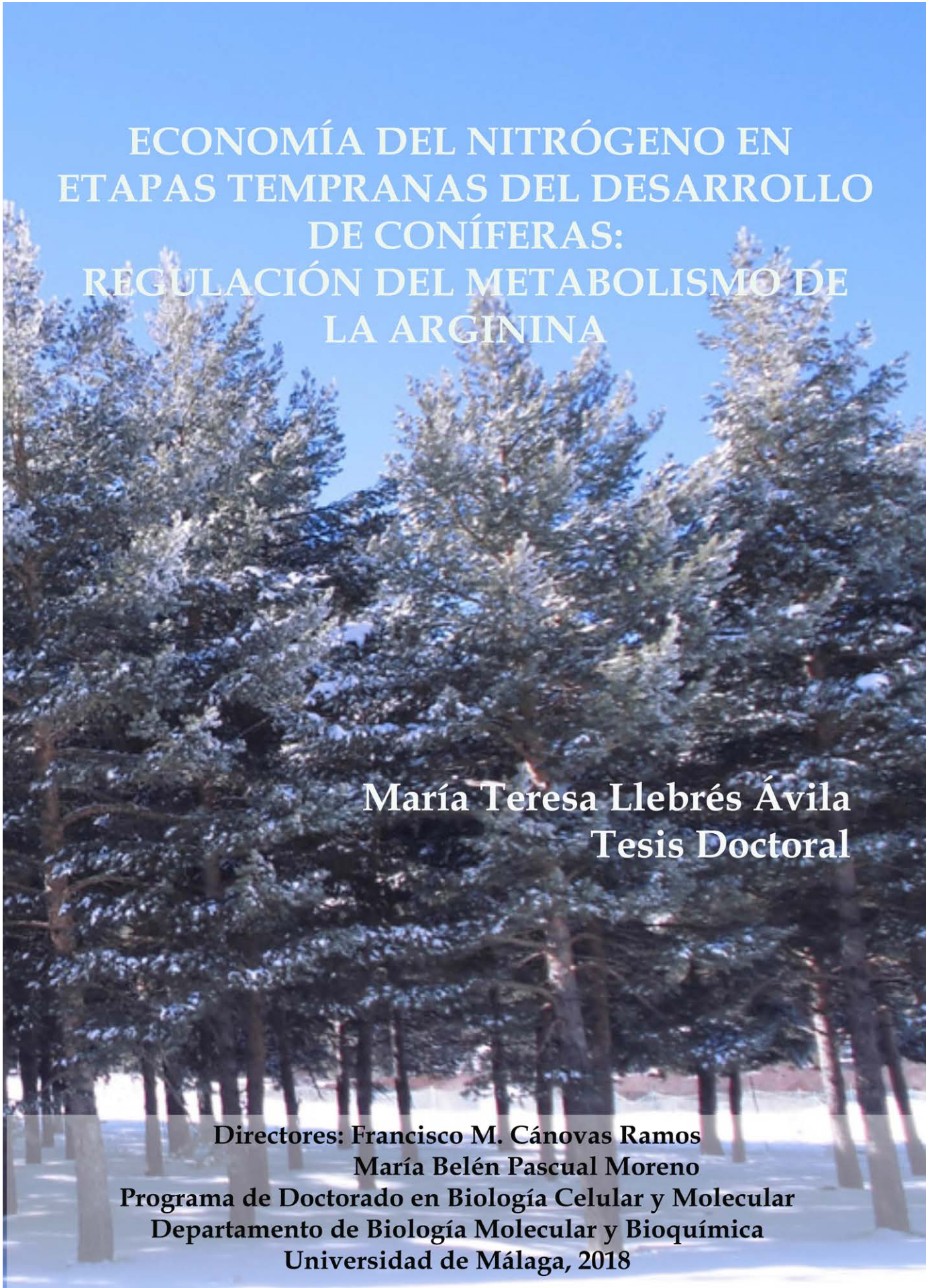


ECONOMÍA DEL NITRÓGENO EN ETAPAS TEMPRANAS DEL DESARROLLO DE CONÍFERAS: REGULACIÓN DEL METABOLISMO DE LA ARGININA

María Teresa Llebrés Ávila

Tesis Doctoral

Directores: Francisco M. Cánovas Ramos
María Belén Pascual Moreno
Programa de Doctorado en Biología Celular y Molecular
Departamento de Biología Molecular y Bioquímica
Universidad de Málaga, 2018

UNIVERSIDAD
DE MÁLAGA

AUTOR: María Teresa Llebrés Ávila

http://orcid.org/0000-0001-9320-4755

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga

ECONOMÍA DEL NITRÓGENO EN ETAPAS TEMPRANAS DEL DESARROLLO DE CONÍFERAS: REGULACIÓN DEL METABOLISMO DE LA ARGININA

MARIA TERESA LLEBRÉS ÁVILA

Tesis Doctoral

Programa de Doctorado en Biología Celular y Molecular
Departamento de Biología Molecular y Bioquímica
Universidad de Málaga, 2018

El Dr. Francisco Miguel Cánovas Ramos, Catedrático del Departamento de Biología Molecular y Bioquímica de la Facultad de Ciencias de la Universidad de Málaga y la Dra. María Belén Pascual Moreno, investigadora de la Universidad de Málaga.

CERTIFICAN:

Que Dña. María Teresa Llebrés Ávila, Licenciada en Ciencias del Mar, ha realizado en el Departamento de Biología Molecular y Bioquímica de la Facultad de Ciencias de la Universidad de Málaga y bajo nuestra dirección el trabajo de investigación recogido en la presente memoria de Tesis Doctoral que lleva por título: ECONOMÍA DEL NITRÓGENO EN ETAPAS TEMPRANAS DEL DESARROLLO DE CONÍFERAS: REGULACIÓN DEL METABOLISMO DE LA ARGININA.

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 Doctora.

Y para que así conste, en cumplimiento de las disposiciones legales vigentes, firmamos el presente certificado en Málaga, 31 de Enero de 2018.

Dr. Francisco M. Cánovas Ramos

Dra. María Belén Pascual Moreno

Los trabajos que se presentan en esta Tesis Doctoral han sido financiados por los proyectos:

- Ministerio de Economía y Competitividad (BIO2012-33797 y BIO2015-69285-R)

- Junta de Andalucía (BIO-474)

- ProCoGen project (FP7-KBBE-289841)

- Plant KBBE program, Scientific and Technological Cooperation in Plant Genome Research (PLE2009-0016)

- The French Ministry of Agriculture (DGAL, N°2014-352, QuaSeGraine project). The project also benefited from the technical support of the XYLOBIOTECH facility (ANR-10-EQPX-16 XYLOFOREST).

AGRADECIMIENTOS

Quisiera dar las gracias especialmente a mi director de Tesis Francisco Cánovas por la oportunidad que me ha dado al permitirme formar parte de su grupo de investigación durante estos años, por confiar en mí y por todo el apoyo que me ha brindado durante la realización de este trabajo.

La realización de esta Tesis no habría sido posible sin Concha Ávila, gracias por todo lo que has hecho por mí y por todo el tiempo que has dedicado en enseñarme, aconsejarme y ayudarme desde que empecé en el laboratorio.

Muchas gracias a mi co-directora de Tesis Belén Pascual por sus consejos, ayuda e implicación durante todo este tiempo, por estar siempre disponible y prestarme el apoyo que necesitaba en algunos momentos complicados, no podría haber tenido mejor jefa en el labo.

A Krystyna Klimaszewska con la que empecé mi formación en el mundo del cultivo in vitro y de la embriogénesis somática, a la que siempre recordaré con especial cariño, thank you so much Krystyna I will always be grateful for your help and dedication.

Quiero agradecérselo también a todos los que han formado parte de este grupo y que durante todo este tiempo me han ayudado tanto. Reme, Fernando, Jorge, Marina, Cañas, Vanessa, Canales y Blanca ¡muchas gracias por todo!.

ÍNDICE

I. INTRODUCCIÓN GENERAL

1. Las plantas, fuente de recursos naturales y biodiversidad

Las plantas son un componente vital de los ecosistemas y son fundamentales para el mantenimiento de la biodiversidad. Desempeñan un papel esencial en la producción de oxígeno y eliminación de emisiones de dióxido de carbono atmosférico, creación y estabilización de suelos, y protección de cuencas hidrográficas. También constituyen la fuente principal de recursos energéticos, de alimentos y de materias primas de uso industrial.

Según el informe de la Secretaría del Convenio sobre la Diversidad Biológica (2009) dos terceras partes de las especies de plantas del mundo se encuentran en peligro de extinción. Esto se debe al aumento de la población humana, la modificación de hábitats y deforestación, una explotación excesiva, la propagación de especies exóticas invasoras, la contaminación y el creciente impacto del cambio climático.

La población mundial actualmente supera los 7.400 millones de personas y el último informe de la FAO estima que en 2050 llegue a sobrepasar los 9.500 millones, lo que conlleva a un incremento dramático de la producción agrícola que pueda cubrir las necesidades alimenticias de una población en constante crecimiento.

El aumento de la población también supone un mayor requerimiento de productos forestales como madera, pulpa, papel y biocombustibles como fuente de energía. La creciente demanda de estos productos será difícil de suplir de manera sostenible y la biodiversidad se verá drásticamente reducida a menos que las nuevas tecnologías en desarrollo permitan una producción sostenible a mayor escala.

2. Los bosques de coníferas y su importancia económica

Las coníferas constituyen el grupo más importante de las plantas gimnospermas, por delante de cícadas, ginkgos y gnetales. Las gimnospermas aparecieron hace unos 350 millones de años y fueron la vegetación dominante durante el final del Paleozoico y comienzos del Mesozoico. Dentro de las coníferas se han identificado 615 especies

distribuidas en 70 géneros pertenecientes a 8 familias. La mayoría de estas especies pertenecen a tres grandes familias: *Pinaceae* con 232 especies, *Podocarpaceae* con 174 y *Cupresaceae* con 135 especies. Las otras cinco familias son *Araucariaceae, Cephalotaxaceae, Phyllocladaceae, Sciadopityaceae y Taxaceae* (Farjon, 2010).

Son plantas de porte leñoso, por lo que su crecimiento requiere mucha materia y energía, y su ciclo vital es lento, lo que provoca que sean plantas, en general, poco adaptadas para la colonización rápida de nuevos ambientes. Probablemente esta es una de las causas de su declive a partir del Cretácico, ya que fueron desplazadas por las angiospermas que comenzaron a diversificarse y ocuparon rápidamente muchos nichos ecológicos (Brodribb et al. 2012). La evolución de las angiospermas supuso una reducción en la diversidad de las coníferas, aunque las familias *Pinaceae* y *Cupressaceae* fueron las menos afectadas ya que continúan ocupando las mismas zonas secas y frías desde el Mesozoico.

Los bosques de coníferas representan aproximadamente un tercio de los bosques mundiales y se encuentran distribuidos formando parte de los paisajes más importantes del mundo (Figura 1). Son mayormente conocidos los extensos bosques del Norte de América, Europa y Asia donde las bajas temperaturas y los largos inviernos favorecen su desarrollo.

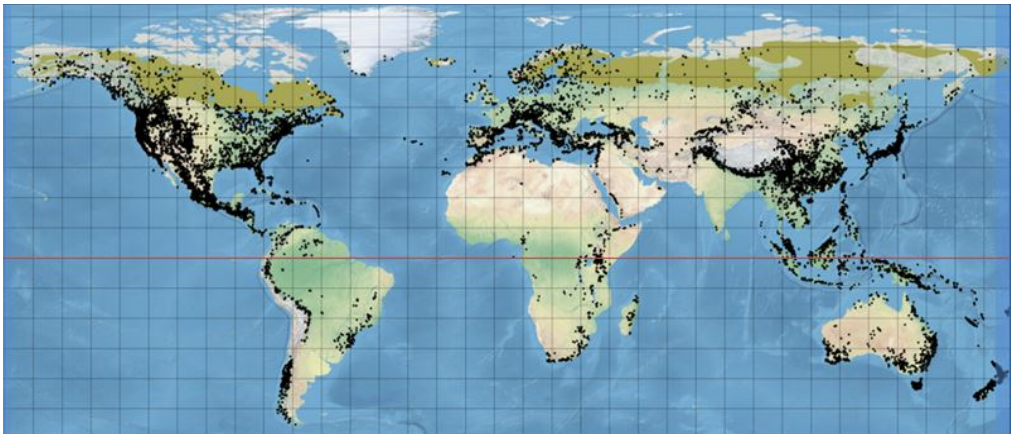


Figura 1. Mapa de distribución de los bosques de coníferas en el mundo. BRAHMS Online Copyright ©1985 - 2017 Department of Plant Sciences, University of Oxford.

La familia *Pinaceae* se distribuye principalmente en el Hemisferio Norte y no se ha identificado ningún registro fósil de esta familia al sur del Ecuador. Los pinos son unas de las coníferas más estudiadas y tienen la mayor diversidad tanto biológica como taxonómica dentro de su género. Su sistema vascular altamente resistente a la congelación y deshielo, combinado con una fisiología fotosintética bien adaptada, ha favorecido que las pináceas hayan ocupado siempre bosques propensos a la congelación (Turnbull et al. 1998; Becker 2000; Brodribb and Feild 2008).

Las coníferas tienen una gran importancia económica, pues representan un recurso de primer orden para la producción de papel y madera en todo el mundo. Hasta la fecha, la única fuente de madera son los bosques naturales, y se calcula que en 40 años habremos acabado con un porcentaje de entre un 20% a un 40% de la superficie forestal actual. Generalmente, los árboles tienen un crecimiento lento, por lo que existe la necesidad de buscar nuevas formas de obtener los productos demandados por la sociedad sin tener que acudir a los bosques nativos. Debido a esto, en los últimos 20 años se está investigando profusamente la biología de los árboles forestales, viéndose reflejado en un paulatino avance biotecnológico dirigido a la domesticación de dichas especies y por tanto, a su cultivo a gran escala, como ya sucede con otras muchas especies de interés agronómico.

Para llevar a cabo esta ardua tarea se deben obtener los conocimientos suficientes y desarrollar la biotecnología adecuada que permita disponer de una fuente de recursos para suplir la demanda existente.

3. Embriogénesis en plantas angiospermas y gimnospermas

Las plantas superiores se reproducen principalmente mediante reproducción sexual dando lugar a individuos que comparten características genéticas procedentes de ambos progenitores. Esto les proporciona la variabilidad génica necesaria para una continua adaptación al medio ambiente.

La mayoría de especies de interés agronómico y forestal pertenecen a dos grandes grupos, las angiospermas, representadas por más de 250.000

especies y las gimnospermas con más de 800 especies. El desarrollo de la semilla se produce de manera diferente en ambos grupos. Mientras que en angiospermas los órganos reproductores se localizan en flores verdaderas, las gimnospermas presentan las semillas al descubierto, y tienen los órganos reproductores alojados en conos o estróbilos. Sin embargo, presentan una serie de aspectos comunes durante el proceso reproductivo que comprende diferentes fases como la floración, polinización, fertilización, embriogénesis y maduración del embrión.

El desarrollo de las plantas puede englobarse en dos fases principales: primero tiene lugar la embriogénesis en sentido estricto, que comprende el desarrollo del embrión cigótico hasta el estado cotiledonario, y posteriormente se produce la maduración de la semilla y su germinación.

Durante la embriogénesis de la angiosperma modelo *Arabidopsis thaliana* el embrión cigótico se divide pasando por diferentes estados de desarrollo: globular, triangular, corazón y torpedo (Jürgens, 1994). Su desarrollo está gobernado desde el inicio por una polaridad celular característica de las angiospermas y una división asimétrica, donde participan tanto células masculinas como femeninas (Figura 2). La doble fertilización generará simultáneamente el embrión y el endospermo dando lugar a una semilla viable (Dodeman et al. 1997).

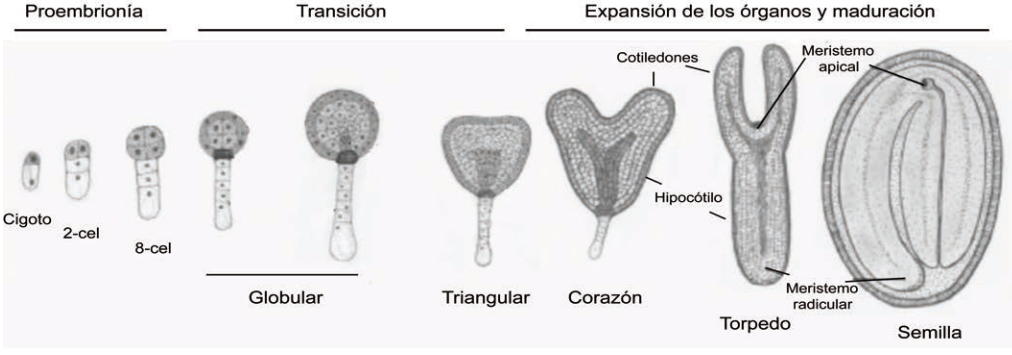


Figura 2. Representación esquemática del desarrollo embrionario en angiospermas (*A. thaliana)*. Adaptación de Von Arnold et al. (2002).

La embriogénesis en gimnospermas se divide en tres fases: proembriogénesis, que corresponde a las etapas previas a la elongación de los suspensores; embriogénesis temprana, antes del establecimiento del meristemo radicular; y embriogénesis tardía, donde se produce la diferenciación de tejidos y órganos en el embrión (Figura 3) (Singh, 1978).

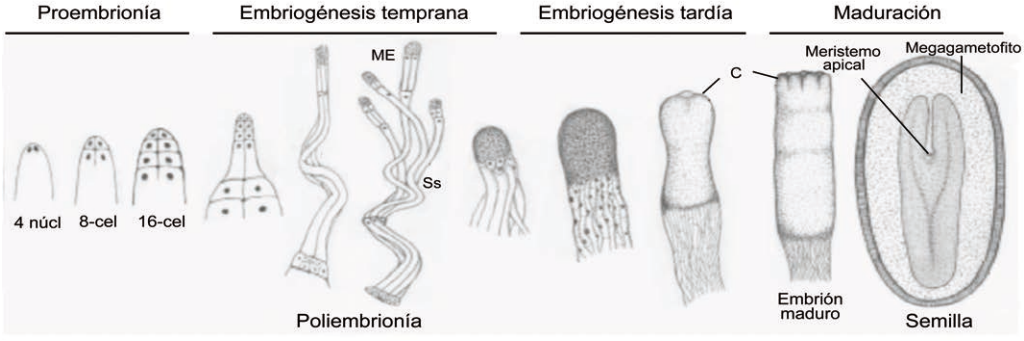


Figura 3. Representación esquemática del desarrollo embrionario en coníferas. Adaptación de Von Arnold et al. (2002). Abreviaturas: ME, Masa embrionaria; Ss, Suspensor secundario; C, Cotiledones.

En coníferas la embriogénesis se caracteriza por la presencia de poliembrionía que puede ser simple o de partición. La poliembrionía simple consiste en la fecundación del óvulo por diferentes granos de polen, produciéndose pre-embriones genéticamente diferentes, que degeneran durante la maduración llegando solo uno de ellos a completar el desarrollo. Sin embargo, en la poliembrionía de partición o *cleavage* los embriones son el resultado de la división de las células apicales de un pre-embrión en cuatro filas, cada una de las cuales puede desarrollarse y generar un embrión genéticamente idéntico (Figura 3) (Bucholz, 1926; Owens y Molder, 1984; Merino et al. 2016).

Se ha postulado que la poliembrionía es un carácter primitivo de las gimnospermas que actúa como mecanismo de selección tras la fecundación de los embriones en función de su vigor (Bucholz, 1926); y también como mecanismo adaptativo para la eliminación de los embriones procedentes de autopolinización (Sorensen, 1982).

Durante la embriogénesis tardía, el embrión dominante es impulsado por el suspensor hacia la región central del megagametofito donde se desarrollará mientras los otros degeneran. El suspensor le proporcionará los nutrientes y reguladores de crecimiento necesarios durante la embriogénesis degenerando una vez que el embrión haya completado su desarrollo (Umehara y Kamada, 2005; Merino et al. 2016).

La embriogénesis también puede reproducirse *in vitro,* de este modo las células precursoras del embrión se desarrollarán sin que tenga lugar el proceso de fertilización. La regeneración de plantas mediante el cultivo *in vitro* puede tener lugar por tres vías: mediante el desarrollo de yemas apicales o axilares, mediante la diferenciación de yemas o raíces adventicias, y mediante embriogénesis somática (Bonga et al. 2010).

4. Embriogénesis somática

El término embriogénesis somática (ES) hace referencia al proceso de formación del embrión a partir de células somáticas. Mediante este sistema de reproducción, se generan nuevos individuos genéticamente idénticos a la planta progenitora. La formación de embriones somáticos, puede tener lugar de manera directa o indirecta. La embriogénesis directa implica la existencia de células somáticas predeterminadas embriogénicamente, que sólo requieren unas condiciones favorables para el inicio del desarrollo embriogénico; en cambio, en la embriogénesis indirecta las células deben sufrir varias divisiones mitóticas en presencia de una auxina durante la inducción, dando lugar a la formación de un callo embriogénico o fase intermedia entre el explanto original y la aparición de embriones somáticos.

Durante el proceso de embriogénesis, los factores que influyen en el desarrollo del embrión maduro son numerosos. En el medio de cultivo, componentes como el agente gelificante, la fuente de carbono y de nitrógeno, y los factores hormonales son cruciales para su desarrollo. Durante este proceso también son importantes la temperatura y las condiciones lumínicas. La identificación de las condiciones óptimas de cultivo puede ser extremadamente difícil debido al amplio número de factores que contribuyen

a la conversión de los embriones somáticos en plantas, siendo muy dependiente de la especie y genotipo (Gutiérrez-Mora et al. 2012).

Los primeros árboles generados mediante técnicas de cultivo *in vitro* se obtuvieron en el inicio de los años 60 utilizando la vía organogénica, es decir, la inducción de yemas axilares o adventicias y su posterior enraizamiento. Sin embargo, durante estos últimos años, la industria forestal ha apostado claramente por la vía de la ES, debido a sus numerosas ventajas como su elevado potencial de multiplicación, la posibilidad de hacer crecer los cultivos en biorreactores, la opción de aplicar técnicas de encapsulación para fabricar semillas sintéticas o la posibilidad de utilizar los cultivos embriogénicos como dianas adecuadas para la transformación genética (Celestino et al. 2005; Thompson, 2015).

La ES está considerada como una de las vías principales para la mejora genética de especies vegetales, y en particular de coníferas. Esto se debe a que permite la propagación masiva de plantas de alto valor comercial mediante la producción de un gran número de embriones en un mismo estado de desarrollo y en relativamente poco tiempo. Además, se ha convertido en un sistema modelo para estudiar la embriogénesis a nivel molecular, celular y de tejidos (Willemsen y Scheres, 2004).

La regeneración de plantas mediante ES se puede considerar como el elemento central de la biotecnología vegetal, ya que da sentido agronómico a otras facetas de la biología celular y molecular. Disponer de una herramienta que permita regenerar plantas es imprescindible para que la clonación, la preservación de material vegetal o la transformación genética se conviertan en biotecnologías de aplicación (Celestino et al. 2005).

4.1. Embriogénesis somática en coníferas

Desde su descubrimiento, hace más de tres décadas en *Picea abies* y *Larix decidua* (Hakman et al. 1985; Chalupa et al. 1985) la ES en coníferas se ha desarrollado en numerosas especies, gran parte pertenecientes a la familia de las pináceas.

La propagación vegetativa de coníferas mediante ES tiene el potencial de ofrecer un suministro estable de plantas superiores para las plantaciones forestales, sin embargo, tras décadas de esfuerzos en investigación, la propagación eficiente de bosques de coníferas adultas no es aún una realidad. Dentro del género *Pinus* se han realizado estudios de ES con al menos 16 especies diferentes, aunque para muchas de ellas la eficiencia durante las fases de iniciación, maduración, regeneración y aclimatación *ex vitro* continúa siendo baja (Klimaszewska et al. 2016)

Actualmente se está evolucionando en la producción industrial y el desarrollo de una gran variedad de líneas embriogénicas de coníferas. Son numerosas las empresas y organizaciones que se dedican a la investigación, selección clonal y propagación de especies forestales de importancia económica, como por ejemplo, de *Pinus taeda* en Arborgen (EEUU), *Pinus radiata* en Arborgen (Nueva Zelanda), *Picea abies, Pinus pinaster* y *Pinus radiata* en FCBA (Francia), y *Picea glauca y Picea abies* en Natural Resources Quebec y JD Irving Inc. (Canadá) (Klimaszewska et al. 2007).

El descubrimiento de la ES en coníferas también ha intensificado la investigación sobre la embriogénesis cigótica. Los estudios donde se comparan los aspectos morfológicos y bioquímicos entre embriones somáticos y cigóticos, han determinado múltiples diferencias, que van desde niveles variables en acumulación de productos de reserva, hasta un perfil de expresión génica incompleto, que indica que los embriones somáticos no alcanzan el nivel de madurez que caracteriza al embrión de la semilla madura (Cairney y Pullman, 2007).

4.2. Etapas de la embriogénesis somática en coníferas

La ES se divide en una serie de etapas en las que tienen lugar fenómenos claves que permiten el desarrollo de los embriones somáticos maduros, a partir de los cuales se regenerará una planta completa. Dichas etapas son: la inducción del cultivo embriogénico, su proliferación, la maduración de los embriones y su germinación (Dunstan et al. 1995).

4.2.1. Inducción

Generalmente, para la iniciación de la ES en coníferas se utilizan como explantos las semillas inmaduras, cuando el embrión cigótico dominante está en un estado de desarrollo precotiledonario, habiendo transcurrido entre 3-6 semanas de la fecundación (Becwar et al. 1990). Actualmente, una de las limitaciones más importantes de la ES es que en muy pocas especies se ha logrado la formación de embriones somáticos empleando tejidos procedentes de individuos adultos, y por tanto con posibilidades de ser seleccionados con fiabilidad. Para solventar este inconveniente, la estrategia que se sigue es la de producir líneas embriogénicas a partir de semillas que provienen de cruzamientos controlados.

La inducción del cultivo embriogénico se suele llevar a cabo en presencia de reguladores de crecimiento (PGRs) en el medio de cultivo, fundamentalmente auxinas y citoquininas, aunque también se ha descrito la inducción del cultivo embriogénico en ausencia de los mismos, tanto en angiospermas (Fernández- Guijarro, 1997) como en gimnospermas (Lelu et al. 1999). En coníferas, los reguladores de crecimiento más utilizados son el ácido 2,4-diclorofenoxiacético (2,4-D) y la benciladenina (BA), y su disponibilidad en el medio de cultivo favorece que las células somáticas adquieran el potencial embriogénico.

Se considera que la embriogénesis somática se ha iniciado cuando hay un crecimiento continuo del tejido embriogénico y se ha producido en una cantidad suficiente para iniciar la proliferación (Figura 4a) (Klimaszewska et al. 2007).

4.2.2. Proliferación

Una vez inducida la embriogénesis somática, el tejido embrionario entra en un proceso cíclico de clonación embriónica, que no requiere, en la mayoría de las especies, la presencia de reguladores del crecimiento (Fernández Guijarro et al. 1995; Cuenca et al. 1999). En cambio en coníferas, se suele utilizar un medio de cultivo de proliferación con la misma composición que el medio de iniciación, incluyendo auxinas y citoquininas.

En esta fase es cuando se debe generar suficiente masa embrionaria,

mediante sucesivos subcultivos, que permita su desarrollo en las siguientes etapas (Figura 4b). El subcultivo del material embriogénico se realiza generalmente cada dos semanas, manteniéndose en oscuridad y 24º C. El subcultivo prolongado puede dar lugar a que el material pierda su capacidad embriogénica (Breton et al. 2006), se produzcan cambios genéticos (Fourré et al. 1997; Tremblay et al. 1999), y aumente el riesgo de contaminación. La solución a estos problemas se consigue mediante la criopreservación del material embriogénico (Cyr, 2000; Hägman et al. 2000), proceso muy importante durante la ES y que trataremos más adelante.

4.2.3. Maduración

Durante la etapa de maduración es preciso mantener un desarrollo embriogénico estable para la formación del embrión somático. En esta fase se produce la expansión de la célula, la acumulación de sustancias de reserva y se adquiere tolerancia a la desecación (Figura 4c). En coníferas, el medio de cultivo durante esta fase es suplementado con ácido abscísico (ABA), necesario para evitar la poliembrionía de partición («cleavage») y lograr que los embriones somáticos, que antes de dicho tratamiento son acúmulos globulares de unas pocas células, lleven a cabo su proceso de histogénesis, diferenciación y maduración hasta alcanzar el estado cotiledonario (Klimaszewska et al. 2002).

Un método alternativo que promueve el desarrollo de los embriones somáticos utilizado en *Pinus strobus*, es incrementado la concentración del agente gelificante en el medio de cultivo, lo que reduce la disponibilidad de agua (Klimaszewska et al. 2000). En otros protocolos descritos utilizan tratamientos con carbón activo (CA) que altera las características físico-químicas del medio de cultivo. El CA provoca el oscurecimiento del medio, y la absorción de sustancias tóxicas y de PGRs (Pullman et al. 2005; Klimaszewska y Cyr, 2002).

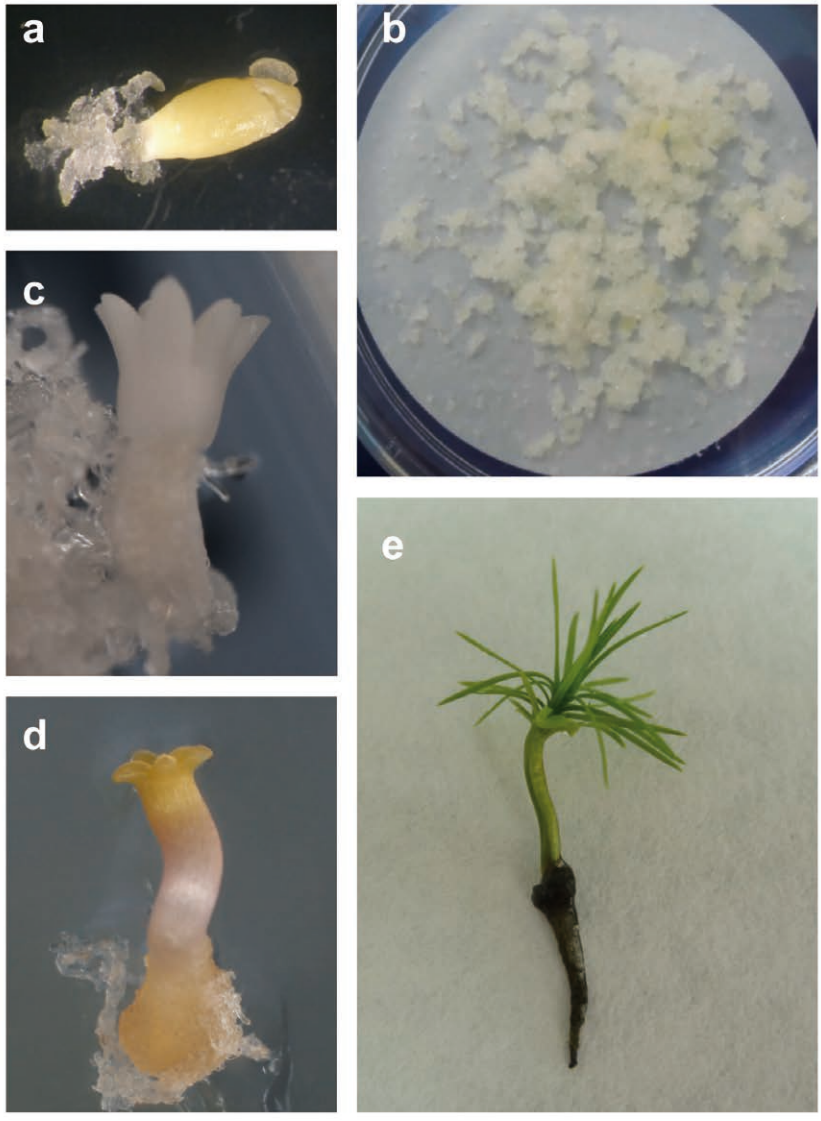


Figura 4. Etapas de desarrollo durante la ES en pino. a) Iniciación del cultivo embriogénico a partir de un embrión cigótico inmaduro (Klimaszewska et al. 2007); b) Proliferación de la masa embrionaria en subcultivo; c) Embrión somático maduro en estado cotiledonario; d) elongación del hipocótilo durante la germinación del embrión somático maduro; e) Plántula somática germinada *in vitro* durante 12 semanas.

4.2.4. Germinación

Por último, los embriones maduros han de germinar y desarrollarse como plantas completas que puedan ser transferidas a suelo. Solamente los embriones somáticos maduros con una morfología normal, con suficientes sustancias de reserva acumuladas y que hayan adquirido tolerancia a la desecación, podrán germinar y convertirse en plantas con éxito (von Arnold et al. 2002).

La germinación se lleva a cabo en medio sólido, con sacarosa, y sin PGRs, y los embriones se mantienen en oscuridad durante los primeros 10 días. Esto favorece la elongación de los hipocótilos y reduce la síntesis de antocianinas (Klimaszewska et al. 2007).

En coníferas, la germinación se completa con el desarrollo del hipocótilo y de los cotiledones (Figura 4d), proceso que tiene lugar tras 12 ó 16 semanas de cultivo (Klimaszewska y Cyr, 2002). La fuente de nitrógeno utilizada en el medio de cultivo constituirá un factor determinante para el desarrollo funcional de la raíz afectando al desarrollo de la planta y su aclimatación en invernadero (Llebrés et al. 2017).

4.3. Embriogénesis somática, criopreservación y transformación genética como herramienta biotecnológica

Además de ser una herramienta fundamental en la micropropagación de especies forestales, la ES también se ha convertido en la mejor vía de regeneración de plántulas a partir de material criopreservado (Janeiro et al. 1996; Engelmann et al. 1997) y de transformación vegetal (Peña y Séguin, 2001). Esto la convierte en una herramienta biotecnológica imprescindible en los estudios de genómica funcional en coníferas (Pascual et al. 2017).

Los cultivos de células somáticas pueden almacenarse largos periodos de tiempo sin que pierdan su viabilidad a temperaturas extremadamente bajas. La criopreservación está considerada como el método más seguro para la conservación de material biológico a largo plazo, debido a que no induce alteraciones genéticas y mantiene el potencial de regeneración. De todos los gases disponibles, el nitrógeno líquido es el más utilizado, ya que

presenta una serie de ventajas como su bajo coste, facilidad de manejo y distribución, y no resulta tóxico para el material almacenado. Esta técnica se considera esencial para la conservación de la biodiversidad vegetal y complementa a las técnicas tradicionales de depósito en bancos y colecciones de semillas.

La ES también ha representado un gran avance en los estudios funcionales de muchos genes de coníferas mediante la transformación vegetal. Además de las coníferas, muchas especies forestales son consideradas recalcitrantes para ser transformadas genéticamente y regeneradas mediante cultivo *in vitro* utilizando las técnicas tradicionales de transformación vegetal. Gracias al desarrollo de nuevas herramientas biotecnológicas y al uso de la embriogénesis somática ha sido posible establecer protocolos eficientes para transferir resistencias, seleccionar caracteres fenotípicos, y mejorar las tasas de producción en materiales de interés. Esta tecnología ha demostrado ser muy interesante para la selección de híbridos resistentes debido a su potencial para la propagación a gran escala de genotipos seleccionados.

5. Control génico y regulación molecular de la ES en coníferas

La iniciación y la maduración de embriones somáticos en coníferas está bajo un fuerte control genético (Park et al. 1993; Lelu-Walter et al. 1999; Klimaszewska et al. 2001; Miguel et al. 2004; Niskanen et al. 2004). Los datos existentes sugieren un efecto materno en la iniciación de los cultivos que puede explicarse, tanto por el genotipo, como por el estado de desarrollo fisiológico y alelo heredado del árbol madre. Por lo tanto, un mejor conocimiento del control genético durante la iniciación de la embriogénesis puede ser esencial para seleccionar genotipos de mayor valor comercial (Trontin et al. 2016).

Esclarecer los eventos moleculares que regulan el desarrollo embrionario en árboles y, especialmente en coníferas, no es una tarea sencilla debido a su gran tamaño, crecimiento lento y su gran genoma. La reciente implementación de métodos cualitativos y cuantitativos en

proteómica, transcriptómica y metabolómica genera oportunidades sin precedentes para desentrañar la compleja red de coordinación de genes reguladores durante el desarrollo embrionario en coníferas. Son numerosos los recursos disponibles en bases de datos, tanto de transcritos como de proteínas para algunas especies de pinos y piceas (Scion-led Project NZ, ProCoGen, SustainePineDB), sin embargo, son escasos los genes que han sido caracterizados durante la ES en coníferas.

La mayoría de los genes que han sido funcionalmente caracterizados durante la ES pertenecen a organismos modelo como *Arabidopsis thaliana*. Esto es debido a que las redes de co-expresión disponibles en las bases de datos actuales para esta especie permiten la identificación de genes implicados en un amplio rango de funciones biológicas como la transcripción, traducción, modificaciones post-traduccionales, respuesta a hormonas, reparación y metilación del DNA.

Los datos transcriptómicos disponibles en coníferas han mostrado tener alta homología con *Arabidopsis*, por lo que podríamos partir de la información de esta especie modelo para el estudio de las coníferas. Como ejemplo, se ha observado que existe una correlación en los perfiles de expresión de embriones cigóticos de *Arabidopsis* y *Pinus pinaster* (Xiang et al. 2011; de Vega-Bartol et al. 2013), con variaciones de expresión asociadas a las diferencias temporales en el desarrollo del embrión entre ambas especies.

En coníferas, el desarrollo de microarrays, tecnologías de secuenciación de nueva generación y PCR cuantitativa a tiempo real, ha permitido identificar algunos genes claves en el proceso de inducción (Klimaszewska et al. 2011; Elhiti et al. 2013), proliferación y maduración de los embriones (Pullman et al. 2003; Stasolla et al. 2004; de Vega-Bartol et al. 2013). También se han caracterizado genes reguladores con funciones de señalización en el megagametofito, genes relacionados con modificaciones de la pared celular, con la respuesta a auxinas y ácido abcísico, con el metabolismo del carbono, así como genes de respuesta a estrés relacionados con la tolerancia a la desecación, defensa y homeostasis (Trontin et al. 2016).

Durante la ES, además de los cambios morfológicos del tejido embrionario, se producen importantes cambios metabólicos que afectan principalmente a la acumulación de carbohidratos de reserva, metabolismo de nucleótidos, metabolismo del carbono y del nitrógeno.

Durante el proceso de maduración y germinación de embriones somáticos en especies de pino se ha demostrado que se producen cambios de expresión de enzimas implicadas en el metabolismo del nitrógeno directamente relacionadas con la síntesis de glutamina, glutamato y arginina (Businge et al. 2012).

6. Importancia de las reservas nitrogenadas en etapas tempranas del desarrollo en coníferas

El nitrógeno es uno de los nutrientes más importantes y su disponibilidad es un factor limitante durante el desarrollo de las plantas (Cole y Rapp, 1981). En la mayoría de las plantas, glutamato, glutamina, asparagina y aspartato, son los aminoácidos donadores y transportadores de nitrógeno producidos por la asimilación de nitrógeno inorgánico (Lea y Miflin, 1980). Estos compuestos se utilizan para transferir nitrógeno desde los órganos fuente hasta los sumideros y fabricar las reservas necesarias que, en periodos de inanición de nitrógeno, serán utilizadas para el crecimiento, defensa y reproducción.

En coníferas, durante la fase de maduración de los embriones, tiene lugar la biosíntesis y deposición de proteínas de reserva, que en el caso de los embriones cigóticos ocurre en el megagametofito. Estas proteínas son ricas en aminoácidos con una relación N/C elevada (Allona et al. 1994), por lo que son fundamentales tanto para el almacenaje, como para el transporte de nitrógeno.

En embriones somáticos maduros se ha descrito que la acumulación de algunas de las proteínas de reserva más abundantes, como vicilinas y leguminas, es mucho menor que en los embriones cigóticos (Morel et al. 2014a). Esto es debido principalmente a la composición del medio de cultivo, que afecta tanto a la acumulación de estas proteínas de reserva, como a la

posterior germinación de los embriones somáticos (Businge et al. 2013). De hecho, la acumulación de estas proteínas, y en concreto de las vicilinas, se ha propuesto como un marcador del estado y calidad de los embriones somáticos maduros (Morel et al. 2014b).

En plantas vasculares, el desarrollo primario comienza con la activación de un grupo de genes que permitirán la utilización de las reservas de carbono y nitrógeno de la semilla (Groome et al. 1991; Stone y Gifford, 1997). La movilización de las reservas nitrogenadas tiene lugar durante la germinación, coincidiendo con la aparición de la raíz, y depende de la activación y síntesis de enzimas clave para la hidrólisis de proteínas de reserva y catabolismo de aminoácidos como la arginina (King y Gifford, 1997). Al ser hidrolizadas suministrarán al embrión el nitrógeno y parte del carbono necesario para su desarrollo**.**

El alto contenido en nitrógeno del material de reserva en coníferas, es suficiente para mantener el crecimiento y el metabolismo en etapas iniciales de desarrollo, de modo que no sería necesaria la incorporación de nitrógeno exógeno hasta que las reservas estuviesen agotadas (Cánovas et al. 1998). El aminoácido más rico en nitrógeno es la arginina, y constituye gran parte del total de aminoácidos de las proteínas de reserva. En etapas tempranas de la germinación, la arginina producida durante la hidrólisis de las proteínas de reserva es canalizada al resto de la planta en forma de glutamina y asparagina para la biosíntesis de los compuestos nitrogenados de la nueva planta (Cánovas et al. 2007). Por lo tanto, la biosíntesis de arginina es considerada una de las rutas metabólicas más relevantes durante la embriogénesis y el desarrollo temprano de las coníferas.

7. Metabolismo de la arginina y su regulación en plantas

La arginina tiene múltiples destinos metabólicos siendo uno de los aminoácidos más versátiles. Su metabolismo está modulado por las actividades de transportadores que mueven la arginina y sus metabolitos derivados a través de membranas inter e intracelulares (Winter et al. 2015).

La ruta de síntesis de arginina en plantas es notable por su complejidad y variabilidad a nivel genético y por su interconexión con otras vías, como la biosíntesis de pirimidinas y poliaminas. Por ello, y a pesar de su importancia, solo ha sido caracterizada parcialmente en plantas. Esta ruta puede dividirse en dos fases, primero la ornitina es sintetizada a partir de glutamato en una ruta cíclica o lineal, y en una segunda fase la arginina es sintetizada a partir de ornitina (Winter et al. 2015) (Figura 5).

Una enzima clave de la ruta de biosíntesis de arginina es la *N*-acetil glutamato quinasa (NAGK). La regulación de esta enzima por proteínas PII fue descrita por primera vez en cianobacterias (Heinrich et al. 2004) y *Arabidopsis* (Burillo et al. 2004). La formación del complejo con proteínas PII reduce la afinidad de NAGK por su inhibidor arginina, aumentando su actividad y por tanto, la propia síntesis de arginina.

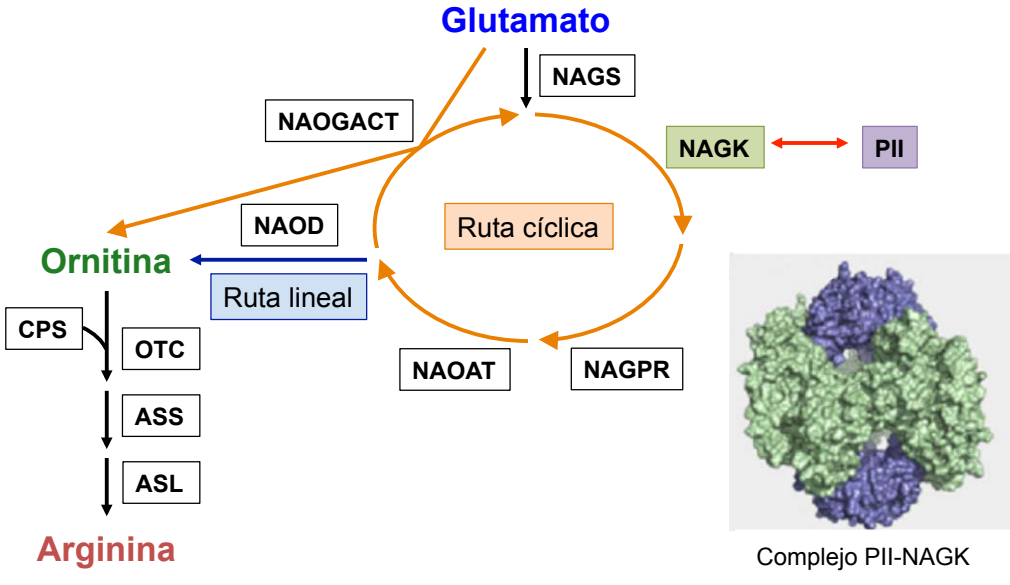


Figura 5. Esquema representativo de la ruta de biosíntesis de arginina en plantas y su regulación por proteínas PII. Principales enzimas implicadas en la síntesis de arginina y estructura del complejo regulador formado por PII (morado) y NAGK (verde) en *S. elongatus* (imagen tomada de Chellamuthu et al. 2013). NAGS, N-acetilglutamato sintasa; NAGK, N-acetilglutamato quinasa; NAGPR, N-acetilglutamato-5-P reductasa; NAOAT, N-acetilornitina aminotransferasa; NAOGACT, N-acetilornitina-glutamato acetiltransferasa; NAOD, N-acetilornitina deacetilasa; CPS, carbamoil fosfato sintetasa; OTC, ornitina transcarbamilasa; ASS, argininosuccinato sinthasa; ASL, argininosuccinato liasa.

La regulación del metabolismo del nitrógeno por proteínas PII está bien caracterizada en cianobacterias (Heinrich et al. 2004), pero su función fisiológica en plantas no es muy conocida. Las proteínas PII están ampliamente distribuidas en la naturaleza estando presentes tanto en procariotas como en eucariotas (Forchhammer, 2004), y su secuencia está muy conservada (Leigh y Dodsworth, 2007). En cuanto a la estructura de estas proteínas se ha determinado que PII es un trímero y NAGK un hexámero. La estructura del complejo cristalino que forma PII con NAGK en procariotas muestra dos trímeros de PII que se encuentran a ambos lados del anillo formado por el hexámero de NAGK (Llácer et al. 2007; Chellamuthu et al. 2013) (Figura 5), y esta estructura es muy similar a la descrita en *Arabidopsis* (Mizuno et al. 2007).

Las proteínas PII actúan como sensores del balance carbono/ nitrógeno y del estatus energético de la célula mediante su unión con ATP y ADP, que compiten por un mismo sitio, así como 2-oxoglutarato (2OG), esqueleto carbonado utilizado en la síntesis de glutamato. Los niveles de 2-oxoglutarato dependen del estado nitrogenado de las células (Muro-Pastor et al. 2001). El 2-oxoglutarato se une sinérgicamente con el ATP (Ninfa y Jiang, 2005), de modo que en ausencia de ATP no hay unión de 2-oxoglutarato.

También se ha descrito que estas proteínas controlan el metabolismo de ácidos grasos (Bourrellier et al. 2010) al interaccionar con acetyl-CoA carboxylasa (ACCasa), y recientemente se ha propuesto que actúan como sensores de glutamina en plantas. En este caso, la regulación de PII se produce por la unión directa de glutamina a un sitio constituido por entre 13-19 residuos conservados en la parte C-terminal de las proteínas PII de plantas, llamado lazo Q (Chellamuthu et al. 2014).

8. Papel de los transportadores de aminoácidos en plantas

Los aminoácidos pueden considerarse como los metabolitos dominantes del metabolismo del nitrógeno en plantas, ya que constituyen el producto inicial en la asimilación primaria. Las plantas incorporan tanto el N inorgánico (amonio y nitrato) como el N orgánico (aminoácidos y péptidos) disponible en el suelo. Se ha demostrado que las coníferas tienen preferencia por la absorción de amonio sobre nitrato, pero también incorporan aminoácidos como fuentes principales de N para su crecimiento (Näsholm et al. 2009; Castro-Rodríguez et al. 2017).

La síntesis, uso y degradación de aminoácidos tiene lugar en diversos compartimentos celulares: cloroplastos, citosol, peroxisomas, mitocondrias y vacuolas. Por lo tanto, es necesario que los aminoácidos sean transportados a través de las membranas celulares para conectar las diferentes rutas metabólicas. El N inorgánico procedente del suelo es incorporado rápidamente en forma de aminoácidos en la raíz y en las hojas maduras (Andrews, 1986). Mientras que parte del N es utilizado en la biosíntesis de proteínas o como precursor de moléculas esenciales en estos tejidos, la mayoría es transportado por el sistema vascular de las plantas desde los sitios de asimilación primaria hasta otros órganos o tejidos para satisfacer sus necesidades nutricionales. Estos tejidos, que incluyen las hojas en desarrollo, meristemos y órganos reproductivos, deben por lo tanto importar los aminoácidos necesarios para su crecimiento y desarrollo.

La mayoría del transporte asociado a membranas biológicas es mediado por proteínas de membrana que contienen dominios hidrofóbicos con estructura de alfa-hélice y en algunos casos, de lámina-beta. Las plantas poseen múltiples transportadores de aminoácidos o permeasas con diferentes propiedades, en función de su distribución en diferentes tejidos, especificidad de sustrato y afinidad, e incluso por su regulación endógena o ambiental (Williams y Miller, 2001).

En los últimos años, mediante expresión heteróloga de genes en *Xenopus* y en *Saccharomyces cerevisiae*, se han identificado y caracterizado muchos de estos transportadores, la mayoría de ellos de *Arabidopsis*.

El progreso en los estudios genómicos ha permitido la identificación directa de numerosos transportadores de aminoácidos en otras especies de interés como *Oryza sativa (Zhao et al. 2012), Selaginella moellendorffii (Wipf et al. 2012*), *Populus trichocarpa (Wu et al. 2015*), *Glycine max (Cheng et al. 2016*), *Solanum tuberosum (Ma et al. 2016*) *y Ricinus communis (Xie et al. 2016*). En el genoma de *Arabidopsis* se han identificado unos 100 genes que codifican transportadores de aminoácidos (Pratelli et al. 2014). La mayoría pertenecen a la familia de permeasas de aminoácidos y auxinas ATF, dentro de la cual se encuentra la familia de permeasas de aminoácidos AAP, que cuenta con ocho miembros en *Arabidopsis y* es la familia mejor caracterizada. Mediante análisis de mutantes se ha determinado que los AAPs de *Arabidopsis* están implicados en el transporte de larga distancia a través del floema y en el transporte de aminoácidos en el endospermo de la semilla (Schmidt et al. 2007; Sanders et al. 2009; Hunt et al. 2010),desempeñando un papel regulador en la biosíntesis de proteínas de reserva (Hirner et al. 1998).

La regulación del contenido de aminoácidos, su flujo y transporte a través de la planta son críticos para la adaptación al estatus carbono/ nitrógeno, desarrollo y defensa (Pratelli et al. 2014). Todavía no está muy claro cómo las plantas detectan los aminoácidos del medio, así como las rutas de señalización correspondientes, y aunque los genes candidatos son numerosos, por el momento no ha sido caracterizado un sensor de nitrógeno o de aminoácidos inequívoco en plantas (Dinkeloo et al. 2017).

9. Bibliografía

Allona I, Collada C, Aragoncillo C (1994). 2S arginine-rich proteins from *Pinus pinaster* seeds. *Tree Physiol* 14: 211-218.

Andrews M (1986). The partitioning of nitrate assimilation between root and shoot of higher plants. *Plant, Cell & Environment*, 9: 511-519.

Becker P, Meinzer FC, Wullschleger SD (2000). Hydraulic limitation of tree height: a critique. Functional Ecology 14: 4-11.

Becwar MR, Nagamani R, Wann SR (1990). Initiation of embryogenic cultures and somatic embryo development in Loblolly pine (*Pinus taeda*). Can. J. For. Res. 20: 810-817.

Bonga JM, Klimaszewska KK, von Aderkas P (2010). Recalcitrance in clonal propagation, in particular of conifers. Plant Cell Tiss Org Cult 100: 241-254.

Bourrellier ABF, Valot B, Guillot A, Ambard-Bretteville F, Vidal J, Hodges M (2010). Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate regulated by interaction of PII with the biotin carboxyl carrier subunit. *Proc. Natl Acad. Sci. USA*, 107, 502-507.

Breton D, Harvengt L, Trontin JF, Bouvet A, Favre JM (2006). Long-term subculture randomly affects morphology and subsequent maturation of early somatic embryos in maritime pine. *Plant Cell Tiss Organ Cult* 87: 95.

Brodribb TJ, Feild TS (2008). Evolutionary significance of a flat-leaved *Pinus* in Vietnamese rainforest. New phytologist 178: 201-209.

Brodribb TJ, Pittermann J, and Coomes DA (2012). Elegance versus Speed: Examining the Competition between Conifer and Angiosperm Trees. Int. J. Plant Sci. 173(6):673-694.

Buchholz JT (1926). Origin of cleavage polyembryony in conifers. Bot Gaz. 81:55-71.

Burillo S, Luque I, Fuentes I, Contreras A (2004). Interactions between the nitrogen signal transduction protein PII and N-acetyl glutamate kinase in organisms that perform oxygenic photosynthesis. J. Bacteriol. 186: 3346-3354.

Bush DR (1999). Amino acid transport. In: Singh BK (ed) Plant Amino Acids: biochemistry and biotechnology, Marcel Dekker, New York, pp. 305-318.

Businge E, Brackmann K, Moritz T, Egertsdotter U (2012) Metabolite profiling reveals clear metabolic changes during somatic embryo development of Norway spruce (*Picea abies*). *Tree Physiol* 32: 232-244.

Businge E, Bygdella J, Wingslea G, Moritza T, Egertsdotter U (2013). The effect of carbohydrates and osmoticum on storage reserve accumulation and germination of Norway spruce somatic embryos. Physiol Plant 149:273-285.

Cairney J, Pullman GS (2007). The cellular and molecular biology of conifer embryogenesis. New Phytologist, 176: 511-536.

Campbell MM, Brunner AM, Jones HM, Strauss SH (2003). Forestry's fertile crescent: the application of biotechnology to forest trees. Forest Biotechnology 1:141-154.

Cánovas FM, Cantón FR, García-Gutiérrez A, Gallardo F and Crespillo R (1998). Molecular physiology of glutamine and glutamate biosíntesis in developing conifer seedlings. *Physiol. Plant*. 103: 287-294.

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

Cañas RA, de la Torre F, Pascual MB, Avila C, Cánovas FM (2016). Nitrogen Economy and Nitrogen Environmental Interactions in Conifers. *Agronomy*, 6, 26.

Carel W, Vergeldt F, De Jager A, Van H (2006). MRI of long distance water transport: A comparison of the phloem and xylem flow characteristics and dynamics in poplar, castor bean, tomato and tobacco. *Plant Cell & Environment*, 29 (9): 1715-1729.

Castro-Rodríguez V, Cañas RA, de la Torre FN, Pascual MB, Avila C, Cánovas FM (2017). Molecular fundamentals of nitrogen uptake and transport in trees. J Exp Bot. 68: 2489-2500.

Celestino C, Hernández I, Carneros E, López-Vela D, Toribio M (2005). La embriogénesis somática como elemento central de la biotecnología forestal. *Invest Agrar: Sist Recur For* , 14(3): 345-357.

Chalupa V (1985). Somatic embryogenesis and plantlet regeneration from cultured immature and mature embryos of Picea abies (L.) Karst. *Comm Inst For Cech* 14: 57-63.

Chandler JW (2008). Cotyledon organogenesis. *J Exp Bot*, 59(11): 2917-2931.

Chellamuthu, VR, Alva V, Forchhammer K (2013). From cyanobacteria to plants: conservation of PII functions during plastid evolution. *Planta* 237: 451.

Chellamuthu VR, Ermilova E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, et al. (2014). A widespread glutamine-sensing mechanism in the plant kingdom. *Cell* 159: 1188-1199.

Chen YM, Ferrar TS, Lohmeir-Vogel E, Morrice N, MizunoY, Berenger B, Ng KK, Muench DG, Moorhead GB (2006). The PII signal transduction protein of *Arabidopsis thaliana* forms an arginine-regulated complex with plastid N-acetyl glutamate kinase. *J. Biol. Chem.* 281: 5726-5733.

Cheng L, Yuan HY, Ren R, Zhao SQ, Han YP, Zhou QY, Ke DX, Wang YX, Wang L (2016). Genome-wide identification classification, and expression analysis of amino acid transporter gene family in Glycine

max. *Front. Plant Sci*. 7: 515.

Crabeel M, Abadjieva A, Hilven P, Desimpelaere J, Soetens O (1997). Characterization of the *Saccharomyces cerevisiae ARG7* gene encoding ornithine acetyltransferase, an enzyme also endowed with acetylglutamate synthase activity. *Eur. J. Biochem*. 250: 232-241.

Cole DW, Rapp M (1981). Elemental cycling in forest ecosystems. In: Reichle DE (ed). Dynamic properties of forest ecosystems. Cambridge University Press, Cambridge, pp 341-409.

Cuenca B, San-José M, Martínez M, Ballester A, Vieitez A (1999). Somatic embryogenesis from stem and leaf explants of *Quercus robur* L. *Plant Cell Rep.* 18: 538-543.

Cyr D (2000). Cryopreservation: roles in clonal propagation and germplasm conservation of conifers. Engelmann F, Takagi H (eds). Cryopreservation of tropical plant germplasm. Current research progress and application. Japan International Research Centre for Agricultural Sciences. Tsukuba/International Plant Genetic Resources Institute. Rome. pp 261-268.

de Vega-Bartol JJ, Sim.es M, Lorenz WW, Rodrigues AS, Alba R, Dean JFD, Miguel C (2013). Transcriptomic analysis highlights epigenetic and transcriptional regulation during zygotic embryo development of Pinus pinaster. *BMC Plant Biol* 13:123.

Dinkeloo K, Boyd S, Pilot G (2017). Update on amino acid transporter functions and on possible aminoacid sensing mechanisms in plants. *Seminars in Cell & Developmental Biology.* doi.org/10.1016/j.semcdb.2017.07.010.

Dodeman VL, Ducreux G, Kreis M (1997). Zygotic embryogenesis versus somatic embryogenesis. *J Exp Bot* 48:1493-1509.

Dunstan DI, Tautorus TE, Thorpe TA (1995). Somatic embryogenesis in woody plants. In: T.A. Thorpe (Ed.) In Vitro Embryogenesis in Plants, pp. 471-538. Kluwer Academic Publishers, Dordrecht.

Elhiti M, Stasolla C, Wang A (2013). Molecular regulation of plant somatic embryogenesis. *In Vitro Cell Dev Biol-Plant.* 49:631-642.

Engelmann F (1997). *In vitro* conservation methods. In: Biothechnology and Plant Genetic Resources. Callow JA, Ford-Lloyd BV and Newbury HJ (eds). CABI, Oxon, UK, pp. 119-161.

Farjon AA (2010). Handbook of the World's Conifers, Brill Academic Publishers, Leiden/Boston.

Fernández-Guijarro B, Celestino C, Toribio M (1994). Somatic embryogenesis in *Quercus suber* L. En: Biotecnology of Trees (Pardos JA, Ahuja MR, Elena Rossello R, eds). Investigación Agraria. Sistemas y Re- cursos Forestales. Fuera de Serie 4. INIA. Madrid, España. pp. 105-110.

Forchhammer K (2004). Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signal to targets. *FEMS Microbiol.*

Rev. 28, 319-333.

Fourré JL, Berger P, Niquet L, André P (1997) Somatic embryogenesis and somaclonal variation in Norway spruce: mophogenetic, cytogenetic and molecular approaches. Theor. Appl. Genet. 94: 159-169.

Groome MC, Axler SR and Gifford DJ (1991). Hydrolysis of lipid and protein reserves in relation to protein electrophoretic patterns following imbibition. *Physiol. Plant*. 83: 99-106.

Gutiérrez-Mora A, González-Gutiérrez AG, Rodríguez-Garay B, Ascencio-Cabral A and Li-Wei L (2012). Plant Somatic Embryogenesis: Some Useful Considerations, Embryogenesis, Dr. Ken-Ichi Sato (ed.). ISBN: 978-953-51-0466-7.

Häggman H, Aronen T, Ryynänen L (2000) Cryopreservation of embyogenic cultures of conifers. In: Jain SM, Gupta PK, Newton RJ (eds) Somatic embryogenesis in woody plants. Vol 6. Kluwer Academic Publishers, Dordrecht, pp. 707-728.

Hakman I, Fowke LC, von Arnold S, Eriksson T (1985) The development of somatic embryogenesis in tissue cultures initiated from immature embryos of Picea abies (Norway spruce). Plant Sci 38:53-59.

Heinrich A, Maheswaran M, Ruppert U, Forchhammer K (2004). The *Synechococcus elongatus* PII signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. Mol. Microbiol. 52: 1303-1314.

Hirner B. Fischer WN, Rentsch D, Kwart M, Frommer WB (1998). Developmental control of H+/amino acid permease gene expression during seed development of *Arabidopsis*. *Plant. J.* 14: 535-44.

Hunt E, Gattolin S, Newbury HJ, Bale JS, Tseng HM, Barrett DA, Pritchard J (2010). A mutation in amino acid permease AAP6 reduces the amino acid content of the *Arabidopsis* sieve elements but leaves aphid herbivores unaffected. *J. Exp. Bot.* 61, 55-6410.

Janeiro LV, Vieitez AM, Ballester A (1996). Cryopreservation of somatic embryos and embryonic axes of *Camellia japonica* L. Plant Cell Reports 15: 699.

Jürgens G, Ruiz RAT, Laux T, Mayer U, Berleth T (1994). Early Events in Apical-basal Pattern Formation in *Arabidopsis*. In: Coruzzi G, Puigdomènech P (eds). Plant Molecular Biology. NATO ASI Series (Series H: Cell Biology), vol 81. Springer, Berlin, Heidelberg.

King JE and Gifford DG (1997). Amino acid utilization in seeds of loblolly pine during germination and early seedling growth. *Plant Physiol.* 113, 1124-1135.

Klimaszewska K, Bernier-Cardou M, Cyr DR, Sutton BC (2000) Influence of gelling agents on culture medium gel strength, water availability, tissue water potential, and maturation response in embryogenic cultures of

Pinus strobus L. *In Vitro Cell. Dev.Biol.-Plant* 36: 279-286.

Klimaszewska K, Park Y-S, Overton C, Maceacheron I, Bonga JM (2001) Optimized somatic embryogenesis in *Pinus strobus* L. In Vitro Cell Dev Biol-Plant 37:392-399.

Klimaszewska K, Cyr DR (2002) Conifer somatic embryogenesis: I. Development. Dendrobiology 48: 31-39.

Klimaszewska K, Trontin JF, Becwar MR, Devillard C, Park Y-S, Lelu-Walter MA (2007). Recent progress in somatic embryogenesis of four Pinus spp. En: da Silva J, Shima K (eds). Tree and Forestry Science and Biotechnology Global Science Books, 1(1): 11-25.

Klimaszewska K, Overton C, Stewart D, Rutledge RC (2011). Initiation of somatic embryos and regeneration of plants from primordial shoots of 10-year-old somatic white spruce and expression profiles of 11 genes followed during the tissue culture process. *Planta* 233:635-647.

Klimaszewska K, Hargreaves C, Lelu-Walter M-A, Trontin JF (2016) Advances in conifer somatic embryogenesis since year 2000. In: Germanà MA, Lambardi M (eds) In Vitro Embryogenesis in Higher Plants, Methods in Molecular Biology, vol. 1359, Chapter 7, Springer Science+Business Media New York, pp. 131-166.

Lea PJ and Miflin BJ (1980). Transport and metabolism of asparagine and other nitrogen compounds within the plant. In: Miflin Bj (ed) The biochemistry of plants, New York, Academic Press, 5: 569-607.

Leigh JA y Dodsworth JA (2007). Nitrogen regulation in bacteria and archaea. *Annual Reviews of Microbiology* 61, 349-377.

Lelu-Walter MA, Bastien C, Drugeault A, Gouez ML, Klimaszewska K (1999). Somatic embryogenesis and plantlet development in *Pinus sylvestris* and *Pinus pinaster* on medium with and without growth regulators. *Physiologia Plantarum*. 105: 719-728.

Llácer JL, Contreras A, Forchhammer K, Marco-Marín C, Gil-Ortiz F, Maldonado R, Fita I, Rubio V (2007). The crystal structure of the complex of PII and acetylglutamate kinase revelas how PII controls the storage of nitrogen as arginine. *Proceedings of the National Academy of Sciences USA* 104: 17644-17649.

Llebrés MT, Pascual MB, Debille S, Trontin JF, Avila C, Cánovas FM (2017) The role of arginine metabolic pathway during embryogenesis and germination in maritime pine (*Pinus pinaster* Ait.). *Tree Physiol*ogy 3:1-14. doi: 10.1093/treephys/tpx133

Ma H, Cao X , Shi S, Li S, Gao J, Ma Y, Zhao Q, Chen Q (2016). Genome-widesurvey and expression analysis of the amino acid transporter superfamily in potato (Solanum tuberosum L.). Plant Physiol. Biochem. 107: 164-177.

Meinnel T, Schmitt E, Mechulam Y, Blanquet S (1992). Structural and biochemical characterization of the *Escherichia coli argE* gene product. J.Bacteriol. 174: 2323-2331.

Merino I, Malin Abrahamsson M, Lieven Sterck L, Craven-Bartle B, Canovas F, von Arnold S (2016). Transcript profiling for early stages during embryo development in Scots pine. *BMC Plant Biology* 16: 255.

Miguel C, Gonc¸alves S, Tereso S, Marum L, Oliveira MM (2004). Somatic embryogenesis from 20 open-pollinated seed families of Portuguese plus trees of maritime pine. *Plant Cell Tissue Org Cult*. 76:121-130.

Mizuno Y, Moorhead GBG, Ng KKS (2007). Structural basis for the regulation of N-acetylglutamate kinase by PII in Arabidopsis thaliana. *The Journal of Biological Chemistry.* 282: 35733-35740.

Molesini B, Mennella G, Martini F, Francese G, Pandolfini T (2015). Involvement of the putative N-acetylornithine deacetylase from *Arabidopsis thaliana* in flowering and fruit development. *Plant Cell Physiol.* 56: 1084-1096.

Moorhead GBG, Smith CS (2003). Interpreting the plastid carbon, nitrogen and energy status. A role for PII?. *Plant Physiol* 133: 1-7.

Morel A, Teyssier C, Trontin JF et al. (2014a). Early molecular events involved in Pinus pinaster Ait. somatic embryo development under reduced wáter availability: transcriptomic and proteomic analyses. *Physiol Plant*. doi:10.1111/ppl.12158.

Morel A, Trontin JF, Corbineau F et al. (2014b). Cotyledonary somatic embryos of *Pinus pinaster* Ait. most closely resemble fresh, maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses. *Planta* 240: 1075.

Muro-Pastor MI, Reyes JC, Florencio FJ (2001). Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. *J Biol Chem* 276: 38320-38328.

Näsholm T, Kielland K, Ganeteg U (2009). Uptake of organic nitrogen by plants. New Phytologist 182: 31-48.

Ninfa AJ, Jiang P (2005). PII signal transduction proteins: sensors of alpha-ketoglutarate that regulate nitrogen metabolism. *Curr Opin Microbiol*. 8(2):168-73.

Niskanen AM, Lu J, Seitz S, Keinonen K, von Weissenberg K, Pappinen A (2004). Effect of parent genotype on somatic embryogenesis in Scots pine (*Pinus sylvestris*). *Tree Physiol* 24: 1259-1265.

Ortiz-Lopez A, Chang HC, Bush DR (2000). Amino acid transporters in plants. *Biochim. Biophys. Acta* 1465: 275-80.

Owens JN, Molder M (1984). The Reproductive Cycle of Lodgepole Pine. B.C. Min. For. Inf. Serv. Br. Victoria. 29 pp.

Park YS, Pond SE, Bonga JM (1993). Initiation of somatic embryogenesis in white spruce (*Picea glauca*): genetic control, culture treatment effects, and implications for tree breeding. *Theoretical and Applied Genetics* 86: 427-436.

Peña L, Séguin A (2001). Recent advances in the genetic transformation of trees. *Trends in biotechnology* 19(12):500-506.

Pratelli R, Pilot G (2014). Regulation of amino acid metabolic enzymes and transporters in plants. *Journal of Experimental Botany,* 65(19): 5535-5556.

Pullman GS, Johnson S, Peter G, Cairney J, Xu N (2003). Improving loblolly pine somatic embryo maturation: comparison of somatic and zygotic embryo morphology, germination, and gene expression. *Plant Cell Rep.* 21:747-758.

Pullman GS, Gupta PK, Timmis R, Carpenter C, Kreitinger M, Welty E (2005) Improved Norway spruce somatic embryo development through the use of abscisic acid combined with activated carbon. *Plant Cell Rep* 24: 271.

Reinert, J (1958). Untersuchungen uber die morphogenese in gewebekulturen. Ver. Dtsch. Bot. Ges. 71: 15-24.

Sanders A, Collier R, Trethewy A, Gould G, Sieker R, Tegeder M (2009). AAP1 regulates import of amino acids into developing Arabidopsis embryos. *The Plant Journal* 59, 540-552.

Schmidt, R., Stransky, H. and Koch, W. (2007) The amino acid permease AAP8 is important for early seed development in Arabidopsis thaliana. *Planta* 226, 805-813.

Svennerstam H, Ganeteg U, Näsholm T (2008) Root uptake of cationic amino acids by Arabidopsis depends on functional expression of amino acid permease. *New Phytologist* 180: 620-630.

Secretaría del Convenio sobre la Diversidad Biológica (2009). Informe sobre la Conservación de las Especies Vegetales: Una revisión de los progresos realizados en la aplicación de la Estrategia Mundial para la Conservación de Plantas (GSPC). 48 pgs.

Shargool PD, Jain JC, McKay G (1988). Ornithine biosynthesis, and arginine biosynthesis and degradation in plant cells. *Phytochemistry* 27: 1571-1574.

Singh H (1978). Embryology of gymnosperms. Encyclopedia of plant anatomy XII. pp.302 Berlin: Borntrager.

Sorensen, FC (1982). The roles of polyembryony and embryo viability in the genetic system of conifers. *Evolution,* 36: 725-733.

Stasolla C, Bozhkov PV, Chu TM, Van Zyl L, Egertsdotter U, Suarez MF, Craig D, Wolfinger RD, Von Arnold S, Sederoff RR (2004). Variation in transcript abundance during somatic embryogenesis in gymnosperms. *Tree Physiol.* 24(10):1073-85.

Steward F, Mapes M, Mears K (1958). Growth and organized development of

cultured cells. *American Journal of Botany*. 45: 705-708

Stone SL, Gifford DJ (1997). Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early-seedlings growth, I. Storage protein reserves. *Int. J. Plant. Sci*. 158: 727-737.

Thompson (2015). Challenges for the large-scale propagation of forest trees by somatic embryogenesis. Proceedings of the Third International Conference of the IUFRO unit 2.09.02 on "Woody Plant Production Integrating Genetic and Vegetative Propagation Technologies" Vitoria-Gasteiz, Spain, 8-12 September 2014 pp.81-91 ref.32.

Tremblay L, Levasseur C, Tremblay FM (1999). Frequency of somaclonal variation in plants of black spruce (Picea mariana, Pinaceae) and white spruce (Picea glauca, Pinaceae) derived from somatic embryogenesis and identification of some factors involved in genetic instability. *Am. J. Bot.* 86: 1373-1381.

Trontin JF, Klimaszewska K, Morel A, Hargreaves C, Lelu-Walter M-A (2016). Molecular aspects of conifer zygotic and somatic embryo development: a review of genome-wide approaches and recent insights. In: Germanà MA, Lambardi M (eds) *In Vitro Embryogenesis in Higher Plants*, Methods in Molecular Biology, vol. 1359, Chapter 8, DOI 10.1007/978-1-4939-3061-6-8, Springer Science+Business Media New York, pp. 167-207.

Turnbull MH, Tissue DT, Griffin KL, Rogers GND and Whitehead D (1998). Photosynthetic acclimation to long-term exposure to elevated CO2 concentration in *Pinus radiata* D. Don. is related to age of needles. *Plant, Cell & Environment,* 21: 1019-1028.

Umehara M, Kamada H (2005). Development of the embryo proper and the suspensor during plant embryogenesis. *Plant Biotechnology* 22, 253-260.

von Arnold S, Sabala I, Bozhkov P, Dyachok J, Filonova L (2002). Developmental pathways of somatic embryogenesis. *Plant Cell Tiss. Org. Cult*. 69: 233-249.

Wendrich JR, Weijers D (2013). The Arabidopsis embryo as a miniature morphogenesis model. New Phytol. 199:14-25.

Williams LE, Miller AJ (2001). Transporters responsible for the uptake and partitioning of nitrogenous solutes. Annu. Rev. Plant Physiol. *Plant Mol. Biol*. 52:659-88.

Willemsen V, Scheres B (2004). Mechanisms of Pattern Formation in Plant Embryogenesis. *Annual Review of Genetics*, 38: 587-614.

Winter G, Todd CD, Trovato M, Forlani G, Funck D (2015). Physiological implications of arginine metabolism in plants. *Front. Plant Sci.* 6: 534.

Wipf D, Loque D, Lalonde S, Frommer WB (2012). Amino Acid transporter inventory of the selaginella genome. *Front. Plant Sci.* 3: 36.

Wu M, Wu S, Chen Z, Dong Q, Yan H, Xiang Y (2015). Genome-wide survey andexpression analysis of the amino acid transporter gene family in poplar. *TreeGenet. Genom*. 11 (4): 83.

Xiang D, Venglat P, Tibiche C, Yang H, Risseeuw E, Cao Y, Babic V, Cloutier M, Keller W, Wang E, Selvaraj G, Datla R (2011). Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in Arabidopsis. *Plant Physiol*. 156: 346-356.

Xie Y, Zhao JL, Wang CW, Yu AX, Liu N, Chen L, Lin F, Xu HH(2016). Glycinergic-fipronil uptake is mediated by an amino acid carrier system andinduces the expression of amino acid transporter genes in Ricinuscommunis seedlings. *J. Agric. Food Chem.* 64(19): 3810-3818.

Zhao H, Ma H, Yu L, Wang X, Zhao J (2012). Genome-wide survey and expressionanalysis of amino acid transporter gene family in rice (Oryza sativa L.). *PLoSOne*, 7 (11).

II. OBJETIVOS

De acuerdo con los antecedentes que se han recogido en la Introducción, el objetivo general de esta Tesis Doctoral es estudiar la economía del nitrógeno en las etapas tempranas del desarrollo de las coníferas, que incluyen la formación de la semilla y la germinación para dar lugar a una planta con plena capacidad autótrofa.

Este objetivo general se ha abordado mediante el planteamiento de una serie de objetivos específicos que se han incluido en dos fases bien diferenciadas del trabajo experimental que se plantea.

Fase I. Embriogénesis somática de pino, criopreservación y generación de líneas transgénicas.

Objetivos específicos:

1. Establecer los métodos de cultivo *in vitro* que permitan la inducción de embriones somáticos, la maduración y la germinación para su conversión en plantas viables.
2. Utilizar la embriogénesis somática como herramienta biotecnológica para el análisis funcional de genes.

Fase II. Estudio comparativo de las reservas nitrogenadas y el metabolismo de la arginina en embriones somáticos y cigóticos de pino.

Objetivos específicos:

3. Identificar y caracterizar los genes implicados en la biosíntesis y degradación de arginina.
4. Estudiar el papel de las proteínas PII en la regulación de la ruta de síntesis de arginina a través de la interacción con su diana metabólica.
5. Identificar y caracterizar funcionalmente transportadores potenciales de arginina.

III. RESULTADOS Y DISCUSIÓN

ARTÍCULO 1

Root growth of somatic seedlings of hybrid Pinus strobus (L.) and P. wallichiana (A. B. Jacks.) is affected by the nitrogen composition of the germination medium

Llebrés MT, Avila C, Cánovas FM, Klimaszewska K

Trees (2017), doi: 10.1007/s00468-017-1635-2

1. Abstract

Mature somatic embryos of F2 hybrid *P. strobus* x *P. wallichiana* backcrossed with *P. strobus* converted to plants but survival of the somatic plants was not satisfactory prompting the present study on somatic seedling root growth on germination media varying in nitrogen (N) composition. The media were modifications of Litvay's (Litvay et al. 1985) which included two main groups: G1, G2, G3, G4 all contained inorganic N with or without glutamine (Gln) or casein hydrolysate (CH) and G5, G6, G7 contained solely glutamine and/or CH. In addition, G8 was half-strength G1 (with organic N) and G9 was half-strength CD (Campbell et al. 1975) without organic N. The roots of plantlets growing on media containing solely organic N grew about 2.55 times longer than on those containing solely inorganic N or both inorganic and organic N. The longest roots grew on G7 supplemented with CH and on G5 with both CH and Gln. Microarray analysis of somatic plants germinated on G1 versus G7 revealed that depending on the N source the somatic plants displayed changes in the transcriptome resulting in the differential expression of a range of genes involved in essential processes for plant growth and development. Roots grown in the absence of inorganic N were capable of rapid uptake of labelled inorganic ^{15}N during the 2 h incubation in the nutrient solution. The somatic plants from G5 medium acclimatized at 50% higher rate than those from G1 (with both inorganic and organic N) and G2 (solely inorganic N) under standard fertilization regime.

2. Introduction

Somatic embryogenesis (SE) initiated from zygotic embryos of conifers, coupled with cryopreservation of embryo forming cell lines, has proven to be a useful biotechnological tool for mass propagation of genetically improved seeds, valuable hybrid seeds and for species conservation. Within the genus *Pinus*, SE from immature zygotic embryos of *P. strobus* (eastern white pine) has been quite successful due to the relatively high initiation rate of embryonal masses followed by efficient somatic embryo

development/maturation and conversion to somatic plants (Klimaszewska et al. 2000, Garin et al. 2000, Klimaszewska et al. 2001, Klimaszewska et al. 2004). Similarly, zygotic embryos of the F2 hybrids of *P. strobus* x *P. wallichiana* (Himalayan pine) responded at 22 and 52% to the SE initiation protocol in 2007 and 2008, respectively from multiple control-pollinated seed families (Daoust et al. 2009). The goal of the hybridization between these 5-needle pines has been the transfer of white pine blister rust (WPBR, *Cronartium ribicola* Fisch.) resistance from *P. wallichiana* to *P. strobus*. WPBR is an exotic pathogen that was accidentally introduced into North America in the early 20th century and devastated both naturally and artificially regenerated stands in eastern Canada and is still a significant obstacle to the re-establishment of the natural stands that once covered extensive areas. However, the first-generation hybrids (F1) are not as well adapted as native eastern white pines and they remain susceptible to the harsh winter conditions in eastern Canada. Consequently, the present breeding program focuses on the production of F2 backcross hybrids and efforts are being devoted to increasing the native gene component and enhancing adaptability (Daoust et al. 2009). The technology of SE has sparked interest in the selection of resistant white pine hybrids because this approach has a potential for large-scale propagation of selected genotypes and their deployment. Coupled with the cryopreservation, the embryogenic cell lines will make establishment of hybrid clonal tests possible and facilitate the selection of resistant, well-adapted clones.

After maturation, somatic embryos of many pine species are typically germinated on the same nutrient medium as used for earlier stages of SE (Klimaszewska et al. 2016). Many somatic embryos germinate within the first 48 h on a germination medium and later develop roots; however these often cease growing and become black in spite of the continuous development and growth of the shoots. The somatic seedlings, which continue developing, require approximately five to six months of culture to grow 2-cm roots and at this time they can be potted and transferred to a greenhouse. The survival of the somatic plants after potting has been thus far unreliable and has often varied with line (genotype). As a consequence, many somatic plants do not

grow and upon close examination it became clear that the ensuing death was due to the complete cessation of the root growth and to the necrosis of the whole root.

In order to both accelerate and improve the root growth, and ensure survival of the somatic plants in a greenhouse, a study on the modification of the nitrogen (N) composition of the germination medium was conducted. Specifically, we were interested in learning how the organic or inorganic forms of N during the germination and post-germination stages affect the root growth. Consequently, the question was whether the length of the root and the germination medium residual effect would influence the survival of the somatic plants during greenhouse acclimatization. The root length was strongly affected by N composition of the germination medium as well as the expression of nitrogen-related metabolism genes. These results have potential application to generate pine somatic plants with improved ability for nutrient acquisition when potted in a soilless substrate.

3. Material and methods

3.1. Plant material

In 2008, immature seeds of F2 hybrid *Pinus strobus* x *P. wallichiana* backcrossed with *P. strobus* pollen were collected for the initiation of SE. Megagametophytes with enclosed zygotic embryos were excised from the seeds and cultured on the initiation medium (modified Litvay et al. 1985, MLV) as described in Klimaszewska et al. (2001). Two embryogenic lines; 223-2 and 270-2 were amassed, after retrieval from the cryopreserved stock (see Percy et al. 2000 for the cryopreservation protocol), by serial subcultures on medium of the same composition as above. Development and maturation of somatic embryos was conducted for 12-14 weeks without subcultures according to Klimaszewska et al. (2004). The two genotypes produced between 600 to over 900 phenotypically normal somatic embryos per g fresh mass (fm) tissue. These somatic embryos were then used in the germination experiments. Another population of somatic embryos, matured three months after the 1st germination experiment, was germinated on G1, G2 and G5

(described below) and these somatic pants were used in the greenhouse experiment (see below).

3.2. Germination of somatic embryos

The germination media were modifications of MLV, which included the absence of organic (Gln, CH)) or inorganic (KNO_3, NH_4NO_3) sources of N depicted as G1, G2, G3, G4, G5, G6 and G7. Additionally, half-strength MLV (G8) and half-strength CD (G9) (Campbell et al. 1975) with full-strength $FeSO_4$ and disodium salt of ethylenediaminetetraacetic acid (Na_2EDTA) were tested (Table 1 and Table 2). In these media no plant growth regulators were added and pH was adjusted to 5.8. About 20 somatic embryos were placed horizontally in a 90 x 15 mm Petri dish with 20 ml of medium, and 10 Petri dishes with the total of 200 somatic embryos were tested for each of the nine variants of the germination medium. In total 1800 somatic embryos per genotype were cultured in this experiment.

Table 1. Media used for germination of somatic embryos of hybrid white pine

Macronutrients	MLV (mgl^{-1})	1/2 CD (mgl^{-1})
KNO_3	950	170
$Ca(NO_3)_2$	–	340
NH_4NO_3	825	400
$CaCl_2$	8.31	–
$MgSO_4$	450	90
KH_2PO_4	170	85
KCL	–	33
$FeSO_4$	20	20
Na_2EDTA	37.3	37.3
Myo-inositol	100	250
Gln	500	–
CH	1000	–
Sucrose	58 mM	58 mM
Gellan gum	0.5%	0.5%
NO_3^+	13.5 mM	6.0 mM
NH_4^-	2.31 mM	1.13 mM

Composition of MLV standard germination medium (G1) and ½ CD (G9) modified from Campbell and Durzan (1975)

After 6 weeks of culture, approximately 90% of somatic embryos germinated and all somatic seedlings were subcultured onto fresh medium by placing them horizontally on the surface. Due to the increased size of the somatic seedlings, their number per Petri dish was reduced to 10. The Petri dishes were placed in a tilted position (at a 45° slant) in the growth cabinets at 24°C during the whole time of culture. For the first week, the germination was conducted in darkness and for the remaining 11 weeks in a 16h photoperiod at 60 µmol m^{-2} s^{-1} light intensity. After 12 weeks of germination, the root and shoot lengths of all somatic plants were measured. The experiment was repeated with 100 somatic embryos per each of the selected germination medium; G1, G2, G3, G6, and G7 that differed the most in N composition. The results from both experiments were combined for the statistical analysis of root and shoot length (see *Statistical analyses*).

Table 2. Nitrogen composition of the media used for germination of hybrid white pine somatic embryos

Germination medium	Gln (0.5 gl^{-1})	CH (1 gl^{-1})
With inorganic N		
G1	+	+
G2	-	-
G3	+	-
G4	-	+
Without inorganic N[a]		
G5	+	+
G6	+	-
G7	-	+
G8 (1/2 MLV)	+	+
G9 (1/2 CD)	-	-

[a]These media contained 0.5 gl^{-1} KCl to compensate for the withdrawal of KNO_3

3.3. Uptake of labelled nitrogen

After measurements of the root and shoot lengths, the somatic plants were placed in containers with distilled water overnight to eliminate germination medium, which had adhered to their roots. Afterwards, three replicates of 30 somatic plants per each germination medium (90 in total) were incubated in 100 ml of the nutrient solution containing ^{15}N ($^{15}NH_4Cl$ and $K^{15}NO_3$) at 250 µM for 2h at room temperature and with slow agitation. The composition of the nutrient solution was based on Bedell et al. (1999) with modifications that included omission of $CaCl_2$, addition of ^{15}N and replacement of chelated iron compound (Sigma, F-0518) with $FeSO_4$ and Na_2EDTA (Table 3). Microelements were the same as in MLV and the pH was adjusted to 4.5.

Table 3. Composition of the liquid nutrient solution for ^{15}N uptake by the somatic plants of hybrid white pine whose somatic embryos germinated for 12 weeks on media varying in N sources

Compound	mg l^{-1}	Compound	mg l^{-1}
KNO_3	62	KI	4.15
NH_4NO_3	40	H_3BO_3	31
$^{15}NH_4Cl$	6.5	$MnSO_4 \cdot H_2O$	21
$K^{15}NO_3$	12.6	$ZnSO_4 \cdot 7H_2O$	43
Na_2HPO_4 x $2H_2O$	20	$Na_2MoO_4 \cdot 2H_2O$	1.25
KCl	40	$CuSO_4 \cdot 5H_2O$	0.5
$MgSO_4$ x $7H_2O$	20	$CoCl_2 \cdot 6H_2O$	0.13
Ferrous sulfate chelate	5.6		

After 2 h, the somatic plants were rinsed in a solution of 0.5 mM $CaCl_2$ and their roots were excised, oven-dried at 70ºC for 48 h and ground into powder for the analysis of ^{15}N content. The $\delta\ ^{15}N$ ($^0/_{00}$) values were determined in triplicate at the University of Málaga Research Facility (Unit of Atomic Spectrometry) using a flash EA 1112 elemental analyser coupled to a

Delta V Advantage isotope ratio mass spectrometer (Thermo Scientific, Waltham, MA, USA), these values represent nitrogen isotopic composition of the sample relative to that of atmospheric N_2 in ‰. $\delta^{15}N$ ‰ = (*R* sample - *R* standard) / R standard) × 1000 δ‰ where *R* sample is the sample isotope ratio ($^{15}N/^{14}N$) and *R* standard is the $^{15}N/^{14}N$ ratio for atmospheric N_2. Ammonium influx was determined in roots as described by Öhlund et al. (2004). Elemental analysis of % N contents was also conducted.

3.4. Chlorophyll analysis

After the ^{15}N uptake experiment and the removal of the roots from the somatic plants for the analysis (see above), the shoots of the same plants were used for the measurements of the chlorophyll a and b contents. Three samples of 10 shoots, from each germination medium and of each genotype, were oven-dried as described above and ground to powder. Chlorophylls were extracted from 60 mg powdered shoots in 80% acetone following the method described by Graan et al. (1984) and measured spectrophotometrically based on the specific coefficients for chlorophyll a at 664 nm and chlorophyll b at 647 nm. The following equations were used to calculate the micromolar concentrations of chlorophyll a (13.19 x A_{664}) - (2.57 x A_{647}) and chlorophyll b (22.1 x A_{647}) - (5.26 x A_{664}), respectively. These were then expressed as mg g^{-1} fm based on the final volume and fm.

3.5. RNA isolation and microarray hybridization

RNA was extracted according to Canales et al. (2012a) from three samples of somatic plant shoots of 270-1 genotype that developed on G1 or G7 germination medium. Genomic DNA was removed by digestion with DNAse I (Promega Corporation, Madison, WI). RNA concentration and quality were quantified using a NanoDrop© ND-1000 spectrophotometer (NanoDrop Technologies, Inc. Wilmington, USA). For gene expression measurements a cDNA microarray (PINARRAY2) with a set of 8208 spots was used as well as positive and negative controls (Cañas et al. 2015). The spots correspond to cDNAs included in the SustainPineDB (Canales et al. 2014). Heterologous arrays have been successfully used in previous studies of gene expression in

Pinaceae (Van Zyl et al. 2002). Samples of RNA were amplified using the Amino Allyl MessageAmp II aRNA amplification Kit (Ambion) following the manufacturer´s instructions. The antisense amplified RNA was labeled with the CyeDye Post-Labelling Reactive Dye Pack (GE Healthcare, Barcelona, Spain). Prehybridization, hybridization and posthybridization washes were carried out using Pronto Universal Hybridization Kit protocol (Corning). Microarray hybridization, scanning and data acquisition were performed as described by Cañas et al. (2015).

3.6. Real-time quantitative PCR (qPCR)

To validate microarray results, selected differentially expressed genes were tested by qPCR. Synthesis of cDNA was performed using a 5X iScript^{TM} cDNA Synthesis Kit (Bio Rad, CA, USA) with 500 ng of total RNA in a volume reaction of 10 μl following manufacturer's recommendation. The qPCR was carried out in a thermal cycler CFX384 (Bio Rad, Hercules, CA) under the following conditions: 1 cycle of denaturation at 95ºC for 2 min, 40 cycles of denaturation at 95 ºC 1 s and hybridization at 60ºC for 5 s and finally 1 cycle of 30 s at 95ºC, 5 s at 65ºC, 0.5 s at 95 ºC for the generation of the dissociation curve that confirms the specific amplification of each reaction. Reactions were performed in triplicate in a total volume of 10 μl containing 5 μl of SsoFst^{TM} EvaGreen® Supermix (Bio Rad, CA, USA), 2 μl cDNA (5 ng μl^{-1}) and 0.5 μl of 10 mM specific primers. Fluorescence data fitted to the two-parameter mass action kinetic model of PCR (MAK2 model, Boggy et al. 2010), the initial concentrations of each gene were deduced from the model using the R package Ritz et al. (2008) and normalized to the geometric mean of the reference gene.

3.7. Growth in a greenhouse

The greenhouse experiment was carried out a few months later after the germination experiments, when the results from the nine tested media were obtained. Subsequently, a new population of mature somatic embryos of 223-2 and 270-1 was germinated on G1, G2 and G5 (the three contrasting media with respect to N sources) for 12 weeks. Thirty two to 45 somatic

plants of each genotype and from each medium were collected and soaked overnight in deionized water in covered containers. Subsequently they were planted in Rootrainers® growing cells (square of 12.7 cm deep and 3.8 cm wide; in a 32-cell tray) filled with moist 3:1:1 peat: perlite: vermiculite, which was also misted with N-P-K 11-41-08 (Master Plant Prod-Inc, Brampton, ON, Canada) dilution 1/100. The somatic plants were misted with 25 ppm (0.125 gl^{-1}) N-P-K 20-20-20 immediately after potting as well as 2 to 3 times during the day in the mist chamber. After 5 days the plants were placed in a greenhouse under ambient conditions and watered when required. Somatic plant survival was determined after 2 weeks in the greenhouse.

3.8. Statistical methods

Root length, shoot length, chlorophyll, root $\delta^{15}N$ and percent N (%N) were each transformed to their natural logarithm for analysis through mixed linear models. A small constant of 0.5 was added to each root length to avoid logarithms of zero. Initially, each model included fixed effects for genotype, medium and their interaction (as well as chlorophyll type and its interactions with the other factors in the case of chlorophyll), random effects for experiments and genotype × medium × experiment, and a residual error. Random effects were assumed to follow independent normal distributions with zero means and constant variances. Residual errors were also assumed normally distributed and independent from other random effects. Two exceptions were as follows. Firstly, the two types of chlorophyll were measured on the same tissue sample; this was reflected in the model by assuming that the two residual errors for the same sample followed a bivariate normal distribution with a 2×1 vector of zero means and a 2×2 unstructured variance-covariance matrix that allowed for different variances for the two types chlorophyll and non-zero covariance. Secondly, preliminary analyses suggested that the variance of the residual errors for the two nitrogen responses was not constant among media; this variance was thus allowed to differ. The random part of each model was reduced if some variance components were deemed negligible based on Akaike's information criterion (Brown and Prescott 2006). The mean of media with inorganic plus

minus organic nitrogen (G1, G2, G3, and G4) was compared with that of media with solely organic N (G5, G6 and G7). Means per medium were also compared in pairs including media G8 and G9. These tests were adjusted for multiplicity by the simulation method (Westfall et al. 1999). When there was evidence of interaction between medium effects and genotype (or genotype by chlorophyll type combinations), multiple comparisons among means per medium (or per medium and chlorophyll type) were performed for each level or combination of levels of the interacting factors. When there was not, this interaction was removed from the model to alleviate the estimation problem associated with the imbalance of the factorial treatment structure (Milliken and Johnson 2009). Means were computed on the logarithmic scale and back-transformed for presentation. Means and ratios quoted in the text are followed by their estimated 95% confidence limits in parentheses separated by a comma (lower, upper). All computations were performed with SAS (SAS Institute Inc., Cary, NC, USA), particularly the MIXED and GLIMMIX procedures.

4. Results

4.1. Root and shoot growth

Elongation of radicles in somatic embryos occurred within 48 h followed by both the hypocotyl and cotyledon elongation. The majority of the somatic embryos (90 to 95%) developed into plants after 12 weeks with the root length strongly influenced by the N composition of the germination medium while the shoots were not affected (Fig. 1a, b).

The statistical analysis of genotype × medium interaction did not reach significance and was removed from the models for root and shoot lengths (Table 4). Restricting the comparison to media (G1, G2, G3, G4, G5, G6 and G7), roots were 1.38 (1.03, 1.74) times longer on somatic plants of 270-1 than of 223-2 ($P = 0.01$) (Table 5). Roots grew about 2.55 (1.98, 3.58) times longer on media that contained solely organic N (G5, G6 and G7) than on those containing solely inorganic N or both forms of N (G1, G2, G3 and G4) ($P < 0.0001$).

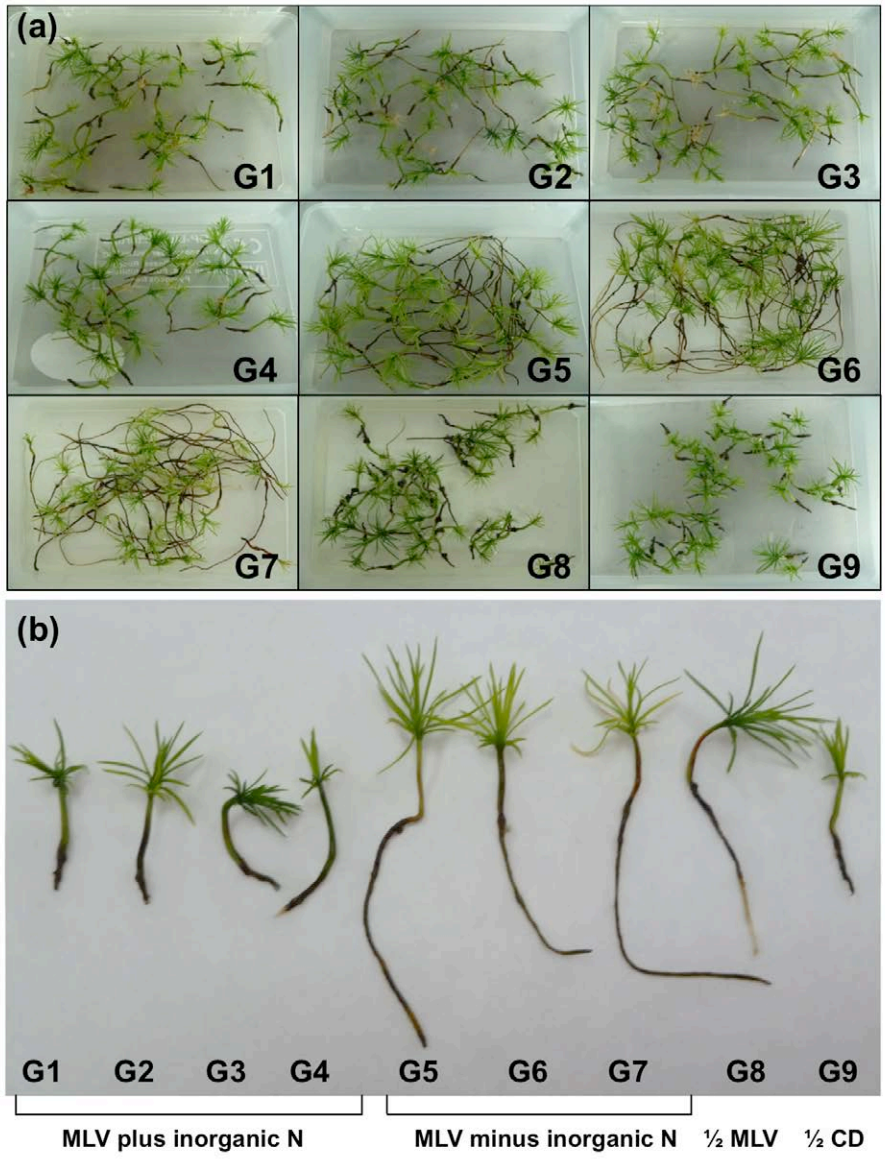


Figure 1. Somatic plants of hybrid white pine 223-2 genotype developed from somatic embryos after 12-week culture on G1-G9 germination media. The somatic plants collected en masse in water **(a)** and single somatic plants each from G1 to G9. Note roots of different lengths **(b)**. Bar =2.4 cm **(a)** and 1.7 cm **(b)**.

Table 4. Analysis of variance of the natural logarithm of root length, shoot length, chlorophyll content, $\delta^{15}N$ and %N contents in somatic plants of hybrid white pine whose somatic embryos germinated for 12 weeks on media varying in N sources

Source of variation	Root length	Shoot length	Chlorophyll	$\delta^{15}N$	%N
Genotype (G)	0.01	<.0001	0.67	<.0001	0.47
Medium (M)	<.0001	0.08	<.0001	<.0001	<.0001
G × M	0.94	0.98	<.0001	<.0001	<.0001
Type (T)			<.0001		
G × T			0.59		
M × T			<.0001		
G × M × T			<.0001		

Table 5. Effect of the nitrogen composition of the somatic embryo germination medium on the root growth of hybrid white pine somatic plants after 12 weeks of culture

Genotype	Medium	Mean	95% confidence limits	
			Lower	Upper
223-2		7.44	2.99	17.53
270-1		10.30	3.94	25.72
	G1-G4	5.78	2.16	14.31
	G5-G7	14.76	6.27	33.87
	G1	6.84	2.93	15.21
	G2	5.02	2.08	11.32
	G3	5.61	2.47	12.08
	G4	5.77	2.52	12.48
	G5	20.57	9.56	43.61
	G6	14.51	6.79	30.38
	G7	23.47	11.15	48.82
	G8	6.96	3.10	14.96
	G9	6.85	3.05	14.72

Back-transformed mean root lengths (mm) and their 95% confidence limits per genotype, per medium group and per each medium (G1 to G4; inorganic N ± organic N, G5 to G7; solely organic N, G8; ½ MLV and G9; ½ CD

Mean root lengths among the media were between 5.02 mm (2.08, 11.32) to 23.47 mm (11.15, 48.82) and the significance of all pair wise differences between mean root lengths per medium is shown in Table S1. The mean root lengths on G8 (1/2 MLV) and G9 (1/2 CD) media were highly similar in spite of the difference in their total N concentration (Table 5). The roots grew the most on G7 and G5, which contained solely CH or both CH and Gln, respectively. Variations in germination media had no apparent effect on shoot length (P = 0.09) whose means were 16.2 and 19.6 mm in 270-1 and 223-2 respectively, 1.2 (1.13, 1.30) times longer in 223-2 than those in 270-1 (P < 0.0001) (Table 6).

Table 6. Effect of the nitrogen composition of the somatic embryo germination medium on the shoot growth of hybrid white pine somatic plants after 12 weeks of culture

Genotype	Medium	Mean	95% confidence limits	
			Lower	Upper
223-2		19.69	14.82	26.15
270-1		16.29	11.99	22.14
	G1-G4	18.10	13.33	24.53
	G5-G7	17.77	13.26	23.74
	G1	18.89	14.50	24.61
	G2	17.98	13.80	23.43
	G3	19.87	15.54	25.41
	G4	17.74	13.96	22.55
	G5	18.91	14.67	24.38
	G6	18.73	14.65	23.95
	G7	17.69	13.83	22.61
	G8	16.05	12.63	20.40
	G9	15.04	11.83	19.12

4.2. Uptake of labelled inorganic N

Evidence that the interaction between genotype and medium affected both $\delta^{15}N$ in root tissues and %N led to conduct pair wise comparisons among media for each genotype ($P < 0.0001$) for both $\delta^{15}N$ and %N (Tables S2, S3). Over media G1-G7, mean $\delta^{15}N$ of root tissue from 223-2 was 1.2 (1.14, 1.26) times higher than that of root tissue from 270-1 ($P < 0.0001$) (Table 7).

Mean $\delta^{15}N$ content of the 223-2 roots grown on media with solely organic N was 4.3 (3.9, 4.8) times higher than that from media with inorganic N ± organic N ($P = 0.0008$). For roots of 270-1, the corresponding ratio was 2.36 (2.20, 2.52) ($P = 0.01$) (Table 7). This suggested that roots grown in the absence of inorganic N were capable of rapid uptake of labelled inorganic ^{15}N during the 2 h incubation in the nutrient solution. Mean of %N was higher in root tissue grown on media with inorganic N than on those with solely organic N (Table 8). The ratios (inorganic to organic N, for this response) were 2.05 (1.83, 2.30) for the 223-2 genotype ($P = 0.02$), and 1.55 (1.13, 1.18) for the 270-1 genotype ($P = 0.03$).

4.3. Chlorophyll content in shoots of somatic plants

The three-way interaction between genotype, media and chlorophyll type was highly significant ($P < 0.0001$) (Table 4). Among shoots of somatic plants of 223-2, type a chlorophyll was 1.21 (1.12, 1.30) times more abundant in shoot tissue grown on media with inorganic N ±organic N than in tissue grown on media with solely organic N ($P = 0.0017$) (Table S4). The effect of inorganic N on type b chlorophyll in shoots from the same genotype was also significant ($P = 0.03$) but not as strong as that for type a; the ratio of mean type b chlorophyll content between inorganic and solely organic N based media was 1.15 (1.07, 1.24). In shoots from genotype 270-1, the corresponding ratios were 1.23 (1.16, 1.31) for type a chlorophyll ($P < 0.0001$) and 1.21 (1.14, 1.29) for type b chlorophyll ($P < 0.0001$). Pair wise comparisons among media were conducted for each combination of genotype and chlorophyll type (Table S5).

Table 7. ^{15}N content in the roots of somatic plants of hybrid white pine whose somatic embryos germinated on media with varying sources of N for 12 weeks

Genotype	Medium	Mean	95% confidence limits	
			Lower	Upper
223-2		218.74	210.38	227.43
270-1		156.93	153.11	160.86
223-2	G1-G4	118.13	110.43	126.37
	G5-G7	512.62	469.47	559.74
270-1	G1-G4	114.47	108.64	120.60
	G5-G7	270.02	256.57	284.17
	G1	105.75	96.301	116.12
	G2	120.32	104.92	137.99
	G3	118.97	106.96	132.32
	G4	120.80	120.32	121.29
	G5	404.80	402.46	407.16
	G6	336.05	284.56	396.86
	G7	678.96	652.84	706.13
	G8	176.53	175.89	177.17
	G9	170.14	169.33	170.96

Back-transformed mean shoot length (mm) and their 95% confidence limits per genotype and medium (G1 to G4; inorganic N ± organic N, G5 to G7; solely organic N, G8; ½ MLV and G9; ½ CD)

4.4. Gene expression analysis in somatic plants

As the absence of inorganic N in the germination media promoted the growth of longer roots, gene expression analysis was conducted in somatic plants to further understand the molecular basis of the observed differences. Intact RNA was extracted from somatic plants of the 270-1 genotype germinated on media with (G1) and without (G7) inorganic N and the microarray analyses were performed to identify differentially expressed genes under these two contrasting germination conditions. The results represent relative expression levels comparing G7 shoots with G1 (the standard germination medium) (Table S6).

Table 8. %N content in the roots of somatic plants of hybrid white pine whose somatic embryos germinated on media with varying sources of N for 12 weeks.

			95% confidence limits	
Genotype	Medium	Mean	Lower	Upper
223-2		4.96	4.77	5.16
270-1		5.23	5.10	5.36
223-2	G1-G4	6.60	6.46	6.75
	G5-G7	3.22	2.87	3.61
270-1	G1-G4	6.23	6.12	6.34
	G5-G7	4.02	3.76	4.30
	G1	6.79	6.72	6.87
	G2	5.70	5.40	6.02
	G3	6.46	6.31	6.63
	G4	6.76	6.64	6.88
	G5	3.74	3.71	3.78
	G6	3.33	2.69	4.12
	G7	3.12	2.84	3.42
	G8	5.80	5.70	5.90
	G9	5.11	5.03	5.20

Back-transformed %N means and their 95% confidence limits per genotype, medium group and each medium (G1 to G4; inorganic N ± organic N, G5 to G7; solely organic N, G8; ½ MLV and G9; ½ CD)

From a total of 5987 differentially expressed genes, 250 genes were differentially expressed by two fold; 181 genes were upregulated and 69 genes were downregulated in somatic plants germinated on G7 compared with those from G1. Figure 2 shows the distribution of differentially expressed genes in functional categories (Figure 2a) and a Mapman diagram of the genes upregulated or repressed in the main metabolic pathways (Figure 2b). Photosynthesis, protein biosynthesis, energy production, nitrogen metabolism and secondary metabolism were the functional categories containing the highest number of differentially expressed genes. Genes involved in light reactions, Calvin cycle, photorespiration, UDP-sugar metabolism and starch biosynthesis were upregulated in G7 somatic plants.

Metabolic pathways that were activated included glycolysis, ammonium assimilation and amino acid biosynthesis, transamination and the biosynthesis of flavonoids and phenylpropanoids. Downregulated genes in G7 somatic plants were those involved in cell wall formation, glutathione metabolism and tetrapyrrol synthesis.

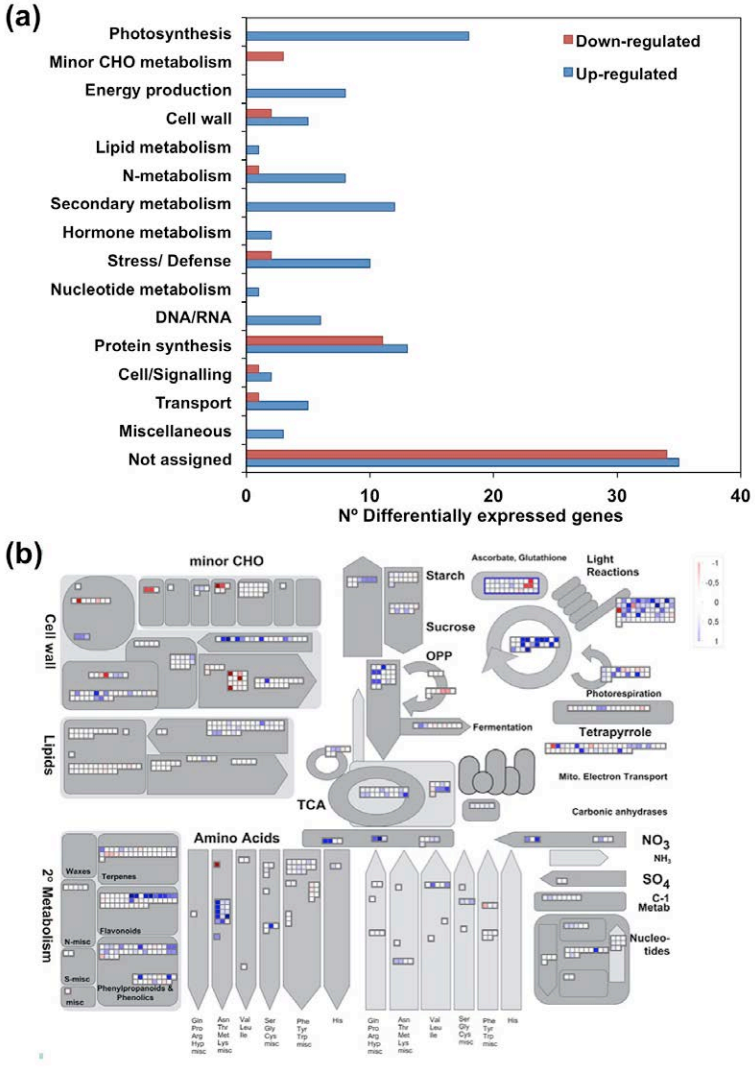


Figure 2. Microarrays of 270-1 somatic plants germinated on G1 (with inorganic and organic N) or on G7 medium (solely CH) showing the distribution of differentially expressed genes in functional categories (a) and a Mapman diagram of the genes upregulated or repressed in the main metabolic pathways (b). Blue scale indicates up-regulated genes and the red scale indicates down-regulated genes in the shoots of G7 somatic plants compared with shoots from G1 medium.

4.5. Validation of transcriptome analysis

To validate the transcriptome profiles, 10 differentially expressed genes were selected for accurate determination of their expression levels using qPCR analysis. These genes are mainly involved in photosynthesis, energy production, N-metabolism, stress and protein synthesis. *PpChS* (chalcone synthase), *Pp40sRP* (40s ribosomal protein s19), *PpPSII* (photosystem ii 10 kDa chloroplast), *PpChl* (chlorophyll a b-binding protein), *PpGS1a* (glutamine synthetase 1a) and *PpGS1b* (glutamine synthetase 1b) genes were highly-expressed in G7, and *PpEAP* (embryo-abundant protein), *Pp24kdP* (24 kDa seed maturation protein), *PpNADH* (NADH:ubiquinone oxidoreductase family protein), *PpEF1*β (elongation factor 1-beta) that were down-regulated genes in G7 had the same expression levels, albeit with slight variations, in both microarray and qPCR analysis (Figure S1). Overall, the results of the qPCR expression analysis supported microarray data. Sequences of specific primers are listed in Table S7.

4.6. Growth of somatic plants in a greenhouse

Somatic plants developed on three contrasting media G1 (inorganic and organic N), G2 (solely inorganic N) and G5 (solely organic N) were grown in a greenhouse to test the influence of the root length and N nutrition during somatic embryo germination and development on their acclimatization and survival. Of 45 somatic plants each of 223-2 and 270-1 that were planted and acclimatized in a greenhouse, 64 and 59%, respectively continued to grow when germinated on G5 (Table 9). Conversely, 33% of 223-2 and 25 to 26% of 270-1 somatic plants from G1 and G2 survived, respectively. Measurements of the root and shoot lengths, and dry mass (dm) of all the surviving plants after 2 weeks in the potting mix were highly similar (data not shown). This suggested that apparently some somatic plants could compensate for the initial slow root growth (or cessation of the root growth) with rapid growth under standard fertilization regime in a greenhouse. Our observation also suggested a trend that the highest survival of somatic plants might be linked with the longest roots at the time of potting.

Table 9. Survival of somatic plants of hybrid white pine, whose somatic embryos were germinated on media varying in N sources for 12 weeks, after 2 weeks in a greenhouse (plants were misted with N:P:K 20:20:20 fertilizer)

Genotype	Germination medium	Number of growing somatic plants / planted somatic plants	Survival %
223-2	G1	15 / 45	33
	G2	15 / 45	33
	G5	29 / 45	64
270-1	G1	11 / 43	25
	G2	9 / 35	26
	G5	25 / 42	59

5. Discussion

The results of our study involving hybrid white pine somatic plants corroborated the contention that the root growth and architecture are dependent, among other factors, on soil composition and particularly on water and mineral nutrients availability (Hodge et al. 2009 and references therein; Lima et al. 2010). Using clonal hybrid white pine somatic embryo-derived plants of two genotypes and in the effort to improve root growth for more efficient acclimatization in a greenhouse we varied the N composition of the germination medium. Conifer somatic embryos are most often germinated on the same media (but without the plant growth regulators) as used for all earlier stages of SE or occasionally on different ones (Klimaszewska et al. 2016 and references therein). The germination of somatic embryos and conversion to somatic plants occurs *in vitro* at various frequencies and the roots are of highly variable lengths. In some somatic seedlings the roots, after initial elongation, cease growing altogether. In the preliminary experiments, we observed that somatic hybrid white pine plants that grew long roots (2 cm and longer) survived acclimatization in a greenhouse at higher numbers. To produce a large number of somatic plants of hybrid white pine in a relatively

synchronized fashion we investigated whether varying the inorganic and organic (amino acids) N composition of the MLV medium would influence the root architecture. In addition we tested half-strength CD medium (G9), which has half of the ionic concentration of NO_3^- and NH_4^+, lacks organic N compounds compared with MLV and also has been used in tissue culture of white spruce (Campbell et al. 1975).

It has been shown that conifers have preference for the uptake of NH_4^+ over that of NO_3^- but can also take up amino acids, which greatly broadens the range of potential N sources for their growth (Näsholm et al. 2009; Castro-Rodríguez et al. 2017). Öhlund and Näsholm (2001) tested organic and inorganic nitrogen compounds on growth and mineral nutrient concentrations of Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*) seedlings. The seedlings were watered with nutrient solutions differing in ammonium, nitrate, arginine and glycine composition at three total nitrogen concentrations: 1, 3 and 10 mM.. The gas chromatography–mass spectrometry analyses of the seedlings from this 100-day lasting experiment showed that the uptake of arginine and glycine was comparable to the uptake of inorganic N in both conifer species.

In the present study, the somatic embryo germination/growth media included two main groups: G1 (standard medium), G2, G3, G4 all contained inorganic N and G5, G6, G7 contained solely organic N. In addition, G8 was half-strength of G1 (with organic N) and G9 was half-strength CD (Campbell et al. 1975) without organic N. The roots of somatic plants growing on media containing solely organic N grew about 2.55 (1.98, 3.58) times longer than on those containing in addition organic or solely inorganic N ($P<0.0001$). However, there was also a difference between genotypes with the roots of 270-1 being 1.38 (1.03, 1.74) times longer than of 223-2 ($P = 0.0158$). Overall, the longest roots grew on G7 supplemented with solely CH and on G5 supplemented with both CH and Gln; 23.47 mm (11.15, 48.82) and 20.57 mm (9.56, 43.61), respectively.

N is critical for plant production and it is required in large amounts during development. The availability of different N sources, together with carbon, was implicated in a differential biomass distribution between roots

and shoots (Bauer et al. 2001; Bown et al. 2010). Moreover, it is well known that the architecture of the root has a profound effect on the capacity of plants to take up nutrients and nutrient uptake is enhanced through the regulation of root growth. For example, increasing nitrate availability reduces primary root elongation (Linkohr et al. 2002). In contrast, low N availability stimulates primary root elongation (Gruber et al. 2013; Wei et al. 2013). Somatic plants from G5, G6, and G7 media exhibited a similar response and developed longer roots than those from media supplied with both inorganic and organic N sources. This implies a considerable investment of organic N and therefore these somatic plants should have an appropriate level of internal N to sustain active root growth. The molecular components that regulate the root architecture in response to N availability have been identified in *Arabidopsis*. In the last few years, research from various laboratories showed that N limitation induces the expression of a complex gene network that includes hormonal genes, in particular those related to auxin production (Kiba et al. 2016).

Microarray analysis revealed that, depending on the N source in the germination medium, somatic embryo plants displayed changes in the transcriptome, resulting in the differential expression of a range of genes involved in different processes. A total of 252 genes were differentially regulated in G7 when compared with G1 somatic plants as a reference. Highly expressed genes included those involved in essential processes for plant growth and development such as photosynthesis, production of metabolic energy, nitrogen metabolism and pathways of secondary metabolism. These results strongly suggest a high demand for organic N that is necessary to support the biosynthesis of nitrogenous compounds for plastid metabolism and vascular development. Two key genes encoding cytosolic glutamine synthetases, *GS1a* and *GS1b* (Avila et al. 2001), were highly expressed in G7 plants suggesting that glutamine biosynthesis is particularly active in these somatic plants. The repression of an asparagine synthetase gene (*AS*, Cañas et al. 2006; Canales et al. 2012b) is also consistent with the above findings suggesting a high ratio of glutamine/asparagine biosynthesis, under conditions of no carbon limitation, to meet the metabolic requirements

for growth in the G7 somatic plants. Recently, it has been reported that the internal N status in plants can be monitored by sensing glutamine levels through the N sensor protein PII (Chellamuthu et al. 2014).

The labelled ^{15}N experiment showed that roots of 223-2 and 270-1 grown in the absence of inorganic N were capable of its rapid uptake, which was 4.3 (3.9, 4.8) ($P = 0.0008$) and 2.36 (2.2, 2.52) ($P = 0.01$) times higher, respectively than that from media with inorganic N. These results indicate that somatic embryo plants germinated in the absence of inorganic N exhibited enhanced ability to uptake N in the form of nitrate and ammonium. High-affinity ammonium transporters (AMT) and nitrate transporters (NRT2/NRT3) have been identified in pine roots that should play an important role in the uptake and transport of these externally supplied N sources (Castro-Rodríguez et al. 2016; 2017).

The shoots were not affected by any of the tested medium and were 1.2 (1.13, 1.3) times longer in 223-2 than in 270-1 ($P = 0.0001$). However, shoots grown on media without inorganic N were slightly chlorotic (had lower amount of chlorophyll a and b) with the most severe symptoms occurring on G7. These findings are consistent with the repression of genes involved in tetrapyrrol biosynthesis that was observed in G7 somatic plants. The lower content of chlorophyll observed in the shoots of somatic plants grown on media without inorganic N is likely reflecting a limitation of N considering that large amounts of N should be allocated to sustain enhanced root growth. For the greenhouse acclimatization experiments, plants germinated on G1, G2 and G5 were chosen and revealed that of the three tested media the highest survival was demonstrated by the plants from G5 medium containing solely organic N sources. The survival was 64 and 59% for 223-2 and 270-1, respectively, which was approximately 50% higher than on the other two media.

6. Conclusions

This study demonstrated that the somatic plants of F2 hybrid eastern white pine were able to develop longer functional roots in the sole presence of organic N, compared with the standard germination medium, as determined by an efficient uptake of $K^{15}NO_3$ and $^{15}NH_4NO_3$. The somatic plants germinated on medium with CH and Glu (G5) survived at 50% higher rate in a greenhouse than those from G1 (inorganic and organic N sources) or G2 (solely inorganic N) media after 2 weeks. These results are encouraging and should be followed by a long-term greenhouse / field experiment to conclude whether N nutrition, during germination, will affect the later growth and phenotype of the somatic trees.

7. Acknowledgements

This work was supported by grants from the Spanish Ministerio de Economía y Competitividad (BIO2015-69285-R) and Junta de Andalucía (BIO-474) to M.T.L.L. We gratefully acknowledge the assistance of Dr. Javier Canales (University of Malaga) with microarray analysis, Mrs. Cathy Overton with the production of mature somatic embryos and acclimatization of somatic plants and Mrs. Michèle Bernier-Cardou for the statistical analyses (Natural Resources Canada, Canadian Forest Service).

8. Author contribution statement

K.K., F.M.C. and C.A. designed the research. M.T.LL. and K.K. carried out the experiments. All the authors analysed data and wrote the manuscript. All authors read and approved the final manuscript. The authors declare that they have no conflict of interest.

9. Supporting information

Table S1. Multiple comparisons among mean root lengths per medium in hybrid white pine somatic plants after 12- week growth on media differing in N sources. Only significant differences are shown. *Adjusted P<0.05

Medium	G5	G6	G7	G8	G9
G1	*		*		
G2	*	*	*		
G3	*	*	*		
G4	*		*		
G5				*	*
G6					
G7				*	*
G8					

Table S2. Comparisons among $\delta^{15}N$ means in roots of hybrid white pine somatic plants per medium for each genotype (223-2 upper part of the table, 270-1 lower part of the table). Only significant differences are shown. *Adjusted P<0.05

Medium	G5	G6	G7	G8	G9
G1	*	*	*	*	
G2	*	*	*		
G3	*	*	*	*	
G4	*	*	*	*	*
G5				*	*
G6					*
G7				*	*
G8					*

Medium	G5	G6	G7	G8	G9
G1	*	*	*		
G2	*		*		
G3	*		*		*
G4	*		*	*	*
G5			*	*	*
G6			*	*	*
G7				*	*
G8					*

Table S3. Comparisons among %N means in roots of hybrid white pine plants per medium for 223-2 (upper part of the table) and 270-1 (lower part of the table) genotype. Only significant differences are shown. *Adjusted P<0.05

Medium	M5	M6	M7	M8	M9
G1	*		*	*	*
G2	*		*		
G3	*		*		*
G4	*		*	*	*
G5				*	*
G6					
G7				*	
G8					*
G1	*		*	*	
G2	*		*		
G3	*		*		*
G4	*		*		*
G5				*	*
G6					
G7				*	*
G8					*

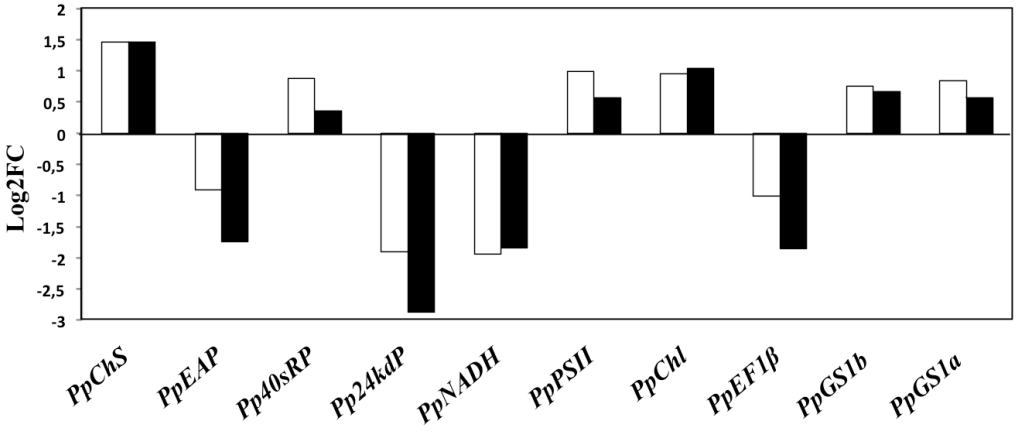


Figure S1. Comparison between microarray and qPCR expression data to validate the microarray hybridizations. Closed bars correspond to Log2FC from microarray data and open bars to Log2FC from qPCR ones. Selected genes from *P. pinaster* (SustainPineDB) were: *PpChS* (chalcone synthase), *PpEAP* (embryo-abundant protein), *Pp40sRP* (40s ribosomal protein s19), *Pp24kdP* (24 kDa seed maturation protein), *PpNADH* (nadh:ubiquinone oxidoreductase family protein), *PpPSII* (photosystem ii 10 kDa chloroplast), *PpChl* (chlorophyll a b-binding protein), *PpEF1*β (elongation factor 1-beta), *PpGS1a* (glutamine synthetase 1a) and *PpGS1b* (glutamine synthetase 1b)

Table S4. Back-transformed means and their 95% confidence limits of chlorophyll content in shoots of hybrid white pine somatic plants per type of chlorophyll and medium, and per genotype.

Effect	Genotype	Type	Medium	Mean	95% confidence limit	
					Lower	Upper
medium*type		a	G1	0.23	0.21	0.24
		b		0.11	0.11	0.12
		a	G2	0.26	0.24	0.28
		b		0.13	0.12	0.14
		a	G3	0.22	0.20	0.23
		b		0.11	0.11	0.12
		a	G4	0.21	0.20	0.23
		b		0.11	0.10	0.12
		a	G5	0.17	0.16	0.18
		b		0.09	0.08	0.10
		a	G6	0.21	0.20	0.23
		b		0.11	0.10	0.11
		a	G7	0.18	0.17	0.19
		b		0.10	0.09	0.10
		a	G8	0.24	0.22	0.26
		b		0.12	0.11	0.13
		a	G9	0.24	0.22	0.26
		b		0.12	0.11	0.13
type	Genotype	a	Medium	0.22	0.21	0.22
		b		0.11	0.11	0.11
gen*med*type	223-2	a	Inorganic N	0.22	0.21	0.23
		b		0.11	0.11	0.12
gen*med*type	270-1	a	Inorganic N	0.24	0.23	0.25
		b		0.12	0.11	0.13
gen*med*type	223-2	a	Organic N	0.18	0.17	0.19
		b		0.10	0.09	0.10
gen*med*type	270-1	a	Organic N	0.19	0.18	0.20
		b		0.10	0.09	0.10

gen*med*type	223-2	a	G1	0.22	0.20	0.25
		b		0.11	0.10	0.13
		a	G2	0.26	0.24	0.29
		b		0.13	0.11	0.14
		a	G3	0.24	0.22	0.27
		b		0.12	0.11	0.13
		a	G4	0.17	0.15	0.19
		b		0.09	0.08	0.10
		a	G5	0.19	0.17	0.21
		b		0.10	0.09	0.11
		a	G6	0.21	0.19	0.24
		b		0.11	0.10	0.12
		a	G7	0.15	0.14	0.17
		b		0.08	0.08	0.09
		a	G8	0.26	0.23	0.28
		b		0.12	0.11	0.14
		a	G9	0.25	0.23	0.28
		b		0.13	0.11	0.14
gen*med*type	270-1	a	G1	0.23	0.22	0.24
		b		0.12	0.11	0.12
		a	G2	0.26	0.23	0.29
		b		0.13	0.11	0.14
		a	G3	0.19	0.17	0.21
		b		0.10	0.09	0.12
		a	G4	0.27	0.24	0.30
		b		0.14	0.12	0.15
		a	G5	0.15	0.14	0.17
		b		0.08	0.07	0.09
		a	G6	0.21	0.20	0.23
		b		0.10	0.10	0.11
		a	G7	0.22	0.20	0.24
		b		0.11	0.10	0.12
		a	G8	0.22	0.20	0.25
		b		0.11	0.10	0.12
		a	G9	0.22	0.20	0.25
		b		0.12	0.11	0.13

Table S5 Multiple comparisons among chlorophyll means in hybrid white pine somatic plants per medium for each genotype (223-2 and 270-1 the two upper tables, respectively) and type of chlorophyll a (3rd table) and chlorophyll b (4th table). Only significant differences are shown.

Medium	G3	G4	G5	G6	G7	G8	G9
G1		*			*		
G2		*	*		*		
G3		*			*		
G4						*	*
G5						*	*
G6					*		
G7						*	*
G8							
G1					*		
G2		*			*		
G3		*			*		
G4						*	*
G5							
G6					*		
G7						*	*
G8							
G1			*				
G2	*		*				
G3		*	*				
G4			*	*			
G5				*	*	*	*
G6							
G7							
G8							
G1			*				
G2			*				
G3		*					
G4			*	*			
G5				*	*	*	*
G6							
G7							
G8							

Table S6 Table of microarray results (excel)

Table S7 Table of primer sequences used for microarray validation (excel)

10. References

Avila C, Suárez MF, Gómez-Maldonado J, Cánovas FM (2001) Spatial and temporal expression of two cytosolic glutamine synthetase genes in Scots pine: functional implications on nitrogen metabolism during early stages of conifer development. Plant J 25:93-102.

Bauer GA, Berntson GM (2001) Ammonium and nitrate acquisition by plants in response to elevated CO_2 concentration: The roles of root physiology and architecture. Tree Physiol 21:137-144.

Bedell JP, Chalot M, Garnier A, Botton B (1999) Effects of nitrogen source on growth and activity of nitrogen-assimilating enzymes in Douglas-fir seedlings. Tree Physiol 19:205-210.

Boggy GJ, Woolf PJ (2010) A Mechanistic Model of PCR for Accurate Quantification of Quantitative PCR Data. PLoS ONE 5(8): e12355

Brown H, Prescott R (2006) Applied Mixed Models in Medicine. 2nd edition, John Wiley & Sons, Chichester, UK.

Bown HE, Watt MS, Clinton PW, Mason EG (2010) Influence of ammonium and nitrate supply on growth, dry matter partitioning, N uptake and photosynthetic capacity of *Pinus radiata* seedlings. Trees 24:1097-1107.

Campbell RA, Durzan DJ (1975) Induction of multiple buds and needles in tissue culture of *Picea glauca*. Can J Bot 53:1652-1657.

Canales J, Avila C, Cantón FR, Pacheco-Villalobos D, Díaz-Moreno S, Ariza D, Molina-Rueda JJ, Navarro-Cerrillo RM, Gonzalo Claros MG, Cánovas FM (2012a) Gene expression profiling in the stem of young maritime pine trees: detection of ammonium stress-responsive genes in the apex. Trees 26:609-619.

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

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

Cañas RA, de la Torre F, Cánovas FM, Cantón FR (2006) High levels of asparagine synthetase in hypocotyls of pine seedlings suggest a role of the enzyme in re-allocation of seed-stored nitrogen. Planta 224:83-95.

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

Castro-Rodríguez V, Assaf-Casals I, Pérez-Tienda J, Fan X, Avila C, Miller A, Cánovas FM (2016) Deciphering the molecular basis of ammonium

uptake and transport in maritime pine. Plant Cell Environ 39:1669-82.

Castro-Rodríguez V, Cañas RA, de la Torre FN, Pascual MB, Avila C, Cánovas FM (2017) Molecular fundamentals of nitrogen uptake and transport in trees. J Exp Bot 68:2489-2500.

ChellamuthuVR, Ermilov E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, Hartmann MD, Forchhammer KA (2014) Widespread glutamine-sensing mechanism in the plant kingdom. Cell 159: 1188-1199.

Daoust G, Klimaszewska K, Plourde D (2009) Somatic embryogenesis, a tool for accelerating the selection and deployment of hybrids of eastern white pine (*Pinus strobus*) and Himalayan white pine (*Pinus wallichiana*) resistant to white pine blister rust (*Cronartium ribicola*). In: Noshad David; Noh Eun Woon; King, John; Sniezko, Richard A. (eds) Breeding and Genetic Resources of Five-Needle Pines. Proceedings of the Conference 2008, Yangyang, Korea. Korea Forest Research Institute, Seoul 104p. ISBN 978-89-8176-605-4 (93520).

Garin E, Bernier-Cardou M, Isabel N, Klimaszewska K, Plourde A (2000) Effect of sugars, amino acids, and culture technique on maturation of somatic embryos of *Pinus strobus* on medium with two gellan gum concentrations. Plant Cell Tiss Org Cult 62:27-37.

Graan T, Ort DR (1984) Quantitation of the rapid electron donors to P700, the functional plastoquinone pool, and the ratio of the photosystems in spinach chloroplasts. J BiolChem. 259:4003-14010.

Gruber BD, Giehl RFH, Friedel S, von Wiren N (2013) Plasticity of the Arabidopsis root system under nutrient deficiencies. Plant Physiol 163:161-179.

Hodge A, Berta G, Doussan C, Merchan F, Crespi M (2009) Plant root growth, architecture and function. Plant Soil 321:153-187.

Kiba T, Krapp A (2016) Plant nitrogen acquisition under low availability: regulation of uptake and root architecture. Plant Cell Physiol 57: 707-714.

Klimaszewska K, Bernier-Cardou M, Cyr DR, Sutton BCS (2000) Influence of gelling agents on culture medium gel strength, water availability, tissue water potential, and maturation response in embryogenic cultures of *Pinus strobus* L. In Vitro Cell Dev Biol-Plant 36:279-286.

Klimaszewska K, Park Y-S, Overton C, Maceacheron I, Bonga JM (2001) Optimized somatic embryogenesis in *Pinus strobus* L. In Vitro Cell Dev Biol-Plant 37:392-399.

Klimaszewska K, Morency F, Jones-Overton C, Cooke J (2004) Accumulation pattern and identification of seed storage proteins in zygotic embryos of *Pinus strobus* and in somatic embryos from different maturation treatments. Physiol Plant 121:682-690.

Klimaszewska K, Hargreaves C, Lelu-Walter MA, Trontin JF (2016) Advances in conifer somatic embryogenesis since year 2000. In: Germanà MA,

Lambardi M (eds) In vitro embryogenesis in higher plants, Humana Press, Hatfield, Hertfordshire, pp 131-166.

Lima JE, Kojima S, Takahashi H, von Wirén N (2010) Ammonium triggers lateral root branching in Arabidopsis in an AMMONIUM TRANSPORTER1;3-dependent manner. The Plant Cell 22:3621-3633.

Linkohr BI, Williamson LC, Fitter AH, Leyser O (2002) Nitrate and phosphate availability and distribution have different effects on root system architecture of Arabidopsis. Plant J 29:751-760.

Litvay JD, Verma DC, Johnson MA (1985) Influence of a loblolly pine (*Pinus taeda* L.) culture medium and its components on growth and somatic embryogenesis of the wild carrot (*Daucus carota* L.). Plant Cell Rep 4:325-328.

Milliken GA, Johnson DE (2009) Analysis of messy data, Volume 1, Designed experiments, 2nd edition, CRCPress, Boca Raton FL. USA.

Näsholm T, Kielland K, Ganeteg U (2009) Uptake of organic nitrogen by plants. New Phytologist 182:31-48.

Öhlund J, Näsholm T (2001) Growth of conifer seedlings on organic and inorganic nitrogen sources. Tree Physiol 21:1319-1326.

Öhlund J, Näsholm T (2004) Regulation of organic and inorganic nitrogen uptake in Scots pine (*Pinus sylvestris*) seedlings. Tree Physiol 24:1397-1402.

Percy RE, Klimaszewska K, Cyr DR (2000) Evaluation of somatic embryogenesis for clonal propagation of western white pine. Can J For Res 30:1867-1876.

Ritz C, Spiess AN (2008) qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. Bioinformatics 24:1549-51.

Van Zyl L, von Arnold S, Chen Y, Egertsdotter U, Mackay J, Sederoff R, Shen J, Zelena L, Clapham DH (2002) Heterologous array analysis in *Pinaceae*: hybridization of *Pinus taeda* cDNA arrays with cDNA from needles and embryogenic cultures of *P. taeda*, *P. sylvestris* or *Picea abies.* Comp Funct Genomics 3:306-318.

Wei H, Yordanov YS, Georgieva T, Li X, Busov V (2013) Nitrogen deprivation promotes *Populus* root growth through global transcriptome reprogramming and activation of hierarchical genetic networks. New Phytologist 200: 483-497.

Westfall P, Tobias RD, Rom D, Wolfinger RD, Hochberg RD (1999) Multiple comparisons and multiple tests using the SAS system. SAS Institute Inc., Cary, NC, USA.

ARTÍCULO 2

PpNAC1, a main regulator of phenylalanine biosynthesis and utilization in maritime pine

Pascual MB, **Llebrés MT**, Craven-Bartle B, Cañas RA, Cánovas FM, Avila C

Plant Biotechnology Journal (2017), doi: 10.1111/pbi.12854

María Teresa Llebrés ha contribuido en los resultados relacionados con la producción y mantenimiento de las líneas transgénicas de *P. pinaster*

1. Summary

The transcriptional regulation of phenylalanine metabolism is particularly important in conifers, long-lived species that use large amounts of carbon in wood. Here, we show that the *Pinus pinaster* transcription factor, *PpNAC1,* is a main regulator of phenylalanine biosynthesis and utilization. A phylogenetic analysis classified PpNAC1 in the NST proteins group, and was selected for functional characterization. *PpNAC1* is predominantly expressed in the secondary xylem and compression wood of adult trees. Silencing of *PpNAC1* in *P. pinaster* results in the alteration of stem vascular radial patterning and the downregulation of several genes associated with cell wall biogenesis, and secondary metabolism. Furthermore, trans-activation and EMSA analyses showed that *PpNAC1* is able to activate its own expression and *PpMyb4* promoter, while PpMyb4 is able to activate *PpMyb8*, a transcriptional regulator of phenylalanine and lignin biosynthesis in maritime pine. Together, these results suggest that *PpNAC1* is a functional ortholog of the *Arabidopsis* SND1 and NST1 genes and support the idea that key regulators governing secondary cell wall formation could be conserved between gymnosperms and angiosperms. Understanding the molecular switches controlling wood formation is of paramount importance for fundamental tree biology and pave the way for applications in conifer biotechnology.

2. Introduction

Conifers like other woody plant species irreversibly immobilize large quantities of carbon skeletons in wood during their long life cycles. Particularly in these species, the metabolism of phenylalanine, precursor amino acid for lignin, is essential for secondary cell wall biosynthesis and must be finely regulated at transcriptional level (Pascual *et al*., 2016). To elucidate the transcriptional network controlling wood formation in conifers is of paramount importance for future applications in fundamental tree biology and biotechnology.

Several transcription factors (TF) regulating the biosynthesis of secondary cell wall components have been described in poplar (Sterky *et al*., 2004), eucalyptus (Goicoechea *et al*., 2005; Rengel *et al*., 2009), white spruce (Pavy *et al*., 2005) and pine (Allona *et al*., 1998; Lorenz and Dean, 2002; Villalobos *et al*., 2012; Patzlaff *et al*., 2003a, b; Gómez-Maldonado *et al*., 2004; Bedon *et al*., 2007; Craven-Bartle *et al*., 2013). Focusing on the transcriptional network, the best wood associated transcription factors characterized belong to the MYB and NAC families. Both transcription factor families are the most represented in plants with over 100 members in Arabidopsis, rice, and poplar (Martin and Paz-Ares, 1997; Ooka *et al*., 2003; Hu *et al*., 2010; Nuruzzaman *et al*., 2010). However, the NAC TF family in conifers is underrepresented with over 37 members (Pascual *et al*., 2015).

In *P. taeda,* PtMyb1, PtMyb4, and PtMyb8 activate transcription of genes involved in phenylpropanoid and lignin biosynthesis through the binding to AC elements present in their promoter regions (Patzlaff *et al*., 2003a, b; Gómez-Maldonado *et al*., 2004; Craven-Bartle *et al*., 2013). In addition, the overexpression of these Myb TFs led to ectopic lignin deposition producing plants with an increased secondary-wall thickening (Patzlaff *et al*., 2003a, b; Bomal *et al*., 2008).

A functional role in wood formation has been shown for some NAC proteins (Yamaguchi and Demura, 2010). In *Arabidopsis,* the NAC protein subfamily with the capacity to induce cell wall secondary biosynthesis has been named the VNS family and some members of this family such as NST1 (secondary wall thickening promoting factor1), SND1 (secondary wall-associated NAC domain protein1), VND6 (vascular related NAC domain6) and VND7 act together as key regulators along the entire secondary cell wall biosynthesis program (Zhong *et al*., 2006; Mitsuda *et al*., 2007) and activate a cascade of downstream TFs involving Myb proteins. A similar transcriptional network involving NAC proteins known as WNDs (wood-associated NAC domain TFs) and Mybs has been described to operate in poplar and eucalyptus (Zhong and Ye, 2009; Zhong *et al*., 2011). In *Arabidopsis*, Myb46 and Myb83, targets of SND1, are key regulators of the biosynthesis of cellulose, hemicellulose and lignin, three major secondary cell wall

components (Zhong *et al*., 2007a). Several wood-associated Mybs, such as EgMyb2 from eucalyptus, PtrMyb83 and PtrMyb20 from poplar, and PtMyb4 from pine, are functional orthologs of *Arabidopsis* Myb46 and Myb83 (Patzlaff *et al*., 2003a, b; Goicoechea *et al*., 2005; McCarthy *et al*., 2010). Furthermore, in a previous study, we showed that PpMyb8 from maritime pine regulates phenylpropanoid metabolism (Craven-Bartle et al., 2013). These data suggest that PtMyb4 and PtMyb8/PpMyb8 may be members of a transcriptional cascade controlling lignin biosynthesis in conifers (Bomal et al., 2008).

In this report, we have identified and characterized *PpNAC1* from *P. pinaster,* a potential ortholog of the *Arabidopsis* NST1 and SND1 genes that is expressed in the xylem and compression wood of adult trees. We have found that silencing of *PpNAC1* alters the morphology of *P. pinaster* plantlets, which exhibit delayed growth, thickened hypocotyls and a disorganized vascular structure. Furthermore, we have seen that *PpNAC1* is able to activate its own expression and that of the *PpMyb4* transcription factor, which in turn controls the expression of *PpMyb8*. Altogether, our findings suggest that *PpNAC1* and the downstream transcription factors, *PtMyb4* and *PtMyb8* are involved in a transcriptional regulatory network controlling phenylalanine metabolism in maritime pine.

3. Results

3.1. Isolation and characterization of *PpNAC1* from *P. pinaster*

We have previously identified 37 genes that encode NAC transcription factors in the *P. pinaster* genome (Pascual *et al*., 2015). The NAC TFs involved in vascular development has been referred as the VNS family (Ohtani *et al*., 2011; Xu *et al*., 2014) and their members are divided into three groups namely, VND, NST, and SMB, by phylogenetic analysis. In *P. pinaster,* we have identified 3 *PpVNS*-type genes: *PpNAC1* belonging to the NST group, *PpNAC30* in the VND group and *PpNAC31* of the SMB group (Figure 1). Considering that no VNS genes belonging to the NST group has been identified in gymnosperms, we selected *PpNAC1* for a further functional characterization.

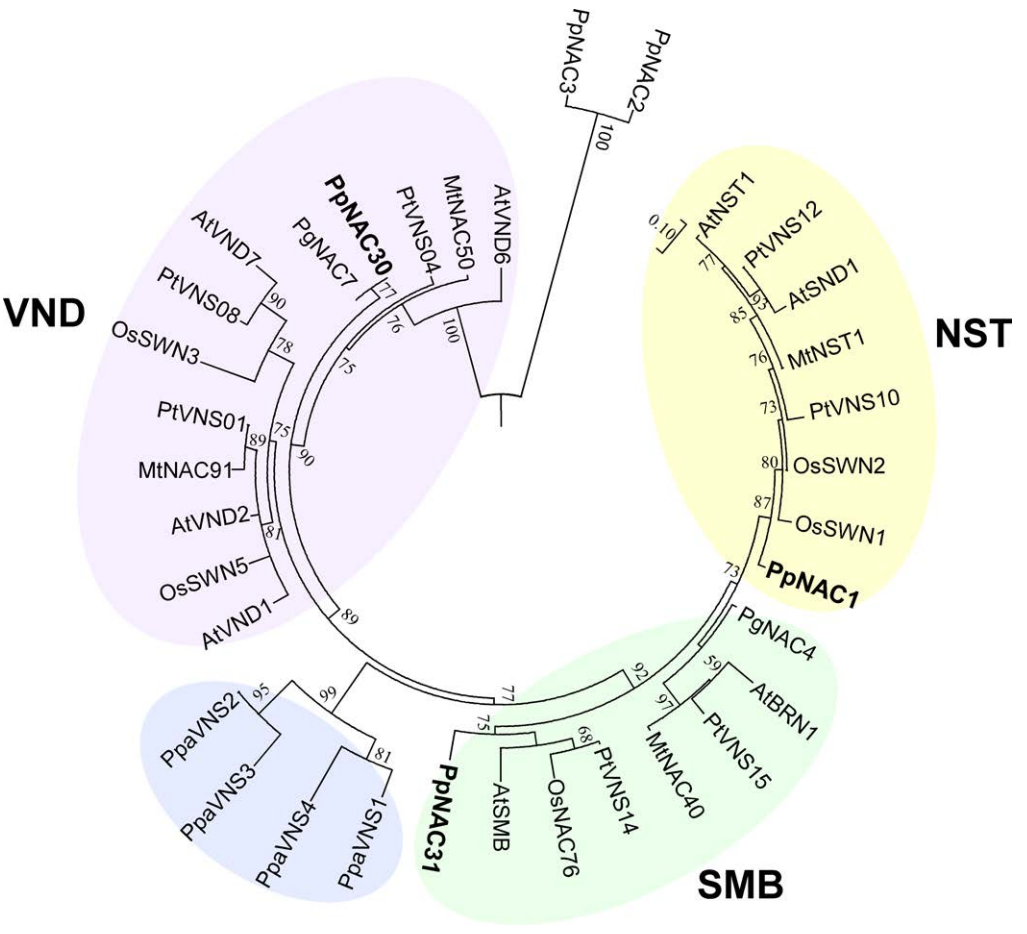


Figure 1. Phylogenetic analysis of full-length VNS proteins. Bootstrapping was performed with 1000 replicates. The VNS proteins are classified in three principal groups: VND, NST and SMB. Abbreviations: Pp, *Pinus pinaster*; At, *Arabidopsis thaliana*; Os, *Oryza sativa*; Pt, *Populus trichocarpa*; Pg, *Picea glauca*; Mt, *Medicago truncatula*; Ppa, *Physcomitrella patens*. Percentage bootstrap values no less than 50% are presented. The accession numbers and sequences of the NAC proteins used in the analysis are available in Supporting Information Table S2.

Using the sequence information available within the *P. pinaster* (http://www.scbi.uma.es/sustainpinedb/sessions/new) and *P. taeda* (https://dendrome. ucdavis.edu/resources/databases/) databases, the cDNA of *PpNAC1* was cloned using PCR and fully sequenced. The full-length cDNA of *PpNAC1* consists of 3252 bp, including a 1434 bp 5´-untranslated region (5´-UTR), an open reading frame (ORF) of 1179 bp encoding a protein of 392 amino acid residues and a 639 bp 3´-untranslated region (3´-UTR). Using

genomic sequence information available in *P. pinaster,* we were able to observe that this gene contains an intron of 569 bp inside the 5´UTR (Figure 2a).

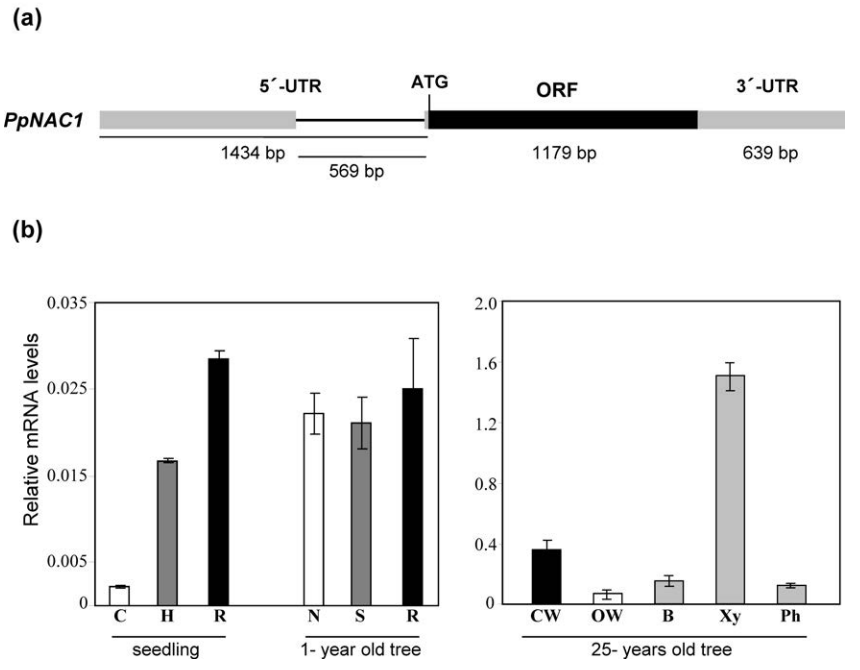


Figure 2. *PpNAC1* cDNA structure and expression analysis. (a) Diagram of the *PpNAC1* cDNA. The coding region of the gene is represented by black boxes; the grey boxes represent untranslated regions and the thinner black line in the 5´-untranslated region represents an intron. The length in base pairs of each region is also shown. (b) Expression of *PpNAC1*. Transcript levels were determined in different organs of *P. pinaster* by qPCR using specific primers (Supporting Information Table S1). The expression data were normalized using *Actin* and *EF1-alpha* as reference genes. Data are mean standard deviation from three biological replicates. C, cotyledons; H, hypocotyl; R, root; N, needles; S, stem; CW, compression wood; OW, opposite wood; B, bark; Xy, secondary xylem; Ph, phloem.

3.2. Expression profiling of *PpNAC1*

The expression pattern of *PpNAC1* gene was analysed by qPCR in different samples from maritime pine seedlings, 1-year-old and 25-years-old trees (Figure 2b). Transcripts were detected in all samples examined, but their relative levels were particularly high in tissues undergoing secondary wall thickening such as the roots and hypocotyls of seedlings and the

needles, stems and roots of 1-year-old trees. Expression analysis in laser-microdissected samples from pine seedling cell types (Cañas *et al*., 2017) revealed that *PpNAC1* was highly expressed in root developing cortex, a tissue with active cell wall biosynthesis (Figure S1). However, the highest mRNA level of *PpNAC1* (50-100-fold of those detected in seedlings) was found in secondary xylem from 25-years-old trees. Furthermore, the transcript accumulation of *PpNAC1* was 5-fold higher in compression than in the opposite wood (Figure 2b).

*3.3. PpNAC1*_RNAi lines

To functionally characterize *PpNAC1* in *P. pinaster*, *PpNAC1*_RNAi lines were generated via somatic embryogenesis using a hairpin construct. We selected a specific fragment of 400 bp of the *PpNAC1* gene to avoid the downregulation of other related targets and cloned it into the gateway vector pBb7GW-I-WG-UBIL. This vector incorporates a BASTA selectable marker and the maize UBIL promoter to drive the expression of the transgene (Figure 3a). Ten independent RNAi transgenic lines were obtained, and the presence of the transgene was confirmed by PCR. The plantlets germinated for 60 days had thickened hypocotyls and poor growth compared with untransformed PN519 plants (Figure 3b). Moreover, their growth was severely delayed, particularly the PN5 and PN9 lines. *PpNAC1* transcript levels were considerably lower in all silenced lines than in control plants, and the two independent lines, PN5 and PN9, showing the highest reduction in *PpNAC1* expression, were selected for further analyses (Figure 3c). We also examined the effects of RNAi inhibition of *PpNAC1* expression on the hypocotyl structure (Figure 3d, e). Cross sections of the stems showed that the vascular morphology of the RNAi plants had a slightly disorganized stem, with a phloem zone expanded with great number of cells and with an altered vascular radial patterning when compared with untransformed PN519 plants.

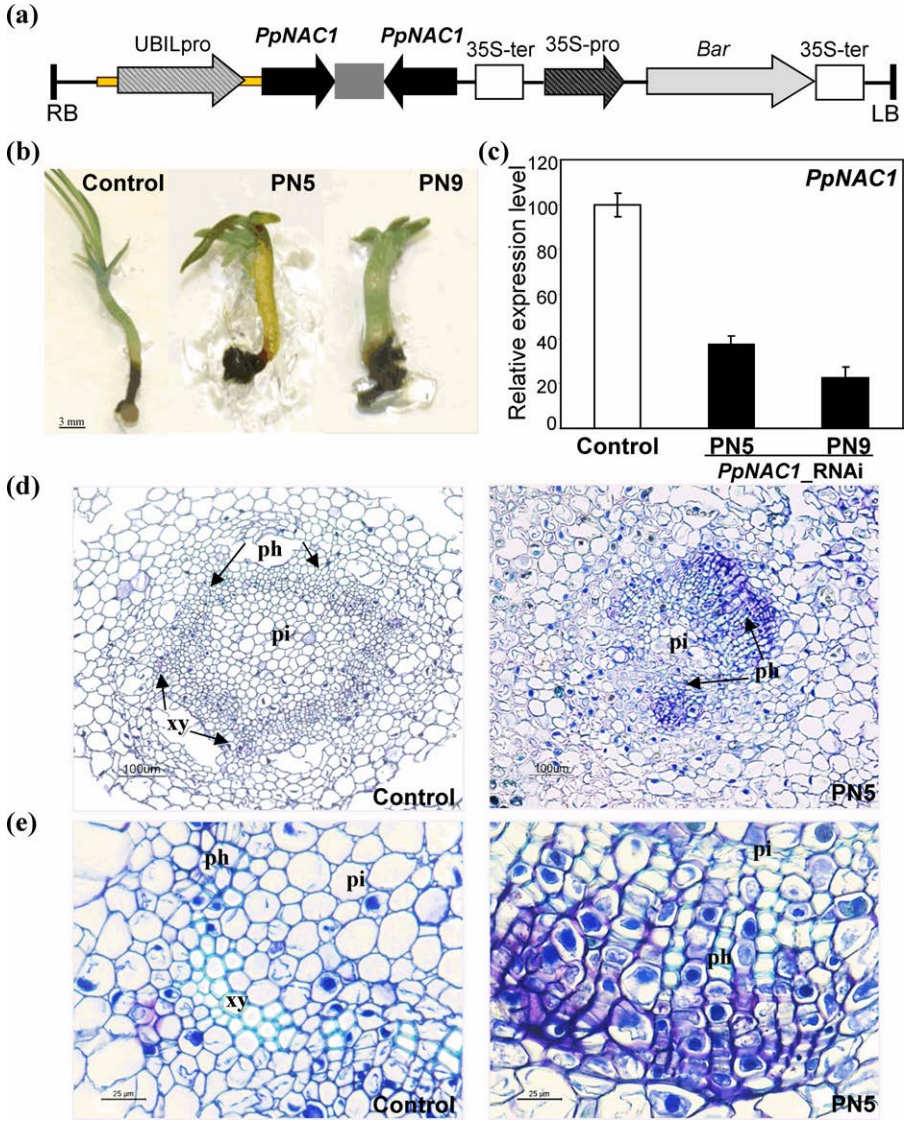


Figure 3. *PpNAC1*_RNAi lines in *P. pinaster*. (a) Diagram of the construct used for the silencing of *PpNAC1*. The black arrow represents the fragment of 400 bp used for gene silencing which expression is driving by the maize UBIL promoter. A grey box between black arrows is a spacer. This vector harbours a BASTA (gen *bar*) selectable marker under the direction of cauliflower mosaic virus 35S gene promoter (CaMV35). (b) Phenotypes of untransformed PN519 (control line) and two independent *PpNAC1* silencing transgenic lines (PN5 and PN9) after 60 days *in vitro* germination in MLV medium. (c) qPCR analysis of *PpNAC1* expression in control (white) and PN5 and PN9 (black) plantlets. Data are mean standard deviation from four biological replicates, each comprised by pooling tissue from three plantlets. Data were normalized to *EF1-alpha* as a reference gene. Expression levels are shown as the percentage of the value obtained for wild-type plants (100%). (d,e) Cross-sections (10 mm) in the hypocotyl of 60d-old *in vitro* germinated plantlets. A phloem zone expanded with great number of cells can be observed in PN5 section. The histological sections were stained with Toluidine blue. Ph, phloem; xy, xylem; pi, pith.

3.4. Transcriptome expression profiling of *PpNAC1*_RNAi *P. pinaster* plantlets

To determine if the downregulation of the *PpNAC1 gene* in *P. pinaster* plantlets resulted in large changes in gene expression, the transcriptomes of silenced and PN519 control plantlets were compared. Total RNA was isolated from PN519, PN5 and PN9 plantlets germinated for 60 days, and the transcriptomes were analyzed using a 4x44K custom array (PINARRAY3). Genes with an adjusted *p- value* below 0.05 and a logarithm fold-change in expression of 0.5 or greater were considered differentially expressed between the transgenic and control lines. For the PN5 line, the number of up- and down-regulated transcripts was 1899 and 1994, respectively. For the PN9 line, the numbers of up- and down regulated transcripts were 1214 and 1474, respectively. For transcriptomic analysis, only genes having differential expression levels in both transgenic lines were considered (Figure 4a; Table S3). Using this criterion, the down-regulation of *PpNAC1* had a strong effect on the pine transcriptome with 928 up-regulated and 1248 down-regulated common genes in both lines (Figure 4a). Functional annotations indicated that many differentially expressed genes were related to cell wall biogenesis, amino acid metabolism and secondary metabolism (phenylpropanoids, flavonoids, and terpenoids) (Figure 4b; Table S4). Key genes for enzymes involved in the monolignol biosynthesis, such as *p*-coumarate 3-hydroxylase (*C3H*), shikimate *O*-hydroxycinnamoyltranferase (*HCT*), caffeoyl-CoA O-methyltransferase (*CcoAOMT*) and cinnamyl-alcohol dehydrogenase (*CAD*), were down-regulated in the transgenic plants. Reduced transcript levels were observed for cellulose synthase *(CESA4),* xyloglucan endotransglucosylase/hydrolase *(XTH)* and laccase (*LAC*) genes involved in secondary cell wall biosynthesis and reassembly.

Genes encoding a-tubulins and putative microtubule-associated proteins (MAPs) were also down-regulated in the RNAi plants. Other genes involved in flavonoid and isoprenoid biosynthesis, such as flavonoid 3´,5´-hydroxylase (*F3´5´H*), geranylgeranyl pyrophosphate synthetase (*GGPS*) and chalcone synthase (*CHS*) were also strongly down-regulated in the transgenic lines. Additionally, transcription factors such as *PpMyb1, PpMyb4* and

PpMyb8 were down-regulated in the RNAi plantlets (Figure S2). The validation of the microarray data was performed by qPCR analysis of 12 differentially expressed genes. Figure 4c shows a comparison of the transcript levels determined by microarray and qPCR analyses, and the results were similar.

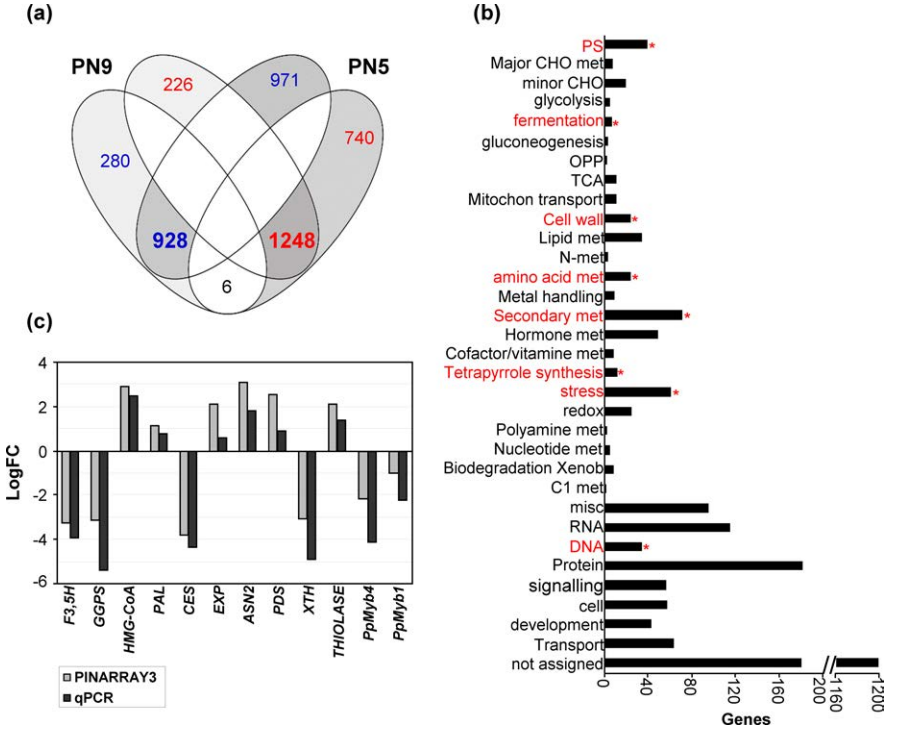


Figure 4. Transcriptome analysis of untransformed PN519 (control) and *PpNAC1*_RNAi plantlets. (a) Venn diagram showing both unique and overlapping (bold) expressed genes of significantly up-regulated (blue) and down-regulated (red) genes between PN5 and PN9 RNAi transgenic plants. (b) Functional enrichment analysis of the functional categories. The horizontal bars represent the number of genes included in each functional category. Functional categories and asterisks in red show significant different categories between the *PpNAC1*_RNAi and control samples with a *P-value* < 0.05 using the Fisher's exact test. (c) Validation of microarray results by qPCR. Fold changes (LogFC) of gene expression in control and *PpNAC1*_RNAi lines (Mean of PN5 and PN9) samples analyzed using PINARRAY3 and qPCR are shown. Positive values correspond to higher expression in *PpNAC1*_RNAi samples and negative values to higher expression in control samples. F3,5H, Flavonoid 3´,5´-hydroxylase; GGPS, Geranylgeranyl pyrophosphate synthetase; HMG-CoA, 3-hydroxy-3-methylglutaryl-coenzyme A reductase; PAL, Phenylalanine ammonia-lyase; CES, cellulose synthase; EXP, Expansin gene; ASN2, Asparagine synthetase 2; PDS, phytoene desaturase; XTH, Xyloglucan endotransglucosylase/hydrolase; THIOLASE, Thiolase family protein; PpMyb4, *P. pinaster* Myb4 transcription factor; PpMyb8, *P. pinaster* Myb8 transcription factor.

3.5. Isolation of *PpNAC1*, *PpMyb4* and *PpMyb8* promoters and *in silico* analysis of putative cis elements

NAC proteins bind to *cis*-acting element containing a consensus sequence of 19-bp named secondary wall NAC binding element (SNBE). This motif is present in many promoters in the *Arabidopsis* genome (Wang *et al*., 2011), including the promoter of Myb46, which is a direct target of SND1 (Zhong *et al*., 2006; 2007a; 2010c). We have isolated the corresponding promoter region of *PpNAC1* (KY451900), which is 1446 pb in length (Figure 5). Simultaneously, and considering that microarray data pointed to a possible coordinated expression of *PpNAC1* and MYB transcription factors regulating lignin biosynthesis, the 5′ flanking regions of *PpMyb4* (KY451898) and *PpMyb8* (KY451899) were isolated and contained 1311 pb and 1083 pb upstream of the ATG, respectively. Figure 5 shows a diagram representing the promoter region for the *PpNAC1*, *PpMyb4* and *PpMyb8* genes and putative *cis* elements identified using PLACE database (http://www.dna.affrc.go.jp/PLACE/). In the *PpNAC1* promoter, at least six putative SNBEs were found as well as one AC element of the AC-II class. In the *PpMyb4* promoter, five putative SNBE sites, one canonical AC element of the AC-III class in the distal region and an AC-rich region (spanning a sequence of 20 bp) proximal to the ATG were found. The upstream region of the *PpMyb8* contained two AC elements of the AC-II class (Box I and Box II) and none SNBE element (Figure 5).

3.6. PpNAC1 protein binds to the SNBE element present on its own promoter to self-activation of gene expression

Previous works have described that SND1 of *Arabidopsis* regulates their own expression, binding directly to an SNBE motif present in its own promoter. To address this possibility in the regulation of *PpNAC1,* we investigated whether PpNAC1 was able to bind its own promoter using electrophoretic mobility shift assays (EMSA). For this analysis, we cloned a 200-pb promoter fragment by PCR containing three putative SNBE motifs present in the upstream region of the gene (from -440 pb to -642 pb, Figure 5). The gel shift observed with PpNAC1 protein was abolished by adding a

competitor DNA proving to be specific (Figure 6a, upper panel). Furthermore, in the trans-activation assay using *P. pinaster* protoplasts, the *PpNAC1* promoter was activated approximately 5-fold when the protoplasts were co-transfected with the effector construct 35S:PpNAC1 (Figure 6a, lower panel). These results indicate that, as occurs in *Arabidopsis* for SND1, the PpNAC1 protein can bind directly to its own promoter, and activates its expression.

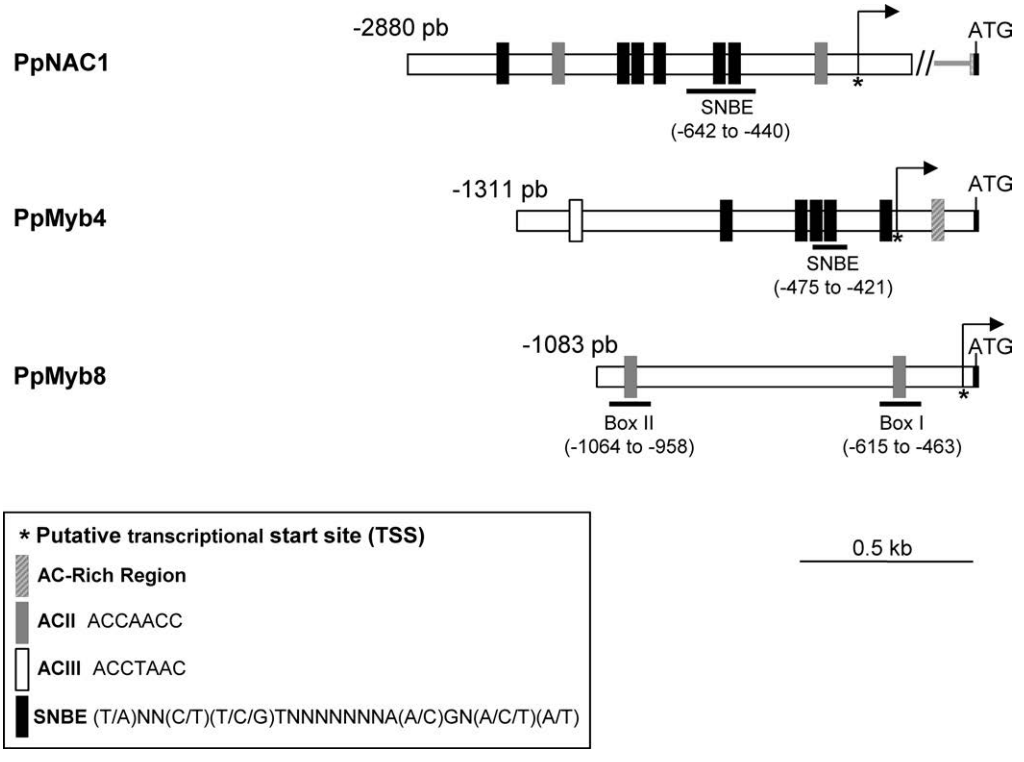


Figure 5. Distribution of SNBE and AC elements in the *PpNAC1*, *Myb4* and *Myb8* promoters. In the schematic representation of the promoters, the position of the transcriptional start site is indicated (arrow). The number to the left of each promoter indicates the relative distance to ATG. The presence of SNBE (black rectangles) and different class AC elements (grey and white rectangles) are showed. An AC-Rich region in the *Myb4* promoter is indicated as striped rectangle. The elements used for further analysis are underlined, and their positions relative to the ATG start codon are indicated.

3.7. PpNAC1 protein binds to the SNBE element present on the *PpMyb4* promoter and activates its expression

In *Arabidopsis,* it has been previously shown that SND1 is able to bind and trans-activate the *Myb46* promoter (Zhong *et al*., 2007a). To examine a possible binding of PpNAC1 to the *PpMyb4* promoter, protein-DNA interactions were analysed using EMSA. PpNAC1 protein was able to bind at a portion of the promoter that contains a SNBE site (from -421 pb to -475 pb, Figure 5) in the *PpMyb4* promoter causing a clearly visible mobility shift (Figure 6b, upper panel).

The complex formed was efficiently competed by the addition of the unlabelled SNBE *Myb4* promoter confirming the specificity of the formed complex. Next, we tested the possible functional significance of the formation of this complex in *P. pinaster* protoplasts. For this assay, pine stem protoplasts were co-transfected with two constructs: the effector construct containing PpNAC1 driven by a tandem duplication of the cauliflower mosaic virus (CaMV) 35S promoter, and the reporter construct containing the *GUS* reporter driven by the *PpMyb4* promoter. As shown in Figure 6b (lower panel), PpNAC1 enhanced *GUS* expression under the *PpMyb4* promoter by approximately 5-fold the levels observed in the absence of the TF.

3.8. PpMyb4 protein binds to the AC-II element present on the PpMyb8 promoter and activates its expression

Previous studies have supported a role of PpMyb8 in the regulation of the phenylalanine pathway (Craven-Bartle *et al*., 2013). *In silico* analysis revealed the presence of two AC-II class elements in the *PpMyb8* promoter: Box I (from -463 pb to -615 pb) and Box II (from -958 pb to -1064 pb) (Figure 5). To test a possible transcriptional control of *PpMyb8* by PpMyb4, EMSA and transactivation analyses were performed. Both *in vitro* and *in vivo* assays (Figure 6c, upper and lower panels) demonstrated a direct interaction of Myb4 protein with the *PpMyb8* promoter resulting in the transcriptional activation of *PpMyb8* by approximately 4-fold the levels observed in the absence of the TF (Figure 6c, lower panel).

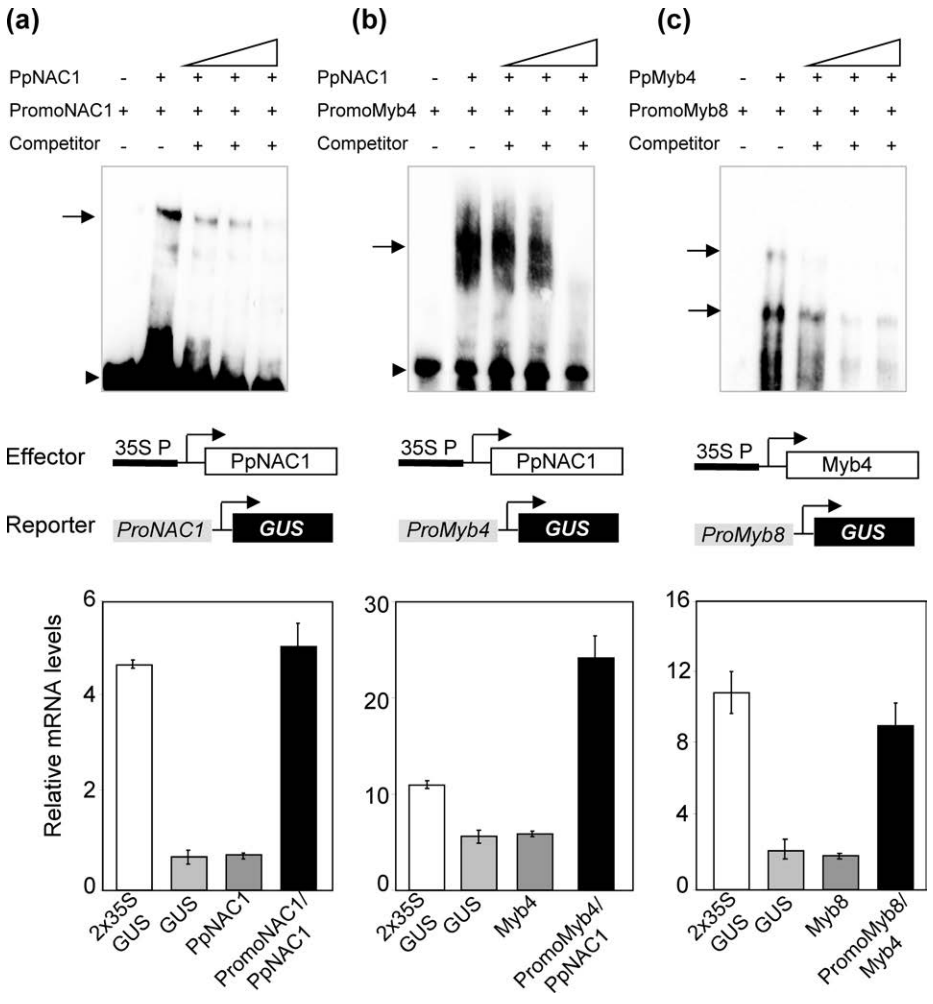


Figure 6. EMSA and transactivation analyses. (a) PpNAC1, (b) PpMyb4 and (c) PpMyb8. In the upper panels, EMSA shows that PpNAC1 binds to their own promoter and to a region in the *Myb4* promoter that contains SNBE sites. The purified protein Myb4 binds to a portion of the *Myb8* promoter that contains AC elements. No band shift was seen in the controls without the addition of protein. Unlabeled promoter fragments in 10, 25 and 50-fold molar excess relative to the labelled probes were used as competitors. Arrows represent shifted complexes. In the lower panels, diagrams of the reporter and effector constructs and transactivation analysis. The effector construct contains the CaMV 35S promoter driven PpNAC1 and PpMyb4. In each reporter construct, the *GUS* gene was driven by the corresponding promoter: *PpNAC1, PpMyb4*, and *PpMyb8*. Transactivation analyses show the PpNAC1-mediated transcriptional activation of *PpNAC1* (a) and *PpMyb4* (b), and the PpMyb4-mediated transcriptional activation of *PpMyb8* (c). The *GUS* RNA levels were assayed in pine stem protoplasts co-transformed with a combination of reporter and effector plasmids at a 1:1 molar ratio. Protoplasts transfected with the reporter or effector construct alone were used as negative controls and the *GUS* RNA levels driven by the 35S promoter were used of positive control (white bar). Bars represent ± SDs from three biological replicates.

4. Discussion

4.1. *PpNAC1* is a potential ortholog of the *NST* genes of angiosperms

The identification of the molecular switches that regulate secondary cell wall biogenesis during wood formation is essential for basic studies and also for the biotechnological manipulation of wood quality and quantity in woody plant species.

Functional studies with wood related NAC TF of *Arabidopsis* such as NST1/2 SND1/2, VND proteins and SOMBRERO (SMB) and BEARSKIN1/2 proteins, have indicated that they are key transcriptional regulators of secondary cell wall (SCW) biosynthesis and have been classified into VND, NST or SMB groups by phylogenetic analysis (Nakano *et al*., 2015). So far no direct correlation has been found between the number of VNS genes present in a plant species and the size of its genome or the abundance of lignified tissues (Zhu *et al*., 2012; Nakano *et al*., 2015). For example, the moss *P. patens* contains 8 VNS genes in its genome (Zhu *et al*., 2012; Xu *et al*., 2014), while *P. trichocarpa* and *E. grandis*, two woody angiosperms, have 16 (Zhong *et al*., 2010b; Ohtani *et al*., 2011) and 6 VNS genes (Myburg *et al*., 2014; Hussey *et al*., 2015), respectively. In conifers, *P. abies* has 4 (Nystedt *et al*., 2013), *P. glauca* 2 (Duval *et al*., 2014) and *P. pinaster* 3 VNS genes (Pascual *et al*., 2015).

The phylogenetic analysis of VNS genes showed that *P. pinaster* presents one gene classified in each group (Figure 1) and points to *PpNAC1* as a potential ortholog of the *SND1* gene of *Arabidopsis*. To our knowledge, no members of the NST group have been previously identified in gymnosperms. In *P. glauca,* only two VNS genes have been identified; *PgNAC7* is a VND-type gene functionally similar to the *AtVND6* and its expression is preferentially associated to vascular tissue in the stem, while the *PgNAC4* clustered with the SMB group, and its expression was clearly predominant in root tips (Duval *et al*., 2014), as described in *Arabidopsis* (Zhong *et al*., 2010b; Ohtani *et al*., 2011).

PpNAC1 is predominantly expressed in the secondary xylem and compression wood of adult trees, tissues undergoing lignin biosynthesis (Figure 2b). This transcript level distribution is consistent with previous observations reported for the PtrWNDs in *Populus* and the SND1, NSTs and VNDs in *Arabidopsis* (Kubo *et al*., 2005; Zhong *et al*., 2007b; Zhong and Ye, 2010). Although the NAC family in conifers (37 putative members in *P. pinaster)* appears to be underrepresented when compared with the more than 100 members in *Arabidopsis* or poplar, the identification of *PpNAC1* suggests that the primary layer of the NAC master switch for secondary cell wall formation has been evolutionarily conserved in vascular plants.

4.2. *PpNAC1*_RNAi lines exhibit altered vascular differentiation

One common strategy for functional characterization of a candidate gene is to down- or up-regulate its expression by genetic transformation. The long generation time and long life span of conifers have been major obstacles to perform reverse genetic approaches in these woody plants, and the functional studies of many conifer genes have been performed in *Arabidopsis* and tobacco (Patzlaff *et al*., 2003b; Newman *et al*., 2004). Nevertheless, advances have been made in the generation of transgenic conifers via somatic embryogenesis, and efficient protocols are currently available for genetic transformation and cryopreservation of embryogenic cell lines and subsequent plant regeneration (Klimaszewska *et al*., 2004; Trontin *et al*., 2007). Using these protocols, we have generated RNAi *P. pinaster* lines for *PpNAC1*. The morphological phenotypes, vascular tissue architecture and expression analysis of RNAi_*PpNAC1* plantlets strongly suggest that *PpNAC1* is associated with vascular development (Figures 3 and 4). In *Arabidopsis*, down-regulation of both SND1 and NST1 genes resulted in loss of secondary cell wall in the xylem fibers of stem, and consequently in a lower stem strength. In addition, several genes involved in the secondary wall biosynthesis were down-regulated in these plants (Zhong *et al*., 2007b). The pine transcriptome was also strongly disturbed in the *PpNAC1*_RNAi plantlets, resulting in altered expression of a range of genes implicated in cell wall biogenesis, amino acid metabolism and secondary metabolism (Figure 4;

Tables S3 and S4). Key genes for enzymes of monolignol biosynthesis, such as *C3H*, *CAD* or *CcoAOMT*, were down-regulated in the transgenic plants. In *P. radiata* has been shown that CcoAOMT is needed for biosynthesis of guaiacyl lignin and its suppression modifies lignin content and composition resulting in a lignin polymer with an unusual subunit composition (Wagner *et al*., 2011). Reduced transcript levels were also observed for *CesA4* and *XTH,* which are involved in secondary cell wall biosynthesis and reassembly during growth and differentiation. XTH modifies xyloglucan, the major hemicellulose present in the primary cell walls of pine trees (Valenzuela *et al*., 2014). Laccase genes were also down-regulated in the RNAi plants. Suppression of *LAC4* and *LAC17* expression in *Arabidopsis* affected lignin biosynthesis mainly in fiber cells of the inflorescence stem (Berthet *et al*., 2011; Zhao *et al*., 2013; Schuetz *et al*., 2014).

Genes encoding a-tubulins and putative microtubule-associated proteins (MAPs) were also down-regulated in the pine RNAi plants. In *Arabidopsis,* the silencing of AtMAP70-5 protein produces atrophied plants exhibiting disorganized vascular elements, suggesting that these proteins are essential for secondary cell wall biogenesis and for the adequate development of xylem (Pesquet *et al*., 2010).

A remarkable finding is that *PpMyb1*, *PpMyb4* and *PpMyb8* genes were down-regulated in the RNAi plantlets (Figure S2). These *Myb* genes are expressed in secondary xylem and have been functionally associated with phenylpropanoid and lignin biosynthesis in *P. glauca* (Bomal *et al*., 2008; Bedon *et al*., 2010), *P. pinaster* (Craven-Bartle *et al*., 2013) and *P. taeda* (Patzlaff *et al*., 2003a, b).

4.3. A transcriptional regulation network controlling phenylpropanoid biosynthesis in maritime pine

It is well known that some NAC proteins (SND1, NST1 and VNDs) act as master regulators of a signalling cascade that involves R2R3-MYBs and regulates vascular development and secondary cell wall formation in *Arabidopsis*. Functional orthologs of this network have been identified in poplar (Lin *et al*., 2013; Zhong *et al*., 2013) and eucalyptus, but functional

studies in conifers are scant. Duval *et al*. (2014) reported that *PgNAC7* could be a master regulator of secondary cell wall biosynthesis in conifer xylem. Recently, it has been proposed that *PgNAC8* could also regulate cellulose biosynthesis in coordination with *PgNAC7* (Lamara *et al*., 2015). Phylogenetic and expression analysis of *PgNAC8* have suggested a role as a potential candidate ortholog of SND2/3 genes regulating complex carbohydrate biosynthesis (Zhong *et al*., 2010a; Lamara *et al*., 2015).

The *in vitro* trans-activation and EMSA analysis showed that PpNAC1 can bind directly to its own promoter to activate transcription through a positive feedback loop (Figure 6a). Likewise, PpNAC1 was able to activate *PpMyb4* expression (Figure 6b), while PpMyb4 was able to activate *PpMyb8* (Figure 6c). Furthermore, the transcript levels of *PpMyb4* and *PpMyb8* were drastically reduced when *PpNAC1* was silenced in the transgenic plants (Figure S2).

The results obtained in this work suggested that a transcriptional cascade similar to the SND1 network defined in *Arabidopsis* (Zhong *et al*., 2006) and poplar (Lin *et al*., 2013; Wang *et al*., 2014) is conserved in conifers (Figure 7). To gain insight into this question we have generated Arabidopsis plants overexpressing PpNAC1. The overexpression of PpNAC1 produced a prominent phenotypic effect in Arabidopsis plants with small rosette size and curled leaves (Figure S3a). In addition, an up-regulation of secondary wall biosynthetic genes was observed (Figure S3b) together with increased transcript levels of secondary cell wall-associated TFs (Figure S3c). This behaviour is similar to that of previously described in Arabidopsis plants overexpressing SND1, NST1, PtrWND2B and PtrWND6B (Mitsuda *et al*., 2005; Zhong *et al*., 2006; Zhong *et al*., 2010).

Moreover, Myb58 and Myb63 transcription factors, regulated by SND1 and Myb46 in Arabidopsis, in turn regulate the expression of genes involved in lignin biosynthesis (Zhou *et al*., 2009). These MYB factors are significantly up-regulated in the PpNAC1 overexpressing Arabidopsis plants, strongly suggesting a functional role of PpNAC1 as ortholog of the Arabidopsis SND1 transcription factor (Figure S3).

PpNAC1 may be a regulator that could act at the first level of transcriptional control in phenylalanine metabolism to promote wood formation, while PpMyb4 would function upstream of PpMyb8 and other MYBs. In conifers, PpMyb8 and PgMyb8 are involved in the coordinated expression of lignin biosynthesis through binding to *cis* regulatory elements present in the promoter region of key genes in the pathway (Craven-Bartle *et al*., 2013; Bomal *et al*., 2014). Moreover, transgenic plants overexpressing *PtMyb8* showed up-regulation of genes encoding phenylpropanoid enzymes and misregulation of several cell wall-related genes (Bomal *et al*., 2008).

PpMyb4 and its orthologs, PtMyb4 and EgMyb2, could be non-specific for regulating lignin biosynthesis because they are also involved in the regulation of cellulose and xylan biosynthesis (Zhong *et al*., 2013). Taken together, the results presented here suggest that PpNAC1, PpMyb4 and PpMyb8 are activators of lignin biosynthesis in maritime pine. The identification of *PpNAC1* as a main regulator of this network involved in wood formation in *P. pinaster* is of great interest for fundamental studies in conifers but also for potential applications in tree biotechnology. In order to increase our knowledge about the transcriptional regulatory network operating in conifers, and given the complexity of the network demonstrated in other species, intensive research is necessary to fully clarify to what extent the transcriptional network could be conserved between gymnosperms and angiosperms.

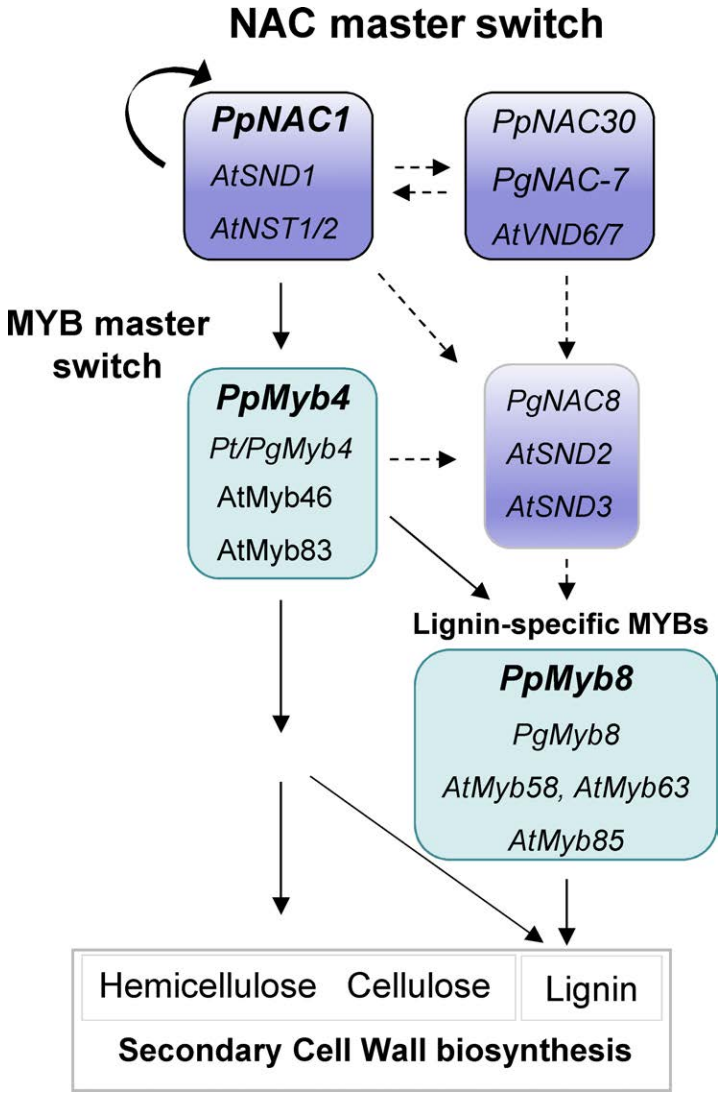


Figure 7. Proposed model of transcriptional regulatory network controlling secondary cell wall biosynthesis in *P. pinaster*. Transcription factors from *Arabidopsis*, *P. pinaster*, and *P. glauca* that have been functionally characterized in the network are shown. Continuous lines indicate that the interactions have been functionally demonstrated, while the dashed lines indicate interactions that have not yet been demonstrated.

5. Experimental Procedures

5.1. Plant material

Pinus pinaster Ait. seeds were supplied by the Centro de Recursos Genéticos Forestales "El Serranillo" (Ministerio de Agricultura, Pesca, Alimentación y Medio Ambiente, Spain). The seed germination and growth of the seedlings were performed as described elsewhere (Pascual *et al*., 2015). Cotyledons, hypocotyls and roots from pine seedlings (one-month-old) and needles, stems and roots from young trees (1-year-old) were collected separately, frozen in liquid nitrogen and stored at -80ºC until use. Samples of bark, xylem, phloem, and compression and opposite wood were collected from maritime pine 25-years-old trees of Sierra Bermeja (Estepona, Spain) (Villalobos *et al*., 2012).

5.2. Constructs and pine transformation

The full-length *PpNAC1* cDNA was cloned from pine seedling hypocotyl RNA by PCR amplification of the specific sequence in the *P. pinaster* database (http://www.scbi.uma.es/sustainpinedb/sessions/new).

To obtain RNA interference (RNAi) we used a hairpin construct. For this, a fragment of 400-bp from the *PpNAC1* gene was amplified using specific Gateway primers, cloned into pDONR207 (Invitrogen, Germany) and introduced into the monocot-specific vector pBb7GW-I-WG-UBIL, which contains a BASTA selectable marker and makes use of maize ubiquitin (UBIL) promoter to drive the expression of the transgene. The *A. tumefaciens* strain C58C1 was transformed by electroporation.

The *P. pinaster* embryogenic cell line PN519 (Trontin *et al*., 2007) has been used and maintained according to Klimaszewska *et al*. (2001). The transformation of PN519 was performed as previously described (Klimaszewska *et al*. 2004) and transferred to proliferation medium containing plant growth regulators. For the maturation of control and transformed embryos, tissues were transferred to maturation medium supplemented with abscisic acid (ABA). The mature embryos were germinated for two months in controlled conditions. Genomic DNA was isolated from Basta-resistant

embryonal tissues and transgenics confirmed by PCR analysis. Independent transgenic lines exhibiting reduced levels of *PpNAC1* transcripts were selected for embryo maturation and production of somatic embryo plants. In addition, the untransformed PN519 line was used as a control. Primers used are listed in Supporting Information Table S1.

5.3. Plantlets growth and histology

Somatic embryos were germinated for 60 days on MLV medium contained 87 mM sucrose. *In vitro* plantlets were frozen into liquid nitrogen and stored at -80ºC until use. For histological analysis, hypocotyls were immediately fixed in 4% (v/v) paraformaldehyde in 0.1M phosphate buffer (pH 7.4) under vacuum (3 times for 15 min). The samples were dehydrated and infiltrated with paraffin for 5 days. Thin sections (10mm) were prepared using a microtome and paraffin-free sections were stained with 1% Toluidine blue.

5.4. Isolation of the promoter regions of *PpNAC1*, *PpMyb4* and *PpMyb8*

Isolation of genomic DNA was performed using the CTAB method (Doyle and Doyle, 1987) and the promoter sequence of *PpNAC1* was obtained by PCR amplification using primers designed from the loblolly pine database (https://dendrome.ucdavis.edu/resources/databases/). The promoters of *PpMyb4* and *PpMyb8* genes were obtained by PCR walking. A list of primers is provided in Supporting Information Table S1.

5.5. RNA isolation and qPCR

The isolation of RNA was performed as described elsewhere (Canales *et al*., 2012). RQ1 RNase-Free DNase (Promega Corporation, Madison, WI) was used for the removal genomic DNA contamination from RNA samples, and cDNA synthesis was performed with iScript Reverse Transcription Supermix (Bio-Rad®). Real-time PCR (qPCR) was performed according to Canales et al. (2012). *Actin* and *elongation factor-1-alpha* (*EF1-a*) were used as reference genes. The gene-specific primers used are listed in Supporting Information Table S1.

The laser capture microdissection procedure and qPCR analysis were carried out as described (Cañas et al. 2017).

5.6. Microarray hybridization

Somatic embryos were germinated for 60 days on MLV medium contained 87 mM sucrose. Two RNAi transgenic lines, PN5 and PN9, as well as the control cell line PN519, were used, with 3 biological replicates per line and 6 somatic embryo plants per replicate.

A custom microarray (PINARRAY3) was used that includes 60-mer oligonucleotides designed using the *P. pinaster* transcriptome (Canales *et al*., 2014). Slides were made by Agilent Technologies and hybridization was performed at 65ºC following the protocol described by Cañas *et al*. (2015). Then, the slides were washed and air-dried. Hybridized slides were scanned and signal intensities were recorded. The differentially expressed genes were identified using the Limma package for R (Smyth GK, 2015). The microarray data are accessible at NCBI´s Gene Expression Omnibus (Edgar *et al*., 2002) through the accession number GSE89341.

5.7. Protein expression and EMSA

The full-length cDNAs of *PpNAC1* and *PpMyb4* were amplified and cloned into the pDEST17 vector (Invitrogen, Germany). The production of the recombinants proteins was induced in the *E. coli* strain BL21-AI at 20ºC for 5 h in presence of 0.2% arabinose. The recombinant PpNAC1 and PpMyb4 proteins were purified by affinity chromatography and used for EMSA with the *PpNAC1, PpMyb4* and *PpMyb8* promoter fragments.

The primers used to amplify the promoter DNA fragments were labelled with biotin at the 5´ terminus and their sequences are provided in Supporting Information Table S1. For EMSAs, 1 µg of purified PpNAC1 or PpMyb4 was incubated at room temperature for 30 min with the biotin-labeled promoter fragment in the binding buffer (10 mM Tris, pH 7.5, 5 mM $MgCl_2$, 2.5% glycerol, 0.05% NP-40, 100 ng ml^{-1} poly(dI-dC)). For competition analysis, unlabeled fragments were added in the reactions as competitors in a 10, 25- or 50-fold molar excess relative to the labelled probes. The samples

were resolved in 5% polyacrylamide non-denaturing gel, electrotransferred onto nylon membranes and signals revealed using a chemiluminescence kit (Thermo Fisher scientific).

5.8. Transient expression analysis in pine protoplasts

The procedure was performed following the protocol described previously (Gómez-Maldonado *et al*. 2004). To prepare the reporter constructs, the promoter sequences of *PpNAC1*, *PpMyb4* and *PpMYB8* were cloned into the pBI221 plasmid replacing the CaMV 35S promoter. Each of the reporter constructs was co-transformed with the corresponding effector construct into pine stem protoplasts according to Gómez-Maldonado *et al*. (2004). After incubation for 16 h in dark, the protoplasts were retrieved by centrifugation at 500 x g for 3 min and frozen in liquid nitrogen. The GUS RNA levels were determined using specific primers presented in Supporting Information Table S1.

5.9. Maximum likelihood phylogenetic analysis

The phylogenetic analysis was conducted with 35 full-length sequences from 7 species: *Pinus pinaster*, *Picea glauca*, *Populus trichocarpa*, *Arabidopsis thaliana*, *Medicago truncatula*, *Oryza sativa*. Four *Physcomitrella patents* NAC proteins were used as out-group to root the tree. Multiple alignments were carried out using MUSCLE v3.8.31 (Edgar, 2004) and tree topology was inferred using maximum likelihood with PhyML (Guindon and Gascuel, 2003; Guidon *et al*., 2005). The bootstrap test was carried out with 1000 replicates. The MEGA 7.0 software (Kumar *et al*., 2016) was used to draw phylogenetic trees. The accession numbers of the NAC sequences are available in Supporting Information Table S2.

6. Acknowledgements

This work was supported by the Plant KBBE program, Scientific and Technological Cooperation in Plant Genome Research (PLE2009-0016), the ProCoGen project (FP7-KBBE-289841), and by grants from the Spanish Ministerio de Economía y Competitividad (BIO2015-69285-R) and Junta de Andalucía (BIO-474). We thank Jean-François Trontin (FCBA Technological Institute, Pierroton, France) for kindly providing the embryogenic line PN519 and Krystyna Klimaszewska (Natural Resources, Otawa, Canada) for advice in pine transformation. We also thank Francisco R. Cantón (Department of Molecular Biology and Biochemistry, Málaga University) for providing the samples of 25-years old tree.

7. Supporting information

Additional supporting information may be found in the online version of this article.

Table S1. Oligonucleotides used in this work.

Table S2. Names, gene accession numbers and sequences of the NAC protein used in the phylogenetic analysis.

Table S3. Microarray results. Up-regulated differential expressed genes are highlighted in red. Down-regulated differential expressed genes are highlighted in blue.

Table S4. Functional enrichment analysis results.

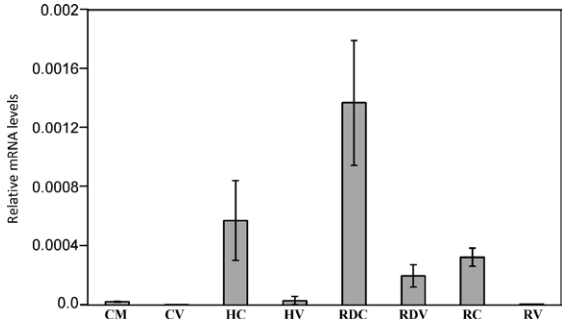


Figure S1. Expression profile of *PpNAC1* in laser-microdissected tissues from one-month-old *P. pinaster* seedlings. The expression data were normalized using *Actin* as reference gene. Data are mean standard deviation from three biological replicates. CM, Cotyledon_mesophyll; CV, Cotyledon_vascular; HC, Hypocotyl_cortex; HV, Hypocotyl_vascular; RDC, Root_developing_cortex; RDV, Root_developing_vascular; RC, Root_cortex; RV, Root_vascular.

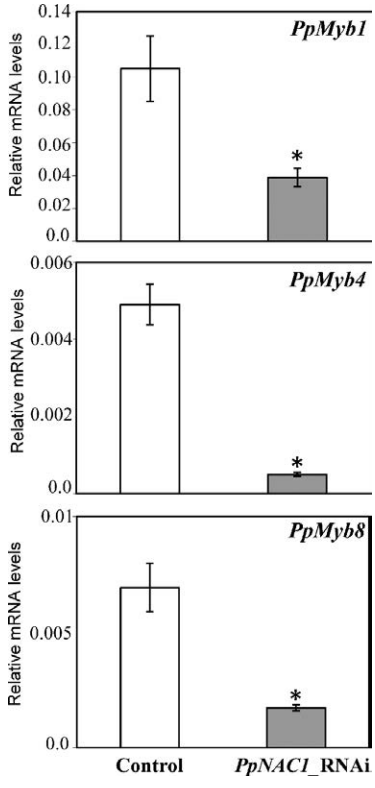


Figure S2. Expression level of *PpMYB1*, *PpMYB4* and *PpMYB8* transcription factors in control (white) and *PpNAC1*_RNAi lines (mean of PN5 and PN9, gray) plantlets. Levels were analyzed by qPCR using *EF1-alpha* and *Actin* as standard gene for normalization. Data are mean standard deviation from three biological replicates. Asterisks indicate statistical differences between the *PpNAC1*_RNAi and control samples with a P-*value* < 0.05 by Student´s test. Error bars represent ± SD.

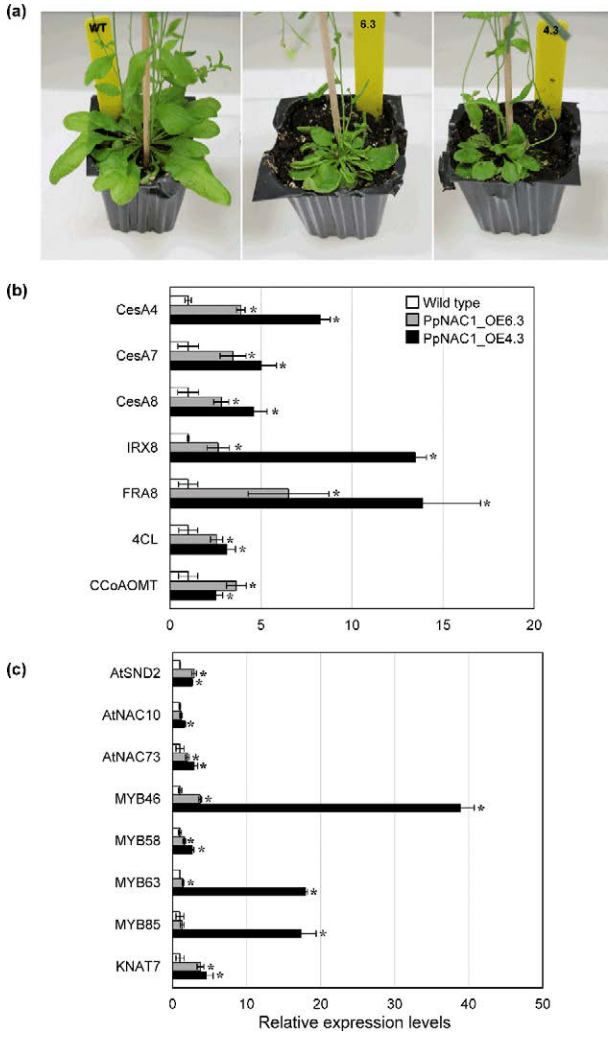


Figure S3. Overexpression of *PpNAC1* in *Arabidopsis* plants**.** The full-length cDNA of *PpNAC1* driven by the CaMV 35S promoter was introduced into wild-type *Arabidopsis*. (a) Seedlings of overexpressors of *PpNAC1*_OE4.3 (middle) and *PpNAC1*_OE6.3 (right) showing curly leaves compared with the wild type (left). (b) Expression level of genes involved in the biosynthesis of cellulose (*CesA4, At5g44030; CesA7, At5g17420* and *CesA8, At4g18780*), xylan (*FRA8, At2g28110* and *IRX8, At5g54690*), and lignin (*CCoAOMT,* At4g34050 and *4CL1, At1g51680*) in seedling of 4-week-old *PpNAC1* overexpressors compared with the wild type (control). (c) Expression levels of secondary wall-associated transcription factors in the seedlings of *PpNAC1* overexpressors compared with the wild type. The Arabidopsis Genome Initiative locus identifiers fo the Arabidopsis genes investigated in this study are as follows: AtSND2 (At4g28500); AtNAC10 (At1g28470); AtNAC73 (At4g28500); MYB46 (At5g12870); MYB58 (At1g16490); MYB63 (At1g7918); MYB85 (At4g22680); KNAT7 (At1g62990). The expression level of each gene in the wild type is set to 1. Asterisks indicate statistical differences between overexpressors and control samples with a P-*value* < 0.05 by Student´s test. Error bars represent ± SD of three biological replicates.

8. References

Allona, I., Quinn, M., Shoop, E., Swope, K., StCyr, S., Carlis, J. et al. (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proc. Natl Acad. Sci. USA,* 95, 9693-9698.

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 Biology*, 7, 17.

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S.D., Schmidt, A., Gershenzon, J. et al. (2010) Subgroup 4R2R3-MYBs in conifer trees: gene gamily expansion and contribution to the isoprenoid- and flavonoid-oriented responses*. Journal Experimental Botany*, 61, 3847-64.

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., Borrega, N., Hervé, J., Blondet, E. et al. (2011) Disruption of *LACCASE4* and *17* results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell*, 23, 1124-1137.

Bomal, C., Bedon, F., Caron, S., Mansfield, S.D., Levasseur, C., Cooke, J.E.K. et al. (2008) Involvement of *Pinus taeda* MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis. *Journal Experimental Botany*, 59, 3925–3939.

Bomal, C., Duval, I., Giguère, I., Fortin, È., Caron, S., Stewart, D., Boyle, B., Seguin, A. and MacKay, J.J. (2014) Opposite action of R2R3-MYBs from different subgroups on key genes of the shikimate and monolignol pathways in spruce. *Journal Experimental Botany*, 65, 495–508.

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

Canales, J., Bautista, R., Label, P., Gómez-Maldonado, J., Lesur, I., Fernández-Pozo, N. et al. (2014) De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. *Plant Biotechnology Journal*, 12, 286-299.

Cañas, R.A., Feito, I., Fuente-Maqueda, J.F., Ávila, C., Majada, J. and Cánovas, F.M. (2015) Transcriptome-wide analysis supports environmental adaptations of two *Pinus pinaster* populations from contrasting habitats. *BMC genomics*, 16, 909.

Cañas, R.A., Li, Z., Pascual, M.B., Castro-Rodríguez, V., Ávila, C., Sterck, L., Van de Peer, Y. and Cánovas, F.M. (2017) The gene expression landscape of pine seedling tissues. *The Plant Journal*, 10.1111/tpj.13617.

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

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

Doyle, J.J. and Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19, 11-15.

Du, J. and Groover, A. (2010) Transcriptional regulation of secondary growth and wood formation. *Journal of Integrative Plant Biology*, 52, 17–27.

Duval, I., Lachance, D., Giguère, I., Bomal, C., Morency, M.J., Pelletier, G., Boyle, B., Mackay, J.J. and Séguin, A. (2014) Large-scale screening of transcription factor-promoter interactions in spruce reveals a transcriptional network involved in vascular development. *Journal Experimental Botany*, 65, 2319-2333.

Edgar, R., Domrachev, M. and Lash, A.E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30, 207–210.

Edgar, R.C. (2004) MUSCLE: multiple sequences alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792–1797.

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 Journal*, 43, 553-567.

Gómez-Maldonado, J., Ávila, C., de la Torre, F., Cañas, R.A., Cánovas, F.M. and Campbell, M.M. (2004) Functional interactions between a glutamine synthetase promoter and MYB proteins. *Plant Journal*, 39, 513-526.

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.

Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 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 Research*, 33, 557-559.

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*. *Plant Biology*, 10, 145.

Hussey, S.G., Saïdi, M.N., Hefer, C.A., Myburg, A.A. and Grima-Pettenati, J. (2015) Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytologist*, 206, 1337–50.

Klimaszewska, K., Lachance, D., Pelletier, G., Lelu, M.A., Séguin, A. (2001) Regeneration of transgenic *Picea glauca*, *P. mariana*, and *P. abies* after cocultivation of embryogenic tissue with *Agrobacterium tumefaciens*. *In Vitro Cellular & Developmental Biology Plant*, 37, 748–755.

Klimaszewska, K., Rutledge, R.G. and Séguin, A. (2004) Genetic transformation of conifers utilizing somatic embryogenesis. In: Peña L, ed. *Methods in Molecular Biology,* pp. 151–164. Totowa, NJ: Humana Press.

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J. et al. (2005) Transcription switches for protoxylem and metaxylem vessel formation. *Genes & Development*, 19, 1855–1860.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, vol. 33, no. 7, pp. 1870–1874.

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 Phytologist*, 210, 240-255.

Lin, Y.C., Li, W., Sun, Y.H., Kumari, S., Wei, H., Li, Q., Tunlaya-Anukit, S., Sederoff, R.R. and Chiang, V.L. (2013) SND1 Transcription factor-directed quantitative functional hierarchical genetic regulatory network in wood formation in *Populus trichocarpa. Plant Cell*, 25, 4324-4341.

Lorenz, W.W. and Dean, J.F.D. (2002) SAGE Profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). *Tree Physiol*ogy, 22, 301–310.

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

McCarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K. et al. (2010) The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. *Plant Cell Physiology*, 51, 1084–1090.

Mitsuda, N., Seki, M., Shinozaki, K., Ohme-Takagi, M. (2005) The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall. *Plant Cell*, 17, 2993-3006.

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K. and Ohme-Takagi, M. (2007) NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. *Plant Cell*, 19, 270 –280.

Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J. et al. (2014) The genome of *Eucalyptus grandis. Nature*, 510, 356-362.

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

Newman, L.J., Perazza, D.E., Juda, L. and Campbell, M.M. (2004) Involvement of the R2R3-MYB, AtMYB61, in the ectopic lignification and dark-photomorphogenic components of the det3 mutant phenotype. *The Plant Journal*, 37, 239-250.

Nuruzzaman, M., Manimekalai, R., Sharoni, A.M., 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, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.C., Scofield, D.G. *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 Journal*, 67, 499–512.

Olsen, A.N., Ernst, H.A., Leggio, L.L. and Skriver, K. (2005) NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science*, 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 Research*, 10, 239–47.

Pascual, M.B., Cánovas, F.M. 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 Biology*, 15, 254.

Pascual, M.B., El-Azaz, J., de la Torre, F., Cañas, R.A., Ávila, C. and Cánovas, F.M. (2016) Biosynthesis and metabolic fate of phenylalanine in conifers. *Frontiers in Plant Science*, 7, 1030.

Patzlaff, A., McInnis, S., Courtenay, A. *et al.* (2003a) Characterisation of a pine MYB that regulates lignification. *Plant Journal*, 36, 743–754.

Patzlaff, A., Newman, L.J., Dubos, C., Whetten, R.W., Smith, C., McInnis, S., Bevan, M.W., Sederoff, R.R. and Campbell, M.M. (2003b) Characterization of PtMYB1, an R2R3-MYB from pine xylem. *Plant Molecular Biology*, 53, 597–608.

Pavy, N., Paule, C., Parsons, L., Crow, J.A., Morencey, M.J., Cooke, J. et al. (2005) Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC Genomics*, 6, 144.

Pesquet, E., Korolev, A.V., Calder, G. and Lloyd, C.W. (2010) The microtubule-associated protein AtMAP70-5 regulates secondary wall patterning in *Arabidopsis* wood cells. *Current Biology*, 20, 744-748.

Rengel, D., Clemente, H.S., Servant, F., Ladouce, N., Paux, E., Wincker, P. *et al*. (2009) A new genomic resource dedicated to wood formation in *Eucalyptus*. *BMC Plant Biology*, 9, 36.

Schuetz, M., Benske, A., Smith, R.A., Watanabe, Y., Tobimatsu, Y., Ralph, J., Demura, T., Ellis, B. and Samuels, A.L. (2014) Laccases direct lignification in the discrete secondary cell wall domains of protoxylem. *Plant Physiology*, 166, 798-807.

Smyth, G.K. (2005) Limma: Linear Models for Microarray Data. In: *Bioinformatics and Computational Biology Solutions using R and Bioconductor, R*. Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S, editors, pp 397-420. Springer, New York.

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A.M. et al. (2004) A populus EST resources for plant functional genomics. *Proc. Natl Acad. Sci USA*, 101, 13951–13956.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24, 1596–1599.

Trontin, J.F., Walter, C., Klimaszewska, K., Park, Y.S. and Lelu-Walter, M.A. (2007) Recent progress in Genetic transformation of four *Pinus* spp. *Transgenic Plant Journal*, 1, 314-329.

Valenzuela, C., Ramos, P., Carrasco, C., Moya-León, M.A. and Herrera, R. (2014) Cloning and characterization of a xyloglucan endotransglycosylase/hydrolase gene expressed in response to inclination in radiata pine seedlings. *Tree Genetics Genomes*, 10, 1305-1315.

Villalobos, D.P., Díaz-Moreno, S.M., el-SS, S., Cañas, R.A., Osuna, D., Van Kerckhoven, S.H., Bautista, R., Claros, M.G., Cánovas, F.M. and Cantón, F.R. (2012) Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. *BMC Plant Biology*, 12, 100.

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., Pears, L. and Ralph, J. (2011) *CCoAOMT* suppression modifies lignin composition in Pinus radiata. *The Plant Journal*, 67, 119-129.

Wang, H., Zhao, Q., Chen, F., Wang, M. and Dixon, R.A. (2011) NAC domain function and transcripcional control of a secondary cell wall master switch. *Plant Journal*, 68, 1104-1114.

Wang, S., Li, E., Porth, I., Chen, J.G., Mansfield, S.D. and Douglas, C.J. (2014) Regulation of secondary cell wall biosynthesis by poplar R2R3

MYB transcription factor PtrMYB152 in *Arabidopsis. Scientific Reports*, 4, 5054.

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

Yamaguchi, M. and Demura, T. (2010) Transcriptional regulation of secondary wall formation controlled by NAC domain proteins. *Plant Biotechnology*, 271, 237-242.

Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., Shao, H., Wang, X., Wang, Z.Y. and Dixon, R.A. (2013) Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in *Arabidopsis*. *Plant Cell*, 25, 3976–3987.

Zhong, R., Demura, T. and Ye, Z.H. (2006) SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell*, 18, 3158–3170.

Zhong, R., Richardson, E.A. and Ye, Z.H. (2007a) The MYB46 transcription factor is a direct target of SND1 and regulates secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell*, 19, 2776–2791.

Zhong, R., Richardson, E.A. and Ye, Z.H. (2007b) Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. *Planta*, 225, 1603-1611.

Zhong, R. and Ye, Z.H. (2009) Transcriptional regulation of lignin biosynthesis. *Plant Signaling & Behavior*, 4, 1028-1034.

Zhong, R. and Ye, Z.H. (2010) The poplar PtrWNDs are transcriptional activators of secondary cell wall biosynthesis. *Plant Signaling & Behavior*, 5, 469–472.

Zhong, R., Lee, C. and Ye, Z.H. (2010a) Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends in Plant Science*, 15, 625-632.

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

Zhong, R., Lee, C. and Ye, Z.H. (2010c) Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. *Molecular Plant*, 3, 1087-1103.

Zhong, R., Lee, C., McCarthy, R.L., Reeves, C.K., Jones, E.G. and Ye, Z.H. (2011) Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant Cell Physiology*, 52, 1856–1871.

Zhong, R., McCarthy, R.L., 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.

Zhou, J., Lee, C., Zhong, R. and Ye, Z.H. (2009) MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis. Plant Cell*, 21, 248-266.

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

ARTÍCULO 3

The role of arginine metabolic pathway during embryogenesis and germination in maritime pine (Pinus pinaster Ait.)

María Teresa Llebrés, María Belén Pascual, Sandrine Debille, Jean François Trontin, Luc Harvengt, Concepción Avila and Francisco M. Cánovas

Tree Physiology (2017) 3, 1-14.

1. Abstract

Vegetative propagation through somatic embryogenesis is critical in conifer biotechnology towards multivarietal forestry that uses elite varieties to cope with environmental and socio-economic issues. An important and still sub-optimal process during *in vitro* maturation of somatic embryos is the biosynthesis and deposition of storage proteins, which are rich in amino acids with high nitrogen content, such as arginine. Mobilization of these nitrogen-rich proteins is essential for the germination and production of vigorous somatic seedlings. Somatic embryos accumulate lower levels of nitrogen reserves than zygotic embryos at a similar stage of development. To understand the molecular basis for this difference, the arginine metabolic pathway has been characterized in maritime pine (*Pinus pinaster* Aiton). The genes involved in arginine metabolism have been identified and GFP-fusion constructs were used to locate the enzymes in different cellular compartments and clarify their metabolic roles during embryogenesis and germination. Analysis of gene expression during somatic embryo maturation revealed high levels of transcripts for genes involved in the biosynthesis and metabolic utilization of arginine. By contrast, enhanced expression levels were only observed during the last stages of maturation and germination of zygotic embryos, consistent with the adequate accumulation and mobilization of protein reserves. These results suggest that arginine metabolism is unbalanced in somatic embryos (simultaneous biosynthesis and degradation of arginine) and could explain the lower accumulation of storage proteins observed during the late stages of somatic embryogenesis.

2. Introduction

Conifers are distributed worldwide and are particularly abundant in the Northern hemisphere, dominating large forest ecosystems and playing essential roles in global carbon fixation as well as the maintenance of biodiversity. Conifers are also of great economic importance since these plants provide a vast range of products of commercial interest, including

wood, pulp, biomass and diverse secondary metabolites (Farjon 2010). Maritime pine (*Pinus pinaster* Ait.) is a broadly planted conifer species in France, Spain and Portugal where it is distributed over approximately 4 million hectares (Bouffier et al. 2013). Maritime pine is also one of the most advanced model trees for genetic and phenotypic studies (Lamy et al. 2014; Plomion et al. 2016), and a large number of molecular and transcriptomic resources are currently available (Canales et al. 2014; Cañas et al. 2015a, 2015b). In addition, biotechnological tools are in development for the mass propagation of maritime pine via somatic embryogenesis in combination with cryopreservation of embryogenic lines (reviewed in Lelu-Walter et al. 2006, 2016; Trontin et al. 2016a). Efficient vegetative propagation in pine would enable rapid deployment and turnover in multivarietal forestry of selected/tested varieties that are better adapted to a changing climate and also to socio-economic considerations (Lelu-Walter et al. 2016). Efficient protocols are also available in pines to achieve the genetic transformation of somatic embryos (SE) and transgenic plant regeneration for reverse genetics applications (Klimaszewska et al. 2004b; Trontin et al. 2007, 2016b). However, maturation of SE remains a critical step in the production of high-quality SE plants in maritime pine. Furthermore, cotyledonary SE typically have reduced conversion rates to plantlets compared to seeds and a lower performance in the field tests during early growth compared with zygotic seedlings (Trontin et al. 2016a). A better understanding of maturation and germination of SE is therefore of paramount importance to improve embryo quality and generate vigorous SE plants that compete with weeds, particularly during the first season after planting in field.

An important process during the maturation phase of embryogenesis is the biosynthesis and deposition of storage proteins. Overall, accumulation of the most abundant storage proteins in maturing and mature SE is much lower than in zygotic embryos (ZE) suggesting an important influence on the quality of SE (Klimaszewska et al. 2004a, 2016; Morel et al. 2014a). In conifer seeds, most of the storage proteins are initially located in the megagametophyte (including in maritime pine, Trontin et al. 2016a) and later, the protein content gradually increases in embryos during maturation. These

storage proteins are characterized by nitrogen (N)-rich amino acids (King and Gifford, 1997). Arginine has the highest N to carbon ratio and is therefore particularly suitable for N storage and transport in living organisms (Llacer et al. 2008; Winter et al. 2015). Arginine constitutes a large portion of the amino acid pool in storage proteins of conifers and therefore arginine biosynthesis is likely a relevant metabolic pathway during pine embryogenesis (Cantón et al. 2005; Cánovas et al. 2007).

Following germination of pine seeds, N reserves are mobilized to support the early stages of plant development until the seedlings initiate autotrophic growth. This mobilization of reserves during germination depends on the activation and synthesis of key enzymes, including those involved in the proteolytic hydrolyzation of storage proteins and catabolism of released amino acids, particularly arginine in a process that is closely synchronized with the emergence of the radicle (King and Gifford 1997). The high accumulation of arginine in the embryo is accompanied by increased arginase (ARG) activity to convert arginine to ornithine and urea (Todd et al. 2001; Todd and Gifford 2003). The subsequent hydrolysis of urea by urease is an important source of ammonium for early seedling development, which is reassimilated into glutamine through the catalytic action of cytosolic glutamine synthetases (GS1a and GS1b) (Avila et al. 1998; Cánovas et al. 2007).

The metabolism of arginine has been relatively unexplored in plant N metabolism. Most of the genes involved in the arginine pathway have been predicted from bacterial and fungal homologs and subsequently identified in *Arabidopsis* (Slocum, 2005). Metabolic conversion of glutamate to arginine occurs through two well-diferentiated pathways, the ornithine pathway and the arginine pathway (Figure 1). The ornithine pathway begins with the acetylation of glutamate into N-acetylglutamate catalyzed by N-acetylglutamate synthase (NAGS). N-acetylglutamate is subsequently phosphorylated, reduced and transaminated to generate N-acetylornithine through the sequential action of the enzymes N-acetylglutamate kinase (NAGK), N-acetylglutamate-5-P-reductase (NAGPR) and N-acetylornithine aminotransferase (NAOAT). Finally, ornithine is produced either by the catalytic activity of N-acetylornithine glutamate acetyltransferase (NAOGACT)

or N-acetylornithine deacetylase (NAOD). In the arginine pathway, ornithine is converted to arginine by the sequential action of ornithine transcarbamylase (OTC), argininosuccinate synthetase (ASS) and argininosuccinate lyase (ASL). Other key enzymes for arginine metabolism include arginase (ARG) and δ-ornithine aminotransferase (δ-OAT), which are involved in the metabolic utilization of arginine.

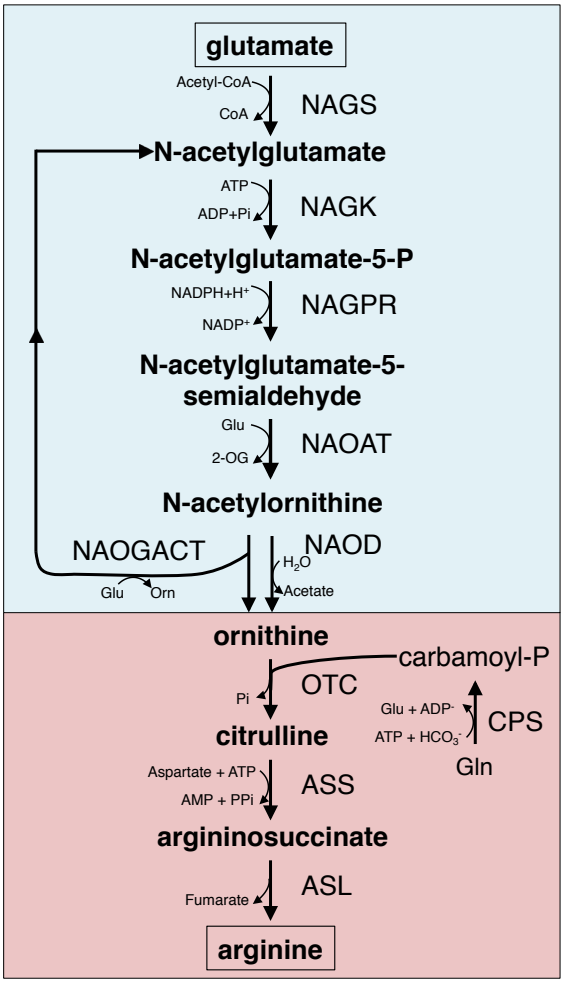


Figure 1.The arginine metabolic pathway. Schematic representation of the metabolic conversion of glutamate to arginine through two well-differentiated pathways: i) the ornithine pathway (in blue) and ii) the arginine pathway (in pink). NAGS, N-acetylglutamate synthase; NAGK, *N*-acetylglutamate kinase; NAGPR, *N*-acetylglutamate-5-P reductase; NAOAT, *N*-acetylornithine aminotransferase; NAOGACT, *N*-acetylornithine-glutamate acetyltransferase; NAOD, *N*-acetylornithine deacetylase; CPS, carbamoyl phosphate synthetase; OTC, ornithine transcarbamoylase; ASS, argininosuccinate synthase; ASL, argininosuccinate lyase.

A limited number of studies have examined the function and molecular regulation of these enzymes. The arginine biosynthetic pathway is primarily located in the plastid, however, many steps still remain poorly characterized, and only a limited number of enzymes have been purified and biochemically characterized (see Winter et al. 2015, for a recent comprehensive review). As arginine is a key amino acid for N storage during embryogenesis and N mobilization during germination, arginine metabolism in conifers deserves special attention.

In the present study, the molecular characteristics of enzymes involved in arginine metabolism, as well as their subcellular localization and their transcriptional levels during development and germination of SE and ZE were investigated in maritime pine.

3. Material and methods

3.1. Plant material

Somatic embryos The embryogenic cell line PN519 of maritime pine (*Pinus pinaster* Ait.) has been used in the present study. This model line initiated at the Institut technologique Forêt, Cellulose, Bois-Construction, Ameublement (FCBA) in 1999 is amenable to both genetic transformation and plant regeneration through somatic embryogenesis and has been extensively characterized during the past 15 years (Lelu-Walter et al. 2016; Trontin et al. 2007, 2016a, 2016b). The cell line was regrown from cryopreserved tissue two months prior to the start of these experiments. Proliferation was performed on modified Litvay medium (MLV) with low PGRs as defined by Klimaszewska et al. (2001), 2.2 µM 2,4-D (2,4-dichlorophenoxyacetic acid) and 2.2 µM BA (6-benzyladenine), and weekly subcultures onto fresh medium were performed one month prior to the maturation experiments. MLV maturation medium is similar to proliferation medium except for the higher the content of sucrose (60 g L^{-1}) and gellan gum (Phytagel, Sigma, 9 g L^{-1}), and plant growth regulators were also replaced with abcisic acid (ABA) at 80 µM. Proliferation and maturation were conducted at 24ºC in darkness inside a culture chamber.

Samples were collected at three different stages of maturation: early-stage translucent SE (ES1, after 4-6 weeks), pre-cotyledonary opaque SE (ES2, after 6-10 weeks) and cotyledonary SE (ES3) that were collected after 14 weeks of maturation.

ES3 cotyledonary embryos were germinated in darkness at 24ºC. Germination medium was identical to MLV but without plant growth regulators and with 5 g L^{-1} gellan gum. Samples were collected after 2 days (2D), 4 days (4D) and 9 days (9D) of germination, frozen in liquid nitrogen and stored at -80 ºC for RNA and protein extraction.

Seeds, zygotic embryos and megagametophytes Zygotic embryos (ZE) were excised from seeds collected from a single maritime pine (*Pinus pinaster* Ait.) seed orchard (Picard, Saint-Laurent-Médoc, France) from July to November 2015. ZE were sampled at different developmental stages according to de Vega-Bartol et al. (2013): pre-cotyledonary ZE (PC, early to mid-July), early cotyledonary ZE (EC, mid to late-July), cotyledonary immature ZE (C, from early August to early September) and cotyledonary, mature ZE (M, November). The ZE samples were frozen in liquid nitrogen and stored at -80 ºC until use.

Mature seeds of maritime pine (*Pinus pinaster Ait*.) provided by the Centro de Recursos Genéticos Forestales “El Serranillo” (Ministerio de Medio Ambiente y Medio Rural y Marino, Spain) were soaked in destilled water for two days with aeration. The megagametophytes were excised and samples of 10 ZE embryos were collected in triplicate (imbibed embryo, EE). Remaining seeds were germinated in vermiculite at 24ºC under a 16 h light/ 8 h dark photoperiod and samples of ten embryos were collected in triplicate after 4 (EG4) and 9 (EG9) days of growth. Samples were frozen in liquid nitrogen and stored at -80ºC until use.

Nicotiana benthamiana L. seeds were sown in pots and cultivated in a controlled growth chamber at 24ºC and 16 h light/8 h dark photoperiod for 5 weeks. This model plant was used for subcellular localization via agroinfiltration.

3.2. Protein extraction

Frozen samples (35-50 mg FW) were homogenized in a lysis buffer with glass beads in a mortar. The buffer contained 50 mM Tris-HCl (pH 8), 2 mM EDTA, 10% (v/v) glycerol, 2% (w/v) SDS, 2% (w/v) PVPP and 5% (v/v) β-mercaptoethanol and it was used in a proportion of 3 μL of buffer per 1 mg of tissue. The homogenates, after incubation 5 min at 95 ºC, were centrifuged for 30 min at 18,000 x g. The supernatants were collected for protein concentration measurements using the Bradford Protein Assay (Bio-Rad). Proteins (20 μg) were analyzed by SDS-PAGE in 12.5% (w/v) acrylamide/bisacrylamide gels followed by Coomassie Brilliant Blue staining. Seven protein bands that differentially accumulated were excised using a sterile scalpel and analyzed by LC-MS/MS.

3.3. Protein identification by mass spectrometry

The protein bands were washed with acetonitrile in 25 mM ammonium bicarbonate until the bands were completely destained. Afterwards, they were vacuum dried and proteins digested with 12.5 ng mL^{-1} trypsin in 25 mM ammonium bicarbonate (de la Torre et al. 2007). Then, the samples were treated with 0.5% (v/v) trifluoroacetic acid (TFA), desalted and concentrated by using mC-18 Spin column (Thermo Scientific). The free peptides were vacuum-dried and solubilised in 2% (v/v) acetonitrile and 0.05% TFA. The peptides were analyzed by nano-flow High-Performance Liquid Chromatography (HPLC)-electrospray tandem mass spectrometry (LC-MS/MS). The separation was performed using a 300 mm x 5 mm Dionex Acclaim PepMap100 C18 column (Thermo Scientific), followed by ionization with the nanospray ion source, and placed into an Orbitrap Fusion mass spectrometer (Thermo Scientific). MS data (Full Scan) were recorded in the positive ion mode over the 400-1500 m/z range.

Data analysis was performed with Proteome Discoverer 2.1 (Thermo Scientific) and combined MASCOT (Matrix Science) and SEQUEST HT searches against the SustainpineDB (Canales et al. 2014).

3.4. Subcellular localization

Full-length cDNAs sequences were obtained in SustainPineDB v.3.0 (http://www.scbi.uma.es/sustainpinedb/sessions/new). Each gene was amplified by PCR using the specific primers listed in Table S1. The resulting PCR product was cloned into the pDONR207 (Invitrogen) and transferred for recombination-based cloning to the final gateway vector pGWB5. All constructs were confirmed by sequencing. Empty plasmid pGWB5 was used as a negative control and pGWB6 (p35S-GFP) as positive control. The *Agrobacterium tumefaciens* strain C58C1 was transformed by electroporation with recombinant plasmids expressing the proteins of interest. *Nicotiana benthamiana* leaves (5 weeks old) were syringe infiltrated with cultures containing pGWB5 constructs mixed with cultures containing P19, both with an optical density at 600 nm of 0.5, according to the procedures described previously (Liu et al. 2002). Subcellular localization of proteins was examined by confocal microscopy 36 to 48 h after agroinfiltration.

Mitotracker ◦ Red FM (Thermo Fisher) was used as red-fluorescent dye to confirm the subcellular localization in the mitochondria (Figure S1). *A* stock solution of the red-fluorescent dye was prepared in DMSO (Sigma). Sections of infiltrated leaves were incubated in 200 nM of the red-fluorescent dye in phosphate buffer saline. Labelling was conducted in the dark at room temperature for 1 hour. Confocal microscopy was performed using a Leica SP5 Laser Scanning Confocal Microscope equipped with HyD and PMT detectors, AOBS (Acousto-Optical beam splitter) and a spectral detection system. GFP fluorescence and chloroplast autofluorescence was detected using Argon laser excitation at 488 nm. The mitochondrial labelling was detected at 581 nm. Images were acquired using either Plan APO 40x 1.30 NA or Plan APO 63x 1.40 NA oil immersion objectives. Laser intensity and detector settings were optimized according to the imaging conditions and GFP signal intensities. The images were processed using Leica LAS and FIJI ImageJ software (version 4.1.1).

3.5. RNA extraction, cDNA synthesis and real-time quantitative PCR (qPCR)

Extraction of RNA was performed as described by Canales et al. (2012) and quantified using a NanoDrop© ND-1000 spectrophotometer. Synthesis of cDNA was performed with 5X iScript™ cDNA Synthesis Kit (Bio-Rad). The qPCR analysis was performed in a thermal cycler CFX384 (Bio-Rad). Each reaction proceded in a total volume of 10 μL, 5 μL of SsoFst™ EvaGreen® Supermix (Bio Rad), 2 μL cDNA (5 ng μL^{-1}) and 0.5 μL of 10 mM of a specific primer. *Actin-7* (18113) was used as a reference gene. Sequences of specific primers are listed in Table S2. Relative expression profiles for each gene were obtained employing the R package (Ritz and Spiess, 2008) and normalized to the reference gene.

4. Results

4.1.The protein profiles differed in the zygotic and somatic embryos of maritime pine

Protein profiles during the embryogenesis and germination of maritime pine somatic (SE) and zygotic (ZE) embryos were compared (Figure 2). The profiles of soluble polypeptides were resolved using SDS-PAGE during the maturation of SE and ZE embryos (Figure 2a). The most abundant proteins in ZE at the cotyledonary immature (C) and cotyledonary mature (M) stages have apparent molecular sizes of 47, 35, 20 and 15 kDa and were previously identified as vicilin-like and legumin-like storage proteins, globulin and albumin using LC–MS/MS, respectively (Morel et al. 2014a). These major N storage proteins were much less abundant in cotyledonary SE (ES3 collected after 14 weeks maturation) than in cotyledonary immature (C) and mature (M) ZE. The relative abundance of vicilin-like and legumin-like storage proteins, globulin and albumin was considerably lower at the pre-cotyledonary opaque (ES2) stage with barely detectable levels at the early-stage translucent (ES1). Three other minor polypeptides of 29.4, 24.7 and 14.9 kDa were also more abundant in ZE. A polypeptide of 17.5 kDa was primarily present in ZE at the cotyledonary mature (M) stage. By contrast, polypeptides of 16.7, 15.8 and

10.9 kDa were clearly more represented in SE with higher relative abundance before the cotyledonary stages (ES1 and ES2 stages). To further explore the molecular basis of these differences in the protein profiles of SE and ZE, the polypeptides of 29.4, 24.7 16.7, 15.8, 14.9 and 10.9 kDa were excised from polyacrylamide gels and subjected to HPLC-electrospray tandem-mass spectrometry analysis (LC-MS/MS).

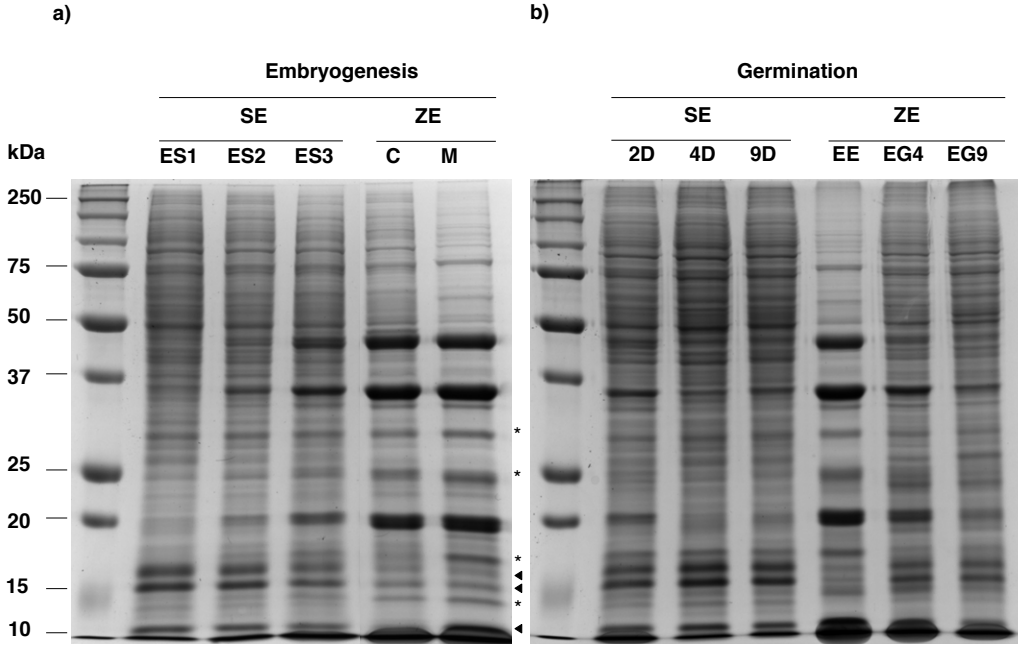


Figure 2. Protein profiles in somatic and zygotic embryos during maturation and germination. Total protein (15 mg) from somatic (SE) and zygotic (ZE) embryos during embryogenesis (a) and germination (b) were fractionated by SDS-PAGE and stained with Coomassie blue. Representative gels are depicted. Asterisks and arrowheads indicate the bands excised from ZE and SE respectively, for mass spectrometry analysis. Molecular masses (kDa) of protein markers are indicated. ES1: early-stage translucent SE; ES2: pre-cotyledonary opaque SE; ES3: cotyledonary SE; C: cotyledonary immature ZE; M: cotyledonary mature ZE; 2D, 4D and 9 D: two, four and nine days in germinating medium, respectively; EE: imbibed zygotic embryos; EG4 and EG9: embryos four and nine days after imbibition.

The most abundant peptides identified in the different analyzed gel bands are listed in Table S3. Notably, the major protein represented in the ES samples corresponded to late embryogenesis abundant proteins (LEA), matching polypeptides of 16.7, 15.8 and 10.9 kDa. Moreover, the peptide

footprints corresponded to different unigenes encoding LEA proteins in the SustainpineDB (Canales et al. 2014).

The profiles of soluble polypeptides during the germination of SE and ZE embryos showed that N storage proteins significantly decreased in ZE after 9 days of germination (Figure 2b). Minor differences in protein profiles were observed between SE and ZE embryos at the last stages of germination examined (4 and 9 days) despite differences in the N storage (Figure 2b).

4.2. Identification of genes encoding enzymes of the arginine metabolic pathway in maritime pine

A first step in the present study was the identification of genes involved in arginine metabolism in the transcriptome of maritime pine through a search in the SustainpineDB (Canales et al. 2014). Seven full-length cDNAs (FLcDNAs) encoding enzymes involved in ornithine biosynthesis were identified, N-acetylglutamate synthase (*PpNAGS*), N-acetylglutamate kinase (*PpNAGK*), N-acetylglutamate-5-P reductase (*PpNAGPR*), N-acetylornithine aminotransferase (*PpNAOGACT*), N-acetylornithine: glutamate acetyltransferase (*PpNAOAT*), N-acetylornithine: glutamate acetyltransferase (*PpNAOGACT*) and N-acetylornithine deacetylase (*PpNAOD*) (Table 1).

A N-terminal sequence for plastid targeting was predicted for *PpNAGS*, *PpNAGK*, *PpNAGPR* and *PpNAOAT* encoding mature polypeptides with high levels of identity (Table S4) to their *Arabidopsis* counterparts (Slocum, 2005). No pre-sequences for organellar targeting were identified in the open reading frames (ORF) of *PpNAOGACT* and *PpNAOD*. Four additional FLcDNAs involved in arginine biosynthesis were identified (Table 1), carbamoyl-P synthetase small subunit (*PpCPS*), ornithine transcarbamoylase (*PpOTC*), argininosuccinate synthetase (*PpASS*) and argininosuccinate lyase (*PpASL*). N-terminal sequences for plastid targeting were predicted in the ORFs for *PpASS* and *PpASL*. Two FLcDNAs encoding enzymes involved in arginine catabolism were also identified, arginase (*PpARG*) and δ-ornithine aminotransferase (*Ppδ-OAT*). The ORFs from these two sequences contained N-terminal for targeting to mitochondria (Table 1).

Table 1. Genes of the arginine metabolic pathway in *Pinus pinaster*.

Name	Gene ID	FL cDNA (bp)	ORF (bp)	Polypeptide (Da)	Subcellular prediction[1]	Processed protein (Da)	pI
PpNAGS	sp_v3.0_unigene5514	2595	1926	69,406	Plastid (79)	61,065	6.1
PpNAGK	sp_v3.0_unigene15977	1824	1059	36,991	Plastid (51)	31,358	8.3
PpNAGPR	sp_v3.0_unigene5400	1924	1248	45,850	Plastid (49)	40,284	6.7
PpNAOAT	sp_v3.0_unigene5428	2030	1482	53,042	Plastid	-	-
PpNAOGACT	sp_v3.0_unigene7147	2222	1524	52,484	-	-	-
PpNAOD	sp_v3.0_unigene1654	1729	1296	47,427	-	-	-
PpCPS	sp_v3.0_unigene8325	1921	1365	49,210	-	-	-
PpOTC	sp_v3.0_unigene6197	1900	1155	42,011	-	-	-
PpASS	sp_v3.0_unigene6320	2311	1539	56,407	Plastid (37)	52,383	6.0
PpASL	sp_v3.0_unigene5329	2254	1590	58,548	Plastid	-	-
PpARG	sp_v3.0_unigene23824	1554	1026	37,303	Mito	-	-
PpδOAT	sp_v3.0_unigene5775	2230	1407	51,234	Mito	-	-

[1]Chloroplast transit peptides prediction and amino acid residues (in brackets) using Predotar and ChloroP databases.

PpNAGS, N-acetylgutamate synthase; *PpNAGK*, N-acetylglutamate kinase; *PpNAGPR* N-acetylglutamate-5-P reductase; *PpNAOAT*, N2-acetylornithine aminotransferase; *PpNAOGACT*, N-acetylornithine:glutamate acetyltransferase; *PpNAOD*, N-acetylornithine deacetylase; *PpCPS*, carbamoyl-P synthetase small subunit; *PpOTC*, ornithine transcarbamoylase; *PpASS*, argininosuccinate synthetase; *PpASL*, argininosuccinate lyase; *PpARG*, arginase; *PpδOAT,* δ-ornithine aminotransferase.

4.3. Subcellular localization of the arginine metabolic pathway in maritime pine

To further characterize the biosynthesis and utilization of arginine, we determined the subcellular localization of the enzymes in the entire pathway. FLcDNAs for all genes identified in the maritime pine transcriptome (Table 1) were PCR-amplified using specific primers (Table S1) and GFP fusions were transiently expressed in *N. benthamiana* leaves. First, the subof enzymes involved in ornithine biosynthesis was examined. The transiently expressed constructs for PpNAGS, PpNAGK, PpNAGPR, PpNAOAT and PpNAOGACT displayed GFP flurorescence associated with chloroplasts as revealed in the corresponding images of cholorophyll red autofluorescence (Figure 3). Examination of the merged images showed co-localization of chlorophyll with the different protein gene products in the chloroplasts visualized in yellow color, confirming plastidic localization. While PpNAGS, PpNAGPR, PpNAOAT and PpNAOGACT displayed a diffuse fluorescence pattern throughout the plastids, PpNAGK showed a punctate distribution of the GFP signal (Figure 3, PpNAGK). PpNAOD fluorescence was primarily distributed throughout the cytosol and no GFP signal was detected in the chloroplasts (Figure 3).

Next, the subcellular localization of enzymes involved in arginine biosynthesis and catabolism was determined (Figure 4). Using a similar approach to that described above, all GFP fusions of PpCPS, PpOTC, PpASS and PpASL were localized to chloroplasts. The co-localization of PpCPS, PpOTC, PpASS and PpASL and chlorophyll in the plastids was indisputably confirmed in the merged images, resulting in chlorophyll and GFP flurorescence. Figure 4 also shows that the GFP-tagged enzymes involved in arginine catabolism, PpARG and PpOAT, were clearly localized outside the chloroplasts but not distributed throughout the cytosol as previously observed for PpNAOD (Figure 3). To identify the precise subcellular localization of PpARG and PpOAT, the mitochondrial prediction derived from the analysis of the ORFs was considered (Table 1). Consequently, the localization of these enzymes was compared with a mitochondrial marker (Mitotracker Red FM), and the results are shown in Figure S1. The co-localization of GFP signals and the marker in the merged images confirmed mitochondrial localization.

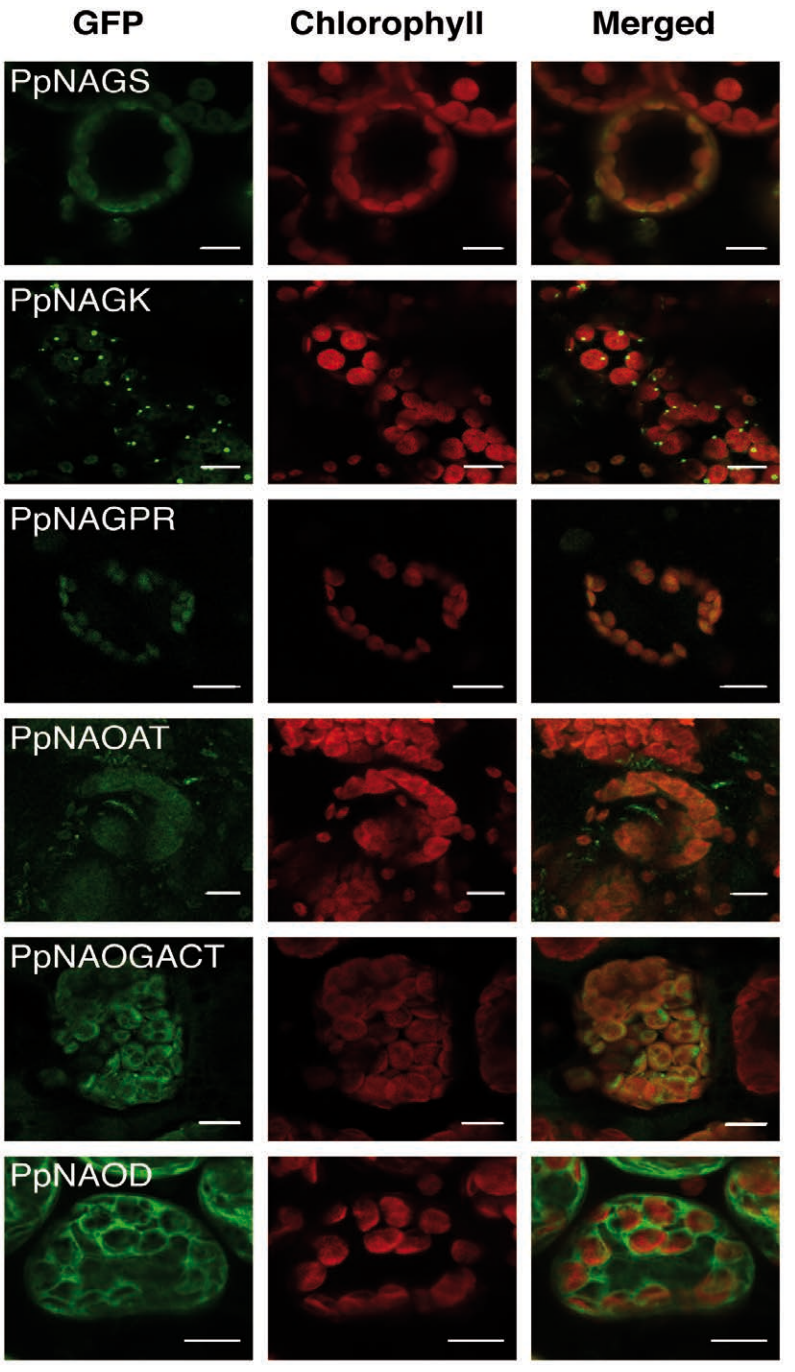


Figure 3. Subcellular localization of enzymes of the ornithine biosynthesis pathway. *Nicotiana benthamiana* leaves were infiltrated with *A. tumefaciens* containing the constructs of interest fused to GFP. Proteins were transiently expressed and their intracellular localizations were determined by confocal laser scanning microscopy. The GFP signal (green) is shown in the first channel, chlorophyll autofluorescence (red) in the second and the third channel shows merged images. Scale bar represents 10 μm.

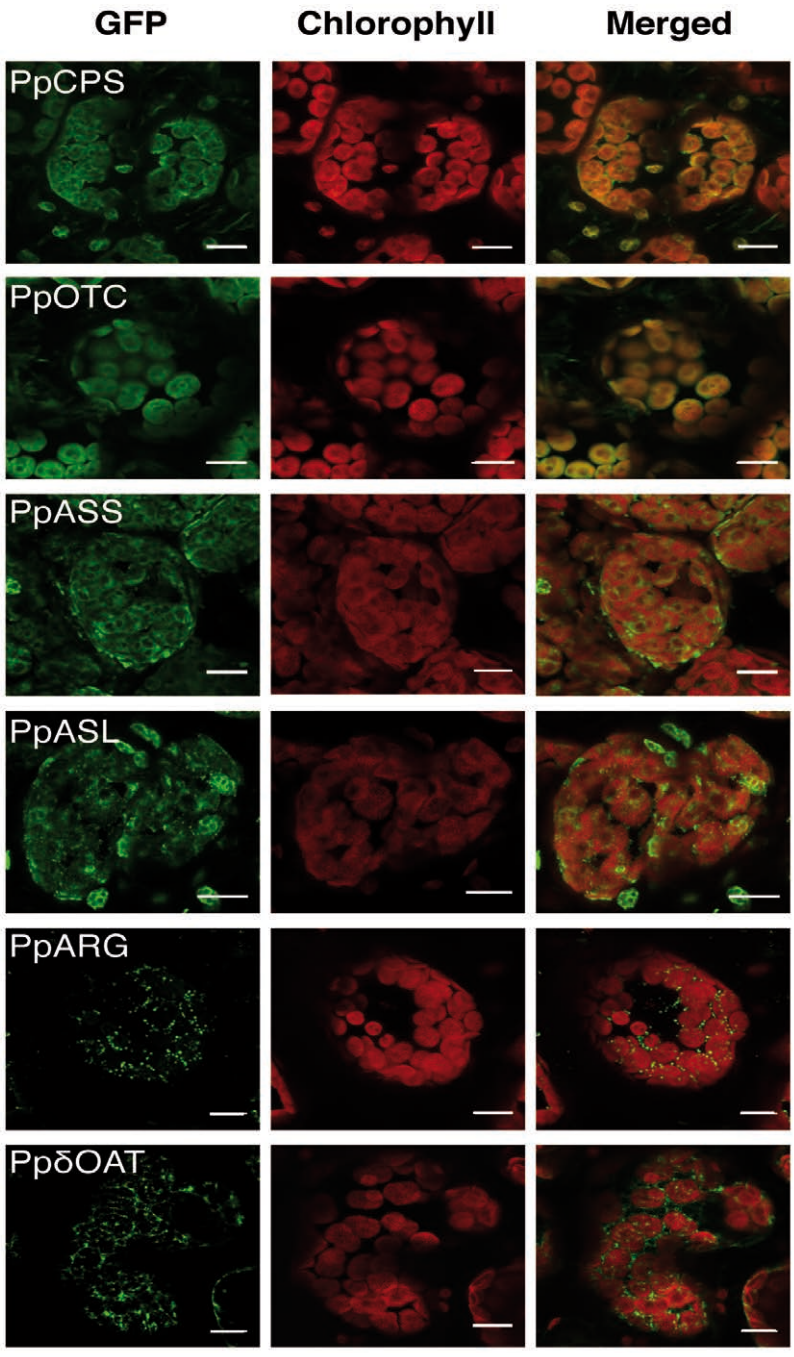


Figure 4. Subcellular localization of enzymes of the arginine metabolic pathway. *Nicotiana benthamiana* leaves were infiltrated with *A. tumefaciens* containing the constructs of interest fused to GFP. Proteins were transiently expressed and their intracellular localization determined by confocal laser scanning microscopy. The GFP signal (green) is shown in the first channel, chlorophyll autofluorescence (red) in the second and the third channel shows the merged images. Scale bar represent 10 μm.

4.4. Transcript levels for enzymes of the arginine metabolic pathway in SE and ZE

To assess the ability of maritime pine embryos for arginine biosynthesis and metabolic utilization, the expression levels of genes involved in the arginine metabolic pathway were compared during the maturation of SE and ZE. Total RNA was extracted from SE and ZE at several stages of development and the relative transcript abundance for all genes was determined by qPCR analysis using specific primers (Table S2). Overall, similar transcript levels were observed for all genes involved in the ornithine (*PpNAGS*, *PpNAGK*, *PpNAGPR*, *PpNAOGACT* and *PpNAOD*) and arginine (*PpCPS*, *PpOTC*, *PpASS* and *PpASL*) biosynthetic pathways at the three stages (ES1, ES2 and ES3) examined during SE maturation (Figure 5b). By contrast, transcript levels for all genes were generally low at early stages of ZE maturation (PC and EC, which are similar to ES2 and ES3, respectively) but significantly increased at the end of maturation (stages C and M) reaching higher levels than those observed in SE. These expression profiles were conserved for all genes involved in ornithine and arginine biosynthesis. Interestingly, transcript levels of *PpARG* and *PpδOAT*, two genes involved in arginine catabolism, were significantly higher during the maturation of SE than during the maturation of ZE (Figure 5b).

4.5. Transcript levels for enzymes of the arginine metabolic pathway during the germination of SE and ZE

To further explore the arginine metabolic pathway in maritime pine embryos, transcript abundance of genes involved in this pathway was also determined using qPCR analysis during the germination of SE and ZE. Overall, similar transcript levels were observed for all genes involved in ornithine and arginine biosynthesis during SE germination (Figure 6). However, in ZE, the expression levels of most genes were much higher in imbibed embryos and subsequently declined following germination. Exceptions to this general profile were *PpNAOD, PpNAOGACT* and *PpASL* genes for which low transcript levels were observed during germination. With regard to arginine catabolism, significant differences in the transcript levels of *PpARG* and *PpδOAT* were observed in SE and ZE embryos at the end of the germination.

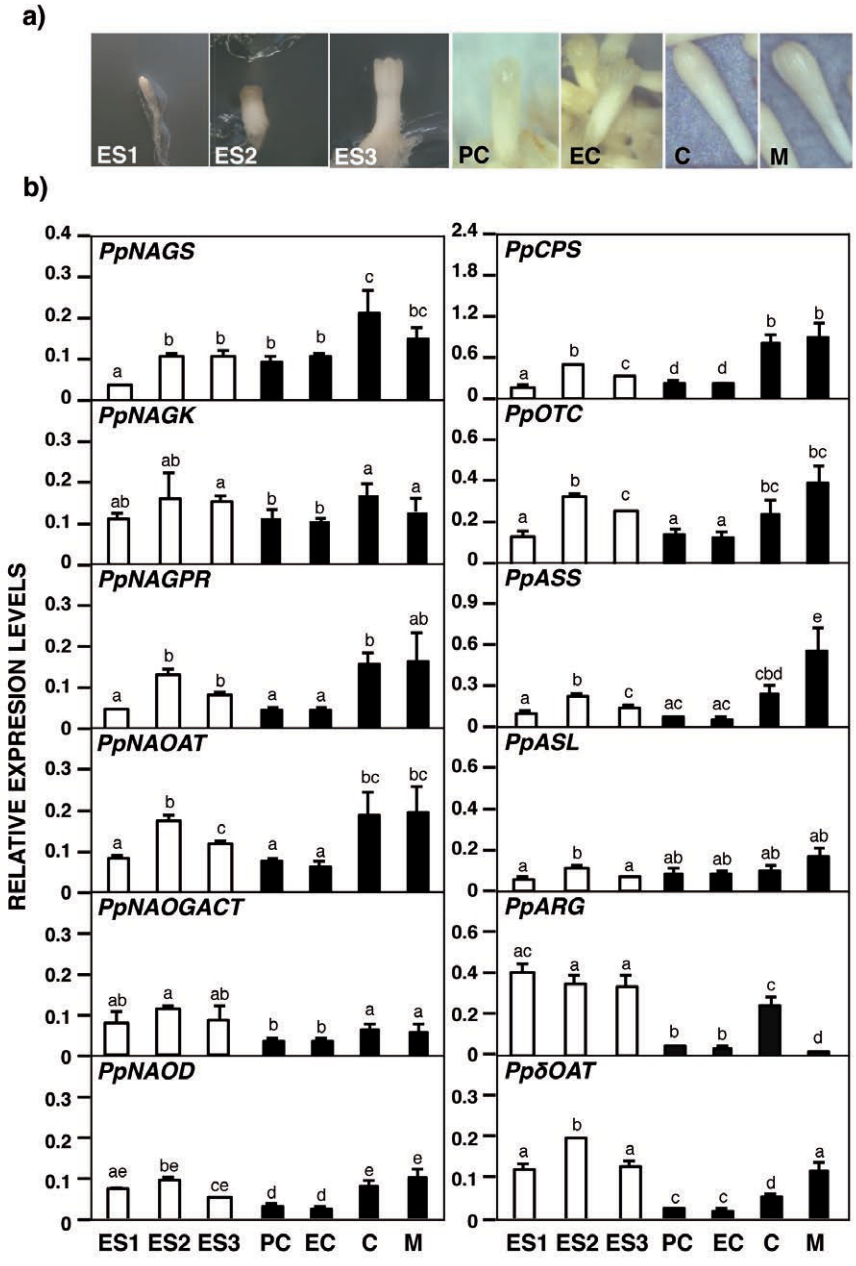


Figure 5. Expression patterns of genes involved in arginine biosynthesis and utilization during last stages of somatic and zygotic embryogenesis in *P. pinaster*. a) Representative images of somatic and zygotic embryos at final stages of embryogenesis (maturation stages). Somatic embryos: ES1, early-stage translucent; ES2, pre-cotyledonary opaque; ES3, cotyledonary. Zygotic embryos: PC, pre-cotyledonary; EC, early cotyledonary; C, cotyledonary immature; M, cotyledonary partially mature. b) qPCR expression analysis in somatic and zygotic embryos during maturation. The expression level for all genes was normalized to that of *actin*-7 as reference gene. Different letters above bars indicate significant differences between samples at P <0.05. Bars represent mean values of three assays, with three biological replicates each ± standard deviation.

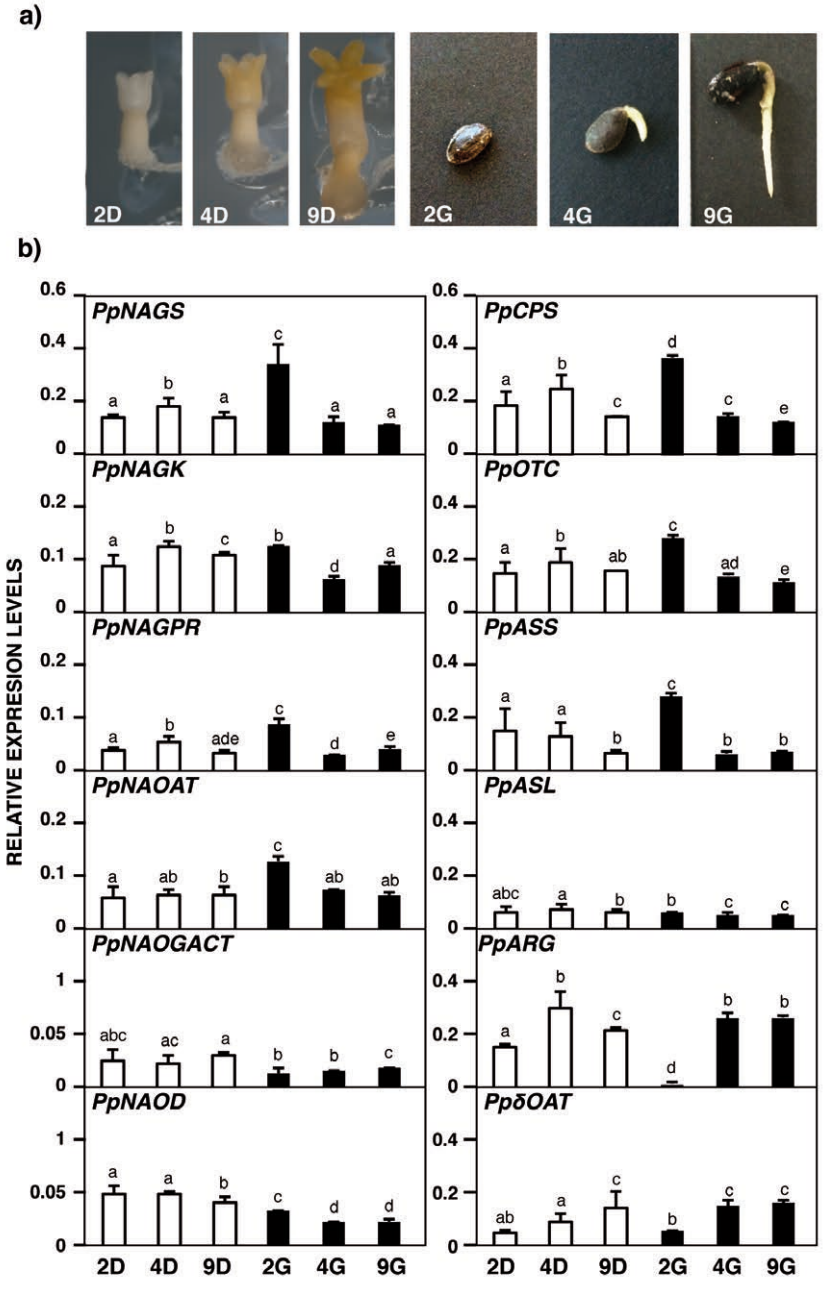


Figure 6. Expression patterns of genes involved in arginine biosynthesis and utilization during germination of somatic and zygotic embryos of *P. pinaster.* a) Representative images of somatic and zygotic embryos at different germination stages. Mature SE in germination media for 2 days (2D), 4 days (4D) and 9 days (9D). Seeds germinated for 2 days (2G), 4 days (4G) and 9 days (9G). b) qPCR expression analysis in somatic and zygotic embryos during germination. The expression level for all genes was normalized to that of *actin-7* as reference gene. Different letters above bars indicate significant differences between samples at P <0.05. Bars represent mean values of three assays, with three biological replicates each ± standard deviation.

5. Discussion

Maritime pine (*P. pinaster*) SE accumulates lower levels of major storage proteins than ZE during the late stages of embryogenesis (Figure 2; Morel et al. 2014a). Moreover, maturation duration (10 to 14 weeks) is not a critical factor for protein accumulation in cotyledonary embryos (Morel el al. 2014a). Similar results have been described in other pine species such as *Pinus strobus* (Klimaszewska et al. 2004a), *Pinus taeda* (Brownfield et al. 2007) and *Pinus sylvestris* (Lelu-Walter et al. 2008). In fact, storage protein content has been suggested as a suitable molecular marker of SE quality, determining to a greater extent the success of germination and survival of SE plants (Klimaszewska et al. 2004a; Tereso et al. 2007; Miguel et al. 2016).

An additional characteristic of maturation process is the acquisition of desiccation tolerance, which in pine SE is externally induced by increasing the osmotic pressure (high sucrose concentration) and lowering the water availability in the culture medium by physical means (high gellan gum concentration). Several extra LEA proteins structurally affiliated with groups 3 and 4 (Amara et al. 2014) were accumulated in SE but were not present in ZE (Table S3). LEA proteins are directly involved in the adaptive response of higher plants (including conifers, reviewed in Miguel et al. 2016, Trontin et al. 2016c) to water deficit and may also stabilize a partially unfolded state, preventing protein aggregation (Goyal et al. 2005; Hand et al. 2011; Furuki et al. 2012). The differential accumulation of LEA suggests that SE do not achieve desiccation tolerance to the same degree as ZE during the last stages of maturation. Morel et al. (2014a) reported similar conclusions after a multi-scale analysis (water content, content of various oligosaccharides, total protein content) of cotyledonary SE and maturing ZE. Cotyledonary SE were most similar to cotyledonary ZE at a partially mature, fresh (undessicated) stage. In particular the ratio between stachyose+raffinose (oligosaccharides of the raffinose family, RFOs, involved in desiccation tolerance) and sucrose significantly increased in ZE during maturation but remained low in cotyledonary SE. Interestingly, Morel et al. (2014a) detected various LEAs and other stress-related proteins (HSPs, heat-shock proteins) that were

proposed as putative generic markers of the fresh, cotyledonary stage of embryo maturation. In *Picea abies*, SE developed on the maturation medium supplemented with sucrose as the main carbon source contained high levels of raffinose and LEA and preferentially accumulated starch and lower levels of storage proteins (Businge et al. 2013). It is suggested that acquisition of desiccation tolerance is promoted through a sucrose-based maturation medium and that it could be a concurrent process for the accumulation of storage proteins. Businge et al. (2013) reported that maturation treatment could induce changes in nitrogen metabolism in mature embryos of *P. abies* through differential expression of key enzymes for glutamine, glutamate and arginine synthesis. Consistently, metabolite profiling also revealed significant amino acids associated with nitrogen metabolism and polyamine biosynthesis during late embryogeny such as ornithine, arginine and asparagine (Businge et al. 2012).

Arginine is very abundant in maritime pine storage proteins (Allona et al. 1992; 1994), and therefore arginine biosynthesis is of paramount importance during embryogenesis. The recent assembly of maritime pine transcriptome (Canales et al. 2014) enabled the identification of genes involved in arginine metabolism. As shown in Table 1, FLcDNA sequences for all enzymes in the ornithine and arginine pathways have been identified. Overall, their molecular characteristics are similar to the corresponding counterparts in *Arabidopsis* (Slocum, 2005, Winter et al. 2015), with minor differences (Table S4). Furthermore, the availability of FLcDNA sequences enabled the prediction of the putative organellar localization of a few but not all enzymes in this pathway. The determination of the precise subcellular compartmentation of the pathway is critical to understand arginine biosynthesis and metabolic utilization in maritime pine. Although the plastidic localization of the ornithine and arginine pathways is assumed through N-terminal prediction of gene sequences, this localization has only been experimentally confirmed for several enzymes (Slocum, 2005). Taking advantage of the availability of FLcDNAs for all enzymes of the pathway, their subcellular localization was systematically determined in *N. benthamiana*. The enzymes of the ornithine and arginine pathways were localized in the plastids,

with the exception of PpNAOD, which was localized in the cytosol (Figures 3 and 4). These results are consistent with previous reports describing the plastidic localization of NAGK (Chen et al. 2006) and NAOAT (Fremont et al. 2013) in *Arabidopsis* and ASL in rice (Xia et al. 2014). Furthermore, these localization studies support the notion that ornithine is synthesized in the chloroplast through the cyclic ornithine pathway and in the cytosol via the enzymatic reaction catalyzed by NAOD. These results imply the existence of separate pools of ornithine in two cellular compartments. In the plastids, ornithine can be metabolically used as precursor for arginine biosynthesis, whereas in the cytosol, ornithine would be channelled for the biosynthesis of polyamines and other nitrogenous compounds, such as alkaloids (Facchini, 2001; Majumdar et al. 2013; Tiburcio et al. 2014). Consistently, the downregulation of NAOD in *Arabidopsis* resulted in decreased levels of ornithine and altered contents of putrescine and spermine (Molesini et al. 2015). Ornithine levels are apparently controlled in plants through the first step in the pathway, a reaction catalyzed by NAGS. In fact, the overexpression of NAGS in tomato resulted in the accumulation of high levels of ornithine (Kalamaki et al. 2009). However, arginine biosynthesis is allosterically regulated by feedback inhibition of NAGK (Slocum, 2005).

From the practical point of view, it is interesting to determine the molecular basis of the lower accumulation and deposition of storage proteins in SE, in spite of the lack of limitation in the availability of N during late embryogenesis (maturation). Is arginine biosynthesis less efficient in SE than in ZE? Are there differences in the expression of genes involved in arginine biosynthesis? To answer these questions, the relative expression levels of all genes involved in the pathway were compared in somatic and zygotic maritime pine embryos. In ZE, gene expression remained low at the initial stages of maturation and increased at the end of this phase, suggesting that arginine biosynthesis is enhanced during the deposition of storage proteins (Figure 5b). By contrast, the observed transcript levels of the ornithine and arginine pathways were high and similar during the three SE maturation stages analyzed (Figure 5b). These results suggest that arginine biosynthesis occurs throughout the maturation of SE. One possible explanation for the

above results could be that arginine degradation was already and concomitantly activated in SE, consistent with the increased levels observed in SE for *PpARG* and *PpδOAT,* two genes involved in arginine catabolism. Moreover, the expression of *ARG* in pine somatic embryos has previously been reported, suggesting an overlap between late embryogenesis and precocious germination (Pérez Rodríguez et al. 2006). Taken together, these results strongly suggest that the pathways for biosynthesis and degradation of arginine simultaneously function in SE and likely reflect the continuous deposition and breakdown of storage proteins. These data support the lower accumulation of storage proteins observed in maritime pine SE compared with ZE (Figure 2a). Consistently, the early mobilization of storage proteins by enhanced protease activity has previously been described in somatic embryos of oil palm (Aberlenc-Bertossi et al. 2008).

The arginine metabolic pathway is regulated at transcriptional and post-transcriptional levels in yeast (Ljungdahl and Daignan-Fornier, 2012). In comparison, little is known about the regulation of the arginine metabolism in plants, particularly at transcriptional level. The results shown in Figure 5 suggest that the expression of genes involved in the ornithine and arginine pathways are coordinately regulated in maritime pine. This assumption is supported by the existence of similar cis-elements in the regulatory region of several genes (Figure S2). However, additional studies are needed to identify specific transcription factors that bind to gene promoters and regulate the transcription of the arginine pathway.

Somatic embryos accumulated higher levels of starch and lower levels of storage proteins than zygotic embryos (Joy et al. 1991; Tereso et al. 2007). In maritime pine somatic embryogenesis, starch accumulation was promoted by low water availability in the maturation medium (Morel et al. 2014a, 2014b). The reduced water availability promoted the downregulation of genes involved in glucose or pentose metabolisms and was associated with increased embryo dry weight and enhanced starch synthesis (upregulation of glucose-1-phosphate adenylyltransferase).

Moreover, it is well documented that sucrose synthase plays an important role in carbohydrate metabolism during maturation (Konrádova et al.

2002) suggesting that somatic and zygotic embryos exhibit a different C/N status, which may explain the differences in the regulation of genes involved in arginine metabolism. The PII protein is a sensor of the C/N status in plants mediated by binding to 2-oxoglutarate and modulation of NAGK activity (Chen et al. 2006). Arginine biosynthesis undergoes the feedback regulation of NAGK (Slocum, 2005; Llacer et al. 2008). When N is abundant, this inhibition is released through interaction with the N sensor protein PII (Chen et al. 2006; Llacer et al. 2008). In addition, PII controls, in a glutamine-dependent manner, the NAGK enzyme, the key step in arginine biosynthesis (Chellamuthu et al. 2014). It is currently unknown whether PII plays a role in the transcriptional regulation of the arginine biosynthetic pathway in plants, but T-DNA insertional mutants in *Arabidopsis* exhibited reduced levels of ornithine, citrulline and arginine (Ferrario-Méry et al. 2006). In pine seedlings, the gene encoding PII-like protein is expressed in different organs, and PII-like protein transcripts are particularly abundant in developing embryos, suggesting a role for PII in the regulation of N metabolism during embryogenesis (Cánovas et al. unpublished). Therefore, the PII protein in maritime pine deserves special attention, particularly considering that a link between PII and storage protein production has previously been proposed (Uhrig et al. 2009).

N is a limiting factor for conifer tree growth and development and mobilization of storage proteins during the germination of SE provides substantially lower levels of N-rich amino acids such as arginine, which are essential during early stages of germination (King and Gifford, 1997). In fact arginine and proline are predominant amino acids representing more than 60% of the total amino acid content in pine zygotic embryos (Cañas et al. 2008). As SE lack the surrounding maternal tissue present around ZE, germination media typically include N supplementation to facilitate germination yield. Llebrés et al. (submitted) recently reported that hybrid white pine SE plants germinated on medium without inorganic N developed functional roots and survived at 50% higher rates than those germinated with inorganic and organic N sources or solely inorganic N. These results suggest a critical role for the N source in the germination of pine somatic embryo plants.

Figure 7 integrates the compartmentation of the GS-GOGAT cycle in conifer cells (Cánovas et al. 2007) and the results presented here for the subcellular localization of enzymes of the arginine metabolic pathway (Figures 3 and 4). In conifers, glutamine biosynthesis occurs in the cytosol and should be transported into the plastid for glutamate biosynthesis. During the maturation stages, increased levels of GS and GOGAT and enhanced levels of glutamine, glutamate, and arginine have been reported in the somatic embryos of white spruce (Joy et al. 1997; Stasolla et al. 2003). In addition, the expression of *GS1b*, a gene encoding cytosolic GS, was induced at the initiation of pine embryo maturation (Pérez-Rodríguez et al. 2006). Glutamine and glutamate are the N donors for the ornithine (Figure 7a) and arginine (Figure 7b) pathways located in the plastid resulting in the biosynthesis of arginine. Arginine is transported outside the plastid for storage protein biosynthesis and deposition into protein vacuoles in the cytosol (Stone and Gifford, 1997). During germination, arginine released from storage protein breakdown or directly transferred from the plastid is catabolized in the mitochondria to generate urea and glutamate (Figure 7b). The hydrolysis of urea by urease has been reported in loblolly pine seedlings (Todd and Gifford, 2002) and the role of δOAT in the biosynthesis of glutamate from ornithine and 2-oxoglutarate has been described (Cañas et al. 2008). The concerted action of urease and GS in the cytosol channels N from arginine to glutamine and asparagine which are the most abundant amino acids in the seedlings (Cánovas et al. 2007; Cañas et al. 2016).

In summary, the arginine metabolic pathway was characterized in maritime pine. The genes involved in the biosynthesis and metabolic utilization of this amino acid were identified; FLcDNAs were isolated and subsequently used to determine the localization of enzymes in different subcellular compartments. The results derived from gene expression profiling strongly suggest that arginine metabolism is deregulated in SE compared with ZE. These changes in N homeostasis were consistent with the lower accumulation of storage proteins observed during the last steps of somatic embryogenesis. Additional studies are needed to identity the regulatory factors involved in the transcriptional regulation of the pathway and unravel

the role of protein sensor PII in arginine metabolism and accumulation of storage proteins. The knowledge acquired from these fundamental studies will help to refine biotechnological procedures for clonal propagation of conifers via somatic embryogenesis.

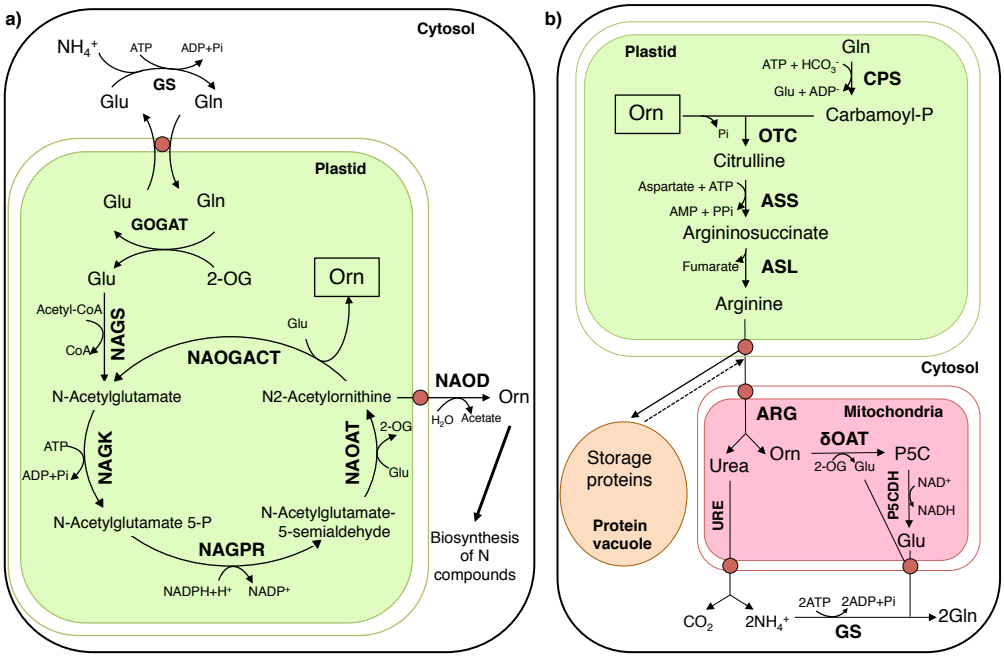


Figure 7. Compartmentation of the GS-GOGAT pathway and arginine metabolism in conifer cells. a) Diagram representing the ornithine biosynthesis pathway in the plastid and the cytosol. GS, glutamine synthetase; GOGAT, glutamate synthase; NAGS, *N*-acetylglutamate synthase; NAGK, *N*-acetylglutamate kinase; NAGPR, *N*-acetylglutamate-5-P reductase; NAOAT, *N*-acetylornithine aminotransferase; NAOGACT, *N*-acetylornithine-glutamate acetyltransferase; NAOD, *N*-acetylornithine deacetylase. b) Diagram representing the plastidic arginine biosynthetic pathway, deposition of storage proteins in the cytosol and arginine mitochondrial degradation. CPS, carbamoyl phosphate synthetase; OTC, ornithine transcarbamoylase; ASS, argininosuccinate synthetase; ASL, argininosuccinate lyase; ARG, arginase; URE, urease; δOAT, ornithine-δ-aminotransferase; P5CDH, P5C dehydrogenase; GS, glutamine synthetase

6. Acknowledgments

This work was supported by grants from the Spanish Ministerio de Economía y Competitividad (BIO2015-69285-R) and Junta de Andalucía (BIO-474). MBP was supported by a postdoctoral contract from Junta de Andalucía (Proyectos de Excelencia Junta de Andalucía, Spain). ZE sampling at FCBA was partially supported by the French Ministry of Agriculture (DGAL, N°2014-352, QuaSeGraine project) and benefited from the support of the XYLOBIOTECH technical facility (ANR-10-EQPX-16 XYLOFOREST). Authors acknowledge Pierre Alazard for cone sampling in maritime pine seed orchard.

7. Supplementary data

Table S1. List of primers used for PCR amplification of genes of the arginine metabolic pathway in maritime pine.

Gene	Fwd_Primer	Rev_Primer
PpNAGS	AAAGCAGGCTTCATGGCTTCAGCCGTCGGA	AGAAAGCTGGGTCGTCAACGATAACCCTA
PpNAGK	AAAGCAGGCTTCATGCGGGCGGCTTCAATG	AGAAAGCTGGGTCACCTGTGATCATCGTTCCAG
PpNAGPR	AAAGCAGGCTTCATGATGGAGTCTACTACC	AGAAAGCTGGGTCAGGAAATAAAGGCTGATG
PpNAOAT	AAAGCAGGCTTCATGGCTTCTTTATACTTT	AGAAAGCTGGGTCTTTTTGTTCATCCAAAGC
PpNAOGACT	AAAGCAGGCTTCATGGATGCTGTAAAAATG	AGAAAGCTGGGTCTGTGGTATACTCTGCA
PpNAOD	AAAGCAGGCTTCATGGCAGCCATCGGAGAA	AGAAAGCTGGGTCTCCTTCAAGTTGTGCAA
PpCPS	AAAGCAGGCTTCATGGCGGCGAAATCATTA	AGAAAGCTGGGTCTCTCTGCTTGTTTTCCC
PpOTC	AAAGCAGGCTTCATGGAAATGGCGGCGTTG	AGAAAGCTGGGTCAAGTCCAAGAGCATGCA
PpASS	AAAGCAGGCTTCATGGCGCATTTGTCGGCA	AGAAAGCTGGGTCTAGTCCACTTTGTAGC
PpASL	AAAGCAGGCTTCATGGAAACAATAGTCACGC	AGAAAGCTGGGTCAGCATTTCCTGATTTG
PpARG	AAAGCAGGCTTCATGGGGTCCATGGGAAAA	AGAAAGCTGGGTCATGAGCCAACTTAGAC
PpδOAT	AAAGCAGGCTTCATGATGCTGCAGAAGGCT	AGAAAGCTGGGTCTTTAATCCGACCGCAAC

Table S2. List of primers used for qPCR analysis of genes of the arginine metabolic pathway in maritime pine.

Gene	Fwd_Primer	Rev_Primer
Actine	ATCTCTCAGCACATTCCAACAG	TATTGCCACCATCATCTCAA
PpNAGS	ACCCGTGGATCAAAGTACTACA	TCGACGCTTTTCAAGTTCCA
PpNAGK	TTCTACGGCCGTCCTTCTAG	CAGAAACTTTCCTGAACTCTGC
PpNAGPR	CCGGTCAAGCACTTCAAAAC	TCGCTCCAGGTCTGTTTACA
PpNAOAT	GCTCCTATGGTTGATGCTGC	ACTGCCTGTTCCAACTCTTCT
PpNAOGACT	GCTCATGGATGGTGGACAAC	TCAGATCACAACCCCATGCT
PpNAOD	GCCACTAAAGATGTCGTCGG	CCCATAGCCTGTTGTTTGCA
PpCPS	GCGGGAAAACAAGCAGAGAT	GCCAGCTTCAATTGCCATTT
PpOTC	GAGCTTTGTTGTGCCATCCA	CTACAAAGCAGCACACACTCT
PpASS	TCCAATCAGAGTTCGAGCCA	GCTGAACTGATGCCTTATTGAC
PpASL	GATCCACTGGCTCTGACTGT	GGCATTCCACTGATCAACAAC
PpARG	CGCGACACAGTTGATGGAAT	TGTCGAACTTCCTCCTGTTACA
PpδOAT	TGAACCATGTGATCGTTGCG	CGGCCCGTCACCTTATACAA

Table S3. Characteristics of proteins differentially accumulated in maritime pine during embryogenesis of somatic and zygotic embryos

Molec. size (kDa)	Description	Sustaine pine Code	Species	Cov %	E value	Identity %	Accession ID	UP*
29.4	Vicilin-like storage protein	unigene42436	*Picea glauca*	89	4,00E-71	32	CAA44873.1	40
	Legumin-like storage protein	unigene42291	*Picea glauca*	95	0.0	75	CAA44874.1	18
24.7	Vicilin-like storage protein	unigene1003	*Picea glauca*		0.0	37	AAB01554.1	33
	Legumin-like storage protein	unigene42291	*Picea glauca*	95	0.0	75	CAA44874.1	19
	Germin-like protein	unigene2649	*Pinus taeda*	99	2,00E-152	92	AAT09853.1	8
17.5	Vicilin-like storage protein	unigene42436	*Picea glauca*	81	2,00E-62	29	CAA44873.1	9
	Legumin-like storage protein	unigene42291	*Picea glauca*	95	0.0	75	CAA44874.1	8
	Late embryogenesis abundant protein	unigene260#	*Picea glauca*	99	9,00E-82	73	ABA54793.1	8

16.7	Vicilin-like storage protein	unigene42436	*Picea glauca*	81	2,00E-62	29	CAA44873.1	17
	Late embryogenesis abundant protein LEA4-1	unigene35634*	*Pinus tabuliformis*	99	2,00E-71	89	AJA33576.1	14
	Late embryogenesis abundant protein	unigene260#	*Picea glauca*	99	9,00E-82	73	ABA54793.1	12
	Late embryogenesis abundant protein LEA1-2	unigene11209#	*Pinus tabuliformis*	99	3,00E-76	97	AJA33568.1	10
	Late embryogenesis abundant protein	unigene11579*	*Picea glauca*	89	2,00E-28	66	AAB01552.1	10
15.8	Vicilin-like storage protein	unigene42436	*Picea glauca*	81	2,00E-62	29	CAA44873.1	20
	Late embryogenesis abundant protein	unigene11579*	*Picea glauca*	89	2,00E-28	66	AAB01552.1	13
	Legumin-like storage protein	unigene42291	*Picea glauca*	95	0.0	75	CAA44874.1	8
	Late embryogenesis abundant protein	unigene37889*	*Picea glauca*	80	2E-102	84	AAA85367.1	7
14.9	Vicilin-like storage protein	unigene42436	*Picea glauca*	89	4,00E-71	32	CAA44873.1	24
	Legumin-like storage protein	unigene42291	*Picea glauca*	95	0.0	75	CAA44874.1	13
	Vicilin-like storage protein	unigene1003	*Picea glauca*		0.0	37	AAB01554.1	10
10.9	Vicilin-like storage protein	unigene42436	*Picea glauca*	81	2,00E-62	29	CAA44873.1	17
	Late embryogenesis abundant protein	unigene48190*	*Picea glauca*		2,9E-17	34	AAA85367.1	17
	Legumin-like storage protein	unigene42291	*Picea glauca*	95	0.0	75	CAA44874.1	7
	Late embryogenesis abundant protein	unigene11579*	*Picea glauca*	89	2,00E-28	66	AAB01552.1	7

*LEA Group 3
#LEA Group 4
UP*: Number of unipeptides detected
Cov: coverage

Table S4. Comparison of enzymes involved in arginine metabolism in *P. pinaster* and *A. thaliana.*

	***A. thaliana* gene ID**	***A. thaliana* protein ID**	***A. thaliana* ORF(aa)**	***P. pinaster* ORF (aa)**	**% identity**
PpNAGS	At2g22910	Q84JF4	609	641	66.3
	At4g37670	B5X4Z4	613	-	66.2
PpNAGK	At3g57560	Q9SCL7	347	352	62.8
PpNAGPR	At2g19940	Q93Z70	401	415	73.8
PpNAOAT	At1g80600	Q9M8M7	457	493	67.3
PpNAOGACT	At2g37500	Q9ZUR7	468	507	70.1
PpNAOD	At4g17830	Q9C5C4	440	431	74.6
PpCPS	At3g27740	Q9LVW7	430	454	74.7
PpOTC	At1g75330	O50039	375	384	70.1
PpASS	At4g24830	Q9SZX3	494	512	78.9
PpASL	At5g10920	Q9LEU8	517	529	71.1
PpARG	At4g08900.1	P46637	342	341	80.2
PpδOAT	At5g46180	Q9FNK4	475	468	68.3

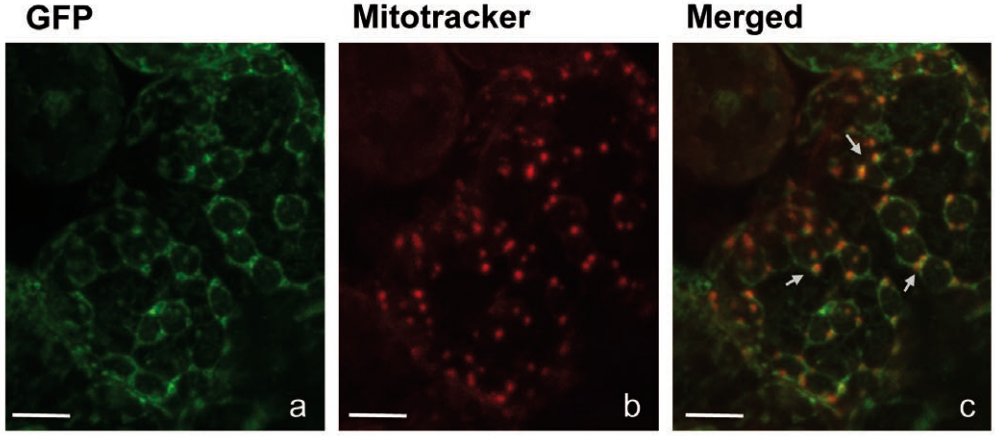


Figure S1. Confocal laser scanning microscopy of GFP-tagged OAT from maritime pine and *N. benthamiana* labelled mitochondria by Mitotracker Red FM. a, green fluorescent protein *(*GFP); b, Mitochondrial marker; c, Overlap between GFP protein and the mitochondrial marker. White arrows indicate co-localized groups of mitochondrias. Scale bar 10 μm.

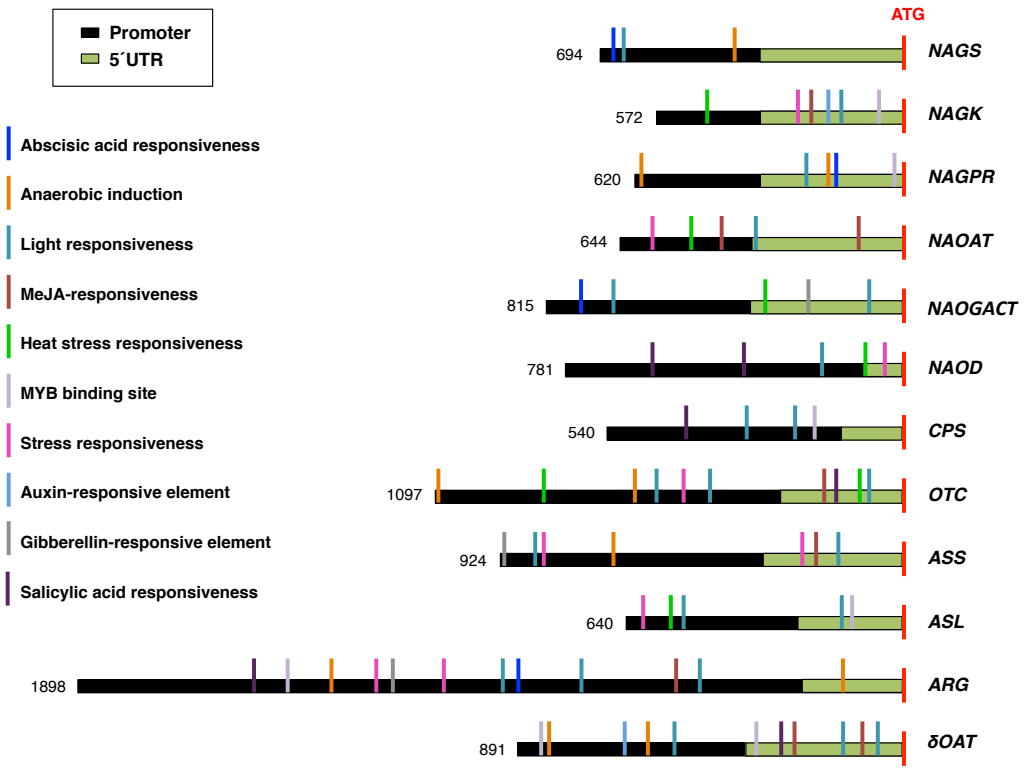


Figure S2. The regulatory regions of genes of the arginine metabolic pathway in maritime pine. Identification of *cis*- acting regulatory elements using the *PlantCARE database* (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

8. References

Avila C, García-Gutiérrez A, Crespillo R, Cánovas FM (1998) Effects of phosphinothricin treatment on glutamine synthetase isoforms in Scots pine seedlings. *Plant Physiol Biochem* 36: 857-863.

Aberlenc-Bertossi F, Chabrillange N, Duval Y, Tregear J (2008) Contrasting globulin and cysteine proteinase gene expression patterns reveal fundamental developmental differences between zygotic and somatic embryos of oil palm. *Tree Physiol* 28: 1157-1167.

Allona I, Collada C, Aragoncillo C (1992) Seed storage proteins from *Pinus pinaster* Ait: Homology of major components with 11S proteins from angiosperms. *Plant Sci* 87: 9-18.

Allona I, Collada C, Aragoncillo C (1994) 2S arginine-rich proteins from *Pinus pinaster* seeds. *Tree Physiol* 14: 211-218.

Amara I, Zaidi I, Masmoudi K, Ludevid MD, Pages M, Goday A, Brini F (2014)

Insights intolate embryogenesis abundant (LEA) proteins in plants: From structure to the functions. *Am J Plant Sci* 5: 3440-3455.

Bouffier L, Raffin A, Alia R (2013) Maritime pine – *Pinus pinaster Ait*. In: Mullin TJ, Lee SJ (eds) Best practice for tree breeding in Europe, 65-75 pp. Skogforsk, Uppsala, ISBN: 978-91-977649-6-4.

Brownfield DL, Todd CD, Stone SL, Deyholos MK, Gifford DJ (2007) Patterns of storage protein and triacylglycerol accumulation during loblolly pine somatic embryo maturation. *Plant Cell Tissue Organ Cult* 88: 217-223.

Businge E, Bygdell J, Wingsle G, Moritz T, Egertsdotter U (2013) The effect of carbohydrates and osmoticum on storage reserve accumulation and germination of Norway spruce somatic embryos. *Physiol Plant* 149: 273-285.

Businge E, Brackmann K, Moritz T, Egertsdotter U (2012) Metabolite profiling reveals clear metabolic changes during somatic embryo development of Norway spruce (*Picea abies*). *Tree Physiol* 32: 232-244.

Canales J, Avila C, Pacheco-Villalobos D, Díaz-Moreno S, Molina-Rueda JJ, Ariza D, Navarro-Cerrillo, Cantón FR, Claros MG, Cánovas FM (2012) Gene expression profiling in the stem of young maritime pine trees: detection of ammonium stress-responsive genes in the apex. *Trees* 26: 609-619.

Canales J, Bautista R, Label P, Gómez-Maldonado J, Lesur I, Fernández-Pozo N, et al. (2014) De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. *Plant Biotechnol J 12*: 286-99.

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

Cañas RA, Villalobos DP, Díaz-Moreno SM, Cánovas FM, Cantón FR (2008) Molecular and functional analyses support a role of ornithine-{delta}-aminotransferase in the provision of glutamate for glutamine biosynthesis during pine germination. *Plant Physiol 148*: 77-88.

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

Cañas RA, Feito I, Fuente-Maqueda F, Avila C, Majada J, Cánovas FM (2015b) Transcriptome-wide analysis supports environmental adaptations of two *Pinus pinaster* populations from contrasting habitats. *BMC Genomics* 16: 909.

Cañas RA, de la Torre F, Pascual MB, Ávila C, Cánovas FM (2016) Nitrogen economy and nitrogen environmental interactions in conifers. *Agronomy* 6: 26.

Cantón FR, Suárez MF, Cánovas FM (2005) Molecular aspects of nitrogen

mobilisation and recycling in trees. *Photosynth Res 83*: 265-278.

Chen YM, Ferrar TS, Lohmeir-Vogel E, Morrice N, Mizuno Y, Berenger B, et al. (2006) The PII signal transduction protein of *Arabidopsis thaliana* forms an arginine-regulated complex with plastid N-acetylglutamate kinase. *J Biol Chem 281*: 5726-5733.

Chellamuthu VR, Ermilova E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, et al. (2014) A widespread glutamine-sensing mechanism in the plant kingdom. *Cell* 159: 1188-1199.

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

Facchini PJ (2001) Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Annu Rev Plant Physiol Plant Mol Biol* 52: 29-66.

Farjon AA (2010) Handbook of the World's Conifers, Brill Academic Publishers, Leiden/Boston.

Ferrario-Méry S, Besin E, Pichon O, Meyer C, Hodges M (2006) The regulatory PII protein controls arginine biosynthesis in *Arabidopsis*. *FEBS Lett* 580: 2015-20.

Fremont N, Riefler M, Stolz A, Schmülling T (2013) The *Arabidopsis TUMOR PRONE5* gene encodes an acetylornithine aminotransferase required for arginine biosynthesis and root meristem maintenance in blue light. *Plant Physiol* 161: 1127-1140.

Furuki T, Shimizu T, Chakrabortee S, Yamakawa K, Hatanaka R, Takahashi T, Kikawada T, Okuda T, Mihara H, Tunnacliffe A, Sakurai M (2012) Effects of Group 3 LEA model peptides on desiccation-induced protein aggregation. *Biochim Biophys Acta* 1824: 891-897.

Goyal K, Walton LJ, Tunnacliffe A (2005) LEA proteins prevent protein aggregation due to water stress. *Biochem J* 388: 151-157.

Hand SC, Menze MA, Toner M, Boswell L, Moore D (2011) LEA proteins during water stress: not just for plants anymore. *Annu Rev Physiol* 73: 115-134.

Joy RW, Yeung EC, Kong L, Thorpe TA (1991) Development of white spruce somatic embryos: I. storage product deposition. *In Vitro Cell Dev Biol Plant* 27: 32-41.

Joy RW, Vogel HJ, Thorpe TA (1997) Inorganic nitrogen metabolism in embryogenic white spruce cultures. A nitrogen 14/15 NMR study. *J Plant Physiol* 151: 306-315.

Kalamaki MS, Alexandrou D, Lazari D, Merkouropoulos G, Fotopoulos V, Pateraki I, et al. (2009) Overexpression of a tomato N-acetyl-L-glutamate synthase gene (*SlNAGS1*) in *Arabidopsis thaliana* results in

high ornithine levels and increased tolerance in salt and drought stresses. *J Exp Bot* 60: 1859-1871.

King JE, Gifford DJ (1997) Amino acid utilization in seeds of loblolly pine during germination and early seedling growth (I. Arginine and arginase activity). *Plant Physiol* 113: 1125-1135.

Klimaszewska K, Park YS, Overton C, MacEacheron I, Bonga JM (2001) Optimized somatic embryogenesis in *Pinus strobus* L. *In Vitro Cell Dev Biol Plant* 37: 392-399.

Klimaszewska K, Morency F, Jones-Overton C, Cooke J (2004a) Accumulation pattern and identification of seed storage proteins in zygotic embryos of *Pinus strobus* and in somatic embryos from different maturation treatments. *Physiol Plant* 121: 682-690.

Klimaszewska K, Rutledge RG, Séguin A (2004b) Genetic transformation of conifers utilizing somatic embryogenesis. In: Peña L (ed) Methods in molecular biology. Totowa, NJ: Humana Press, 151-164.

Klimaszewska K, Hargreaves C, Lelu-Walter M-A, Trontin JF (2016) Advances in conifer somatic embryogenesis since year 2000. In: Germanà MA, Lambardi M (eds) In Vitro Embryogenesis in Higher Plants, Methods in Molecular Biology, vol. 1359, Chapter 7, Springer Science+Business Media New York, pp. 131-166.

Konrádova A, Lipavská H, Albrechtová J, Vreugdenhil D (2002) Sucrose metabolism during somatic and zygotic embryogeneses in Norway spruce: content of soluble saccharides and localisation of key enzyme activities. *J Plant Physiol* 159: 387-396.

Lamy JB, Delzon S, Bouche PS, Alia R, Vendramin GG, Cochard H, Plomion C (2014) Limited genetic variability and phenotypic plasticity detected for cavitation resistance in a Mediterranean pine. *New Phytol* 201: 874-86.

Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2006) Simplified and improved somatic embryogenesis for clonal propagation of *Pinus pinaster* (Ait.). *Plant Cell Rep* 25: 767-776.

Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2008) Clonal plant production from self- and cross-pollinated seed families of *Pinus sylvestris* (L.) through somatic embryogenesis. *Plant Cell Tissue Organ Cult* 92: 31-45.

Lelu-Walter M-A, Klimaszewska K, Miguel C, Aronen T, Hargreaves C, Teyssier C, Trontin JF (2016) Somatic embryogenesis for more effective breeding and deployment of improved varieties in *Pinus* sp.: bottlenecks and recent advances. In: Loyola-Vargas VM, Ochoa-Alejo N (eds) Somatic Embryogenesis Fundamental Aspects and Applications, Springer Verlag, Chap. 19, pp. 319-365. ISBN 978-3-319-33704-3, DOI 10.1007/978-3-319-33705-0-19.

Liu Y, Schiff M, Dinesh-Kumar SP (2002) Virus-induced gene silencing in

tomato. *Plant J* 31: 777-786.

Ljungdahl PO, Daignan-Fornier B (2012) Regulation of amino acid, nucleotide, and phosphate metabolism in *Saccharomyces cerevisiae*. *Genetics* 190: 885-929.

Llacer JL, Fita I, Rubio V (2008) Arginine and nitrogen storage. *Curr Opin Struct Biol 18*: 673-681.

Miguel CM, Rupps A, Raschke J, Rodrigues AS, Trontin JF (2016) Impact of molecular studies on somatic embryogenesis development for implementation in conifer multi-varietal forestry. In: Park Y-S, Bonga JM, Moon H-K (eds), Vegetative Propagation of Forest trees, Korea Forest Research Institute, Seoul, Korea, pp. 373-421. Published online 05/03/2016, http://www.iufro20902.org/. ISBN 978-89-8176-064-9.

Molesini B, Mennella G, Martini F, Francese G, Pandolfini T (2015) Involvement of the putative N-acetylornithine deacetylase from *Arabidopsis thaliana* in flowering and fruit development. *Plant Cell Physiol* 56: 1084-1096.

Morel A, Trontin JF, Corbineau F, Lomenech AM, Beaufour M, Reymond I, Le Metté C, Ader K, Harvengt L, Cadene M et al. (2014a) Cotyledonary somatic embryos of Pinus pinaster Ait. Most closely resemble fresh, maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses. *Planta* 240: 1075-1095.

Morel A, Teyssier C, Trontin JF, Eliášová K, Pešek B, Beaufour M, Morabito D, Boizot N, Le Metté C, Belal-Bessai L, Reymond I, Harvengt L, Cadene M, Corbineau F, Vágner M, Label P, Lelu-Walter MA (2014) Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water availability: transcriptomic and proteomic analyses. *Physiol Plant* 152: 184-201.

Pérez-Rodríguez J, Suárez MF, Heredia R, Ávila C, Breton D, Trontin, JF, Filonova L, Bozhkov P, von Arnold S, Harvength L, Cánovas FM (2006) Expression patterns of two glutamine synthetase genes in zygotic and somatic pine embryos support specific roles in nitrogen metabolism during embryogenesis. *New Phytol* 169: 35-44.

Plomion C, Bartholomé J, Lesur I, Boury C, Rodríguez-Quilón I, Lagraulet H, Ehrenmann F, Bouffier L, Gion JM, Grivet D, de Miguel M, de María N, Cervera MT, Bagnoli F, Isik F, Vendramin GG, González-Martínez SC (2016) High-density SNP assay development for genetic analysis in maritime pine (*Pinus pinaster*). *Mol Ecol Resour* 16: 574-87.

Majumdar R, Shao L, Minocha R, Long S, Minocha SC (2013) Ornithine: The overlooked molecule in the regulation of polyamine metabolism. *Plant Cell Physiol* 54: 990-1004.

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

Bioinformatics 24: 1549-1551.

Stasolla C, van Zyl L, Egertsdotter U, Craig D, Liu W, Sederoff RR (2003) The effects of polyethylene glycol on gene expression of developing white spruce somatic embryos. *Plant Physiol* 131: 49-60.

Slocum RD (2005) Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant Physiol Biochem* 43: 729-745

Stone SL, Gifford DJ (1997) Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early seedling growth. I. Storage protein reserves. *Int J Plant Sci* 158: 727-737.

Uhrig RG, Ng KK, Moorhead GB (2009) PII in higher plants: a modern role for an ancient protein. *Trends Plant Sci* 14: 505-511.

Tereso S, Zoglauer K, Milhinhos A, Miguel C, Oliveira MM (2007) Zygotic and somatic embryo morphogenesis in *Pinus pinaster*: comparative histological and histochemical study. *Tree Physiol* 27: 661-669.

Tiburcio AF, Altabella T, Bitrián M, Alcázar R (2014) The roles of polyamines during the lifespan of plants: from development to stress. *Planta* 240: 1-18.

Todd CD, Cooke JE, Mullen RT, Gifford DJ (2001) Regulation of loblolly pine (*Pinus taeda* L.) arginase in developing seedling tissue during germination and post-germinative growth. *Plant Mol Biol* 45: 555-565.

Todd CD, Gifford DJ (2002) The role of the megagametophyte in maintaining loblolly pine (*Pinus taeda* L.) seedling arginase gene expression in vitro. *Planta* 215: 110-118.

Todd CD, Gifford DJ (2003) Loblolly pine arginase responds to arginine in vitro. *Planta* 217: 610-615.

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

Trontin JF, Teyssier C, Morel A, Harvengt L, Lelu-Walter MA (2016a) Prospects for new variety deployment through somatic embryogenesis in maritime pine. In: Park Y-S, Bonga JM, Moon H-K (eds) Vegetative Propagation of Forest Trees, Korea Forest Research Institute (Seoul, Korea) pp. 572-606. ISBN 978-89-8176-064-9.

Trontin JF, Ávila C, Debille S, Teyssier C, Canlet F, Rueda-López M, Canales J, de la Torre F, El-Azaz J, Pascual B, Cañas R, Boizot N, Le Metté C, Lesage-Descauses MC, Abarca D, Carneros E, Rupps A, Hassani SB, Zoglauer K, Arrillaga I, Mendoza-Poudereux I, Cano M, Segura J, Miguel C, De Vega-Bartol J, Tonelli M, Rodrigues A, Label P, Le Provost G, Plomion C, da Silva Perez D, Harvengt L, Díaz-Sala C, Cánovas FM, Lelu-Walter MA (2016b) Somatic embryogenesis as an enabling technology for reverse genetics: achievements and prospects for breeding maritime pine (*Pinus pinaster* Ait.). Fourth international conference of the IUFRO unit 2.09.02 (somatic embryogenesis and

other vegetative propagation technologies) on “Development and application of vegetative propagation technologies in plantation forestry to cope with a changing climate and environment”, September 19-23 (La Plata, Argentina), pp. 57.

Trontin JF, Klimaszewska K, Morel A, Hargreaves C, Lelu-Walter M-A (2016c) Molecular aspects of conifer zygotic and somatic embryo development: a review of genome-wide approaches and recent insights. In: Germanà MA, Lambardi M (eds) In Vitro Embryogenesis in Higher Plants, Methods in Molecular Biology, vol. 1359, Chapter 8, DOI 10.1007/978-1-4939-3061-6_8, Springer Science+Business Media New York, pp. 167-207.

Vega-Bartol JJ, Simões M, Lorenz WW, Rodrigues AS, Alba R, Dean JFD, Miguel C (2013) Transcriptomic analysis highlights epigenetic and transcriptional regulation during zygotic embryo development of *Pinus pinaster*. *BMC Plant Biol* 13: 123.

Winter G, Todd CD, Trovato M, Forlani G, Funck D (2015) Physiological implications of arginine metabolism in plants. *Front Plant Sci* 6: 534.

Xia J, Yamaji N, Che J, Shen RF, Ma JF (2014) Normal root elongation requires arginine produced by argininosuccinate lyase in rice. *Plant J* 78: 215-226.

ARTÍCULO 4

Structural and functional characteristics of two molecular variants of the nitrogen sensor PII in maritime pine

Manuscrito en preparación (2018)

1. Abstract

High levels of nitrogen are stored as arginine during the last stages of seed formation in maritime pine (*Pinus pinaster* Aiton). The protein sensor PII regulates the feedback inhibition of arginine biosynthesis through interaction with the key enzyme N-acetylglutamate kinase (NAGK). In the present study, the structural and functional characteristics of PpPII have been investigated to get insights into the regulation of arginine metabolism in maritime pine. Two different forms of PII have been identified, PpPIIa and PpPIIb, which differs in their amino acid sequence and most likely corresponds to splicing variants of a single gene in the pine genome. The two PII variants are present in several pine species but were not found in other conifers such as spruce. PpPIIa and PpPIIb are trimeric proteins exhibiting similar tridimensional structures and located in the chloroplasts, where the PII-target enzyme PpNAGK is present. PpPIIa, PpPIIb and PpNAGK have been recombinantly produced and the formation of NAGK-PII complexes investigated. The interaction of PpPIIb and PpNAGK is regulated by glutamine relieving the feedback inhibition of PpNAGK activity by arginine. In contrast, glutamine has a limited effect antagonizing the inhibition by arginine of the PpNAGK-PpPIIa complex. The potential roles of PpPIIa and PpPIIb in the arginine metabolism of pine are discussed.

2. Introduction

Plants, like other organisms, have developed mechanisms that allow them to receive information and respond to changes in carbon and nitrogen levels. These mechanisms regulate the expression of genes and proteins activity involved in the transport and metabolism of nitrogenous and carbon compounds, allowing the plants to optimize the use of their energy sources. This process is particularly important in trees that have to cope with long life cycles.

To sense carbon and nitrogen there are important mechanisms that allow plants to regulate its metabolism and development in response to nutrient status. The mechanisms of sensitivity to carbon/nitrogen balance make the plant to activate genes involved in the assimilation of nitrogen when carbon skeletons are abundant and internal levels of organic nitrogen are low, or on the contrary, decrease the rate of assimilation of nitrogen when the levels of photosynthetic products are low, or internal levels of organic nitrogen are relatively high. All these mechanisms of sensitivity and response to carbon and nitrogen allow plants to use or maintain the energy required to carry out the carbon and nitrogen metabolism as well as the transport of the metabolites involved (Coruzzi and Zhou 2001).

The systems to sense and respond to carbon and nitrogen levels have been the subject of elegant studies on microorganisms such *E. coli* and yeasts for decades (Forchhammer and Lüddecke 2016). In these unicellular models, signalling systems have been found to be complex and multiple (Jiang et al. 1998). Similar studies on plants are more recent and it can be said beforehand that signalling systems in plants are more complex. Plants must capture and respond to carbon and nitrogen metabolites in different tissues and stages of development and modulate this response to ensure adequate adaptation to the changing conditions of the environment (Ferrario-Mery et al. 2006; Ferrario-Mery et al. 2008).

In bacteria, the signalling PII protein serve as central processing units for integrating signals that are antagonistic to carbon and nitrogen status, and use that information to control nitrogen uptake (Ninfa and Atkinson, 2000). In enterobacteria, where PII function has been best studied, it acts as a direct and indirect sensor of various stimuli that cause the formation of different conformations in the protein, each with a regulatory significance on the expression and activity of the enzyme glutamine synthetase (GS) (Arcondeguy et al. 2001). The assimilation of ammonium as glutamine is regulated in bacteria in response to intracellular concentrations of glutamine (a nitrogen signal) and 2-oxoglutarate (2-OG, a carbon signal) (Senior, 1975). Under these conditions, the precise control of GS activity, responsible for the

assimilation of ammonium, will limit the rate of nitrogen assimilation as a whole to maintain the balance with carbon assimilation.

Regarding the structural characteristics of PII, the bacterial protein has a very well conserved primary structure with two conserved domains called domain I and II, which are found in two functionally important regions called loop B and loop T. In *E. coli*, the protein is a homotrimer with the T loop in the domain I of each monomer. This is the urydilation site (Ninfa and atkinson, 2000). The loop B is formed between residues 81-92 of the *E. coli* protein taking part of domain II, which is the most conserved domain. In loop B is located the ATP binding site and also participates in the 2-OG binding.

In plants, the PII protein is a sensor of the C/N status mediated by the binding to 2-OG and glutamine and subsequent regulation of the arginine metabolic pathway (Chen et al. 2006; Chellamuthu et al. 2014). It is well known that arginine biosynthesis in many organisms is feedback regulated by the end product of the pathway, arginine, through allosteric inhibition of the key enzyme N-acetylglutamate kinase (NAGK). However, when N is abundant, this inhibition is relieved through the interaction of NAGK with the N sensor protein PII (Chen et al. 2006; Llacer et al. 2007). Increased levels of glutamine are a signal of N abundance that is sensed by interaction with an specific binding site in a C-terminal segment of the structure of the PII protein that define another important domain, the so-called Q-loop (Chellamuthu et al. 2014).

Arginine is highly abundant in the seed storage proteins of conifers and represents an important source of nitrogen for developing plantlets following seed germination (Todd et al. 2001; Todd and Gifford 2003; Cánovas et al. 2007). The regulation of arginine metabolism has a relevant role in the deposition and mobilization of storage proteins during the embryogenesis and germination of maritime pine (Llebrés et al. 2017). In the present study, the structural and functional characteristics of the protein sensor PII have been investigated in order to get further insights into the regulation of arginine metabolism in pine.

3. Material and methods

3.1. Plant material

Pinus pinaster Aiton seeds were provided by the Centro Nacional de Recursos Genéticos Forestales from "Ministerio de Agricultura, Pesca, Alimentación y Medio Ambiente", Spain). Seeds were imbibed in distilled water for 24 h under continuous aeration and germinated and grown with vermiculite as a substrate under a photoperiod of 16 h light/ 8 h dark at 24 °C. Hypocotyls of pine seedlings were collected after 2 weeks, frozen in liquid nitrogen and stored at −80 °C until use.

Nicotiana benthamiana L. seeds were sown and grown in pots and maintained under a 16 h light/8 h dark photoperiod at 24 °C for 5 weeks.

3.2. Cloning of PpPIIa, PpPIIb and PpNAGK

The sequences of *PpPIIa, PpPIIb* and *PpNAGK* were retrieved from the maritime pine (*Pinus pinaster*) transcriptome available at SustainPineDB v.3.0 (http://www.scbi.uma.es/sustainpinedb/sessions/new). Full-length cDNAs were isolated from pine seedling hypocotyl RNAs by reverse transcription-PCR (RT-PCR) using primers designed from *P. pinaster* sequences. Extraction of RNA was performed as described by Canales et al. (2012) and quantified using a NanoDrop© ND-1000 spectrophotometer. Synthesis of cDNA was performed with 5X iScript™ cDNA Synthesis Kit (Bio-Rad). The primer pairs used for specific amplification were as follows: 5´-AAA AAGCAGGCTTAATGGCTGCCCATCTGCCTC-3´ as a forward primer for *PpPIIa*; 5´-AAAAAGCAGGCTTAATGCCTGCCCATACGCTTT-3´ as a forward primer for *PpPIIb*; and 5´-GAAAGCTGGGTCTTAGTTGCTGTCTGTAACC-3´ as a reverse primer for *PpPIIa* and *PpPIIb*; 5´-AAAGCAGGCTTCATGCGG GCGGCTTCAATG-3´ as a forward primer and 5´-GAAAGCTGGGTCTCA ACCTGTGATCATCGTTC-3´ as a reverse primer for *PpNAGK*. The resulting PCR products were cloned in a pJet1.2 (ThermoFisher Scientific™) vector and completely sequenced.

3.3. Transient expression of GFP fused proteins in *Nicotiana benthamiana*

Forward primers used for PCR amplification were described previously and stop codon were removed from reverse primer sequences; 5´-AGAAAG CTGGGTCGTTGCTGTCTGTAACCTC-3´ for *PpPIIa* and *PpPIIb*, and 5´-AGA AAGCTGGGTCACCTGTGATCATCGTTCCAG-3´ for *PpNAGK*. The resulting PCR products were cloned into the pDONR207 and subcloned into the pGWB5 vector via Gateway Technology (Invitrogen) to produce full-length proteins fused to GFP at their C-termini under the control of the CaMV *35S* promoter according to El-Azaz et al. (2016). Empty plasmid pGWB5 was used as a negative control. The *Agrobacterium tumefaciens* strain C58C1 was transformed by electroporation with recombinant plasmids. *N. benthamiana* leaves (5 weeks old) were syringe infiltrated with cultures containing pGWB5 constructs mixed with cultures containing P19, both with an optical density at 600 nm of 0.5, according to the procedures described previously (Liu et al. 2002). Subcellular localization of proteins was examined by confocal microscopy 36 to 48 hours after agroinfiltration.

3.4. Transient expression of GFP fused proteins in pine protoplasts

To perform transient expression assays in pine protoplasts, fusions of PpPIIa with and without the transit peptide and the GFP reporter gene were constructed using the appropriate primers. Protoplasts were prepared from maritime pine cotyledons by incubation of 1g of fresh cutting tissue in 10 mL of a mixture containing 0.44% (w/v) K3 medium: 0.4% (w/v) cellulase, 0.4% (w/v) macerase (Calbiochem) and 0.4% sucrose. Transformation with the gene constructs was performed by electroporation essentially as previously described by Gómez-Maldonado *et al*. (2004). After culture in dark for 18-24 hours the transformed protoplasts were visualized by laser confocal microscopy using a Leica CLSM microscope. Excitation was conducted with a laser beam at 488 nm. Red autofluorescence of chlorophyll was detected up to 560 nm and green fluorescence of GFP was detected between 505-520 nm.

3.5. Overexpression and purification of recombinant proteins

Open reading frames (ORFs) corresponding to *PpPIIa* and *PpPIIb* minus the region encoding the chloroplast transit peptide were PCR-amplified using primers containing the appropriate restriction sites. Forward primers used were 5´-CGCATATGCAGGCACCAAATGGGACC-3´ for *PpPIIa* and 5´-CGCATATGCAGGCACCAAATGCAACC-3´ for *PpPIIb* both containing *NdeI* restriction site; and 5´-GGCTCGAGGTTGCTGTCTGTAA CCTC-3´ as reverse primer containing *XhoI* restriction site. PCR products were digested and subcloned into the pET30b at *Nde*I and *Xho*I 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 by shaking in Luria Bertani broth containing 50 μg mL^{-1} kanamycin until OD_{600} = 0.6; 0.2% arabinose (w/v) was added, and cultures were further incubated for 5 hours at 30º C with shaking. Cells were pelleted by centrifugation at 4500 *g* and frozen. ORF corresponding to *PpNAGK* minus the region encoding the chloroplast transit peptide was amplified using 5´-AAAAAGCAGGCTTAAGGAAATCCAGAG GCGCCC-3´ as forward primer. The resulting PCR product was cloned into the pDONR207, subcloned into the pDest17 vector via Gateway Technology (Invitrogen) and transformed into *Escherichia coli* strain BL21-AI (ThermoFisher Scientific). Bacterial growth and protein induction requirements were the same as described above but the media contained 100 μg mL^{-1} ampicillin instead. Proteins were purified by binding onto Ni-agarose resin (Protino NiNTA; Macherey-Nagel). Protein concentrations were determined by the Bradford dye-binding method (Bradford, 1976) and analysed by SDS-PAGE to verify purification.

3.6. PpNAGK activity assays

Activity assays with purified recombinant proteins were performed according to the method described previously by Heinrich et al. (2004). The reaction mixture consisted of 400 mM NH_2OH-HCl, pH 7, 20 mM Tris/Cl, pH 7, 20 mM $MgCl_2$, 40 mM NAG, and 10 mM ATP. All assays were done in triplicate in two or more separate experiments and each reaction mixture

contained 0.65 μg of PpNAGK enzyme. Reactions were initiated by the addition of reaction mixture, and the incubation was carried out at 37 °C in a volume of 150 μL and terminated by the addition of 150 μL of stop mixture (1:1:1 of 5% (w/v) FeCl3·6 H_2O in 0.2 M HCl, 8% (w/v) trichloroacetic acid, and 0.3 M HCl). After standing for 5 min at room temperature, the tubes were centrifuged and the absorbance was immediately measured at 540 nm. The assay was linear throughout the 30-min period. The blank reactions did not contain N-acetyl glutamate. One unit of enzyme activity is defined as one mmol of product produced in 1 min calculated with a molar absorption coefficient of 456 $M^{-1}cm^{-1}$ at 540 nm for the N-acetylglutamylhydroxamate-Fe^{+3} complex.

To study the effect of PpPII on the PpNAGK activity, various amounts of PpPII were mixed with PpNAGK, and the mixture was placed at 4 °C for 10 min to allow complex formation. The reaction was initiated by the addition of the reaction mixture into the PpNAGK-PpPII complex mixture. For studying the PpNAGK activity in the presence of PpPII and its effector 2-OG, PpPII was incubated with 5 mM ATP and various amounts of 2-OG for 3 min at 4 °C followed by the addition of PpNAGK. The inhibitory effect of arginine on PpNAGK was studied adding various amounts of arginine into the PpNAGK solution, and the mixture was allowed to stand for 5 min before initiating the reaction. To test the effect of arginine on the PpNAGK-PpPII complex, PpPII was first added into the diluted PpNAGK solution, and the two proteins were placed at 4 °C for 10 min to allow complex formation followed by the addition of arginine. The enzyme mixture was allowed to stand for 5 min before the initiation of assays. The arginine effect was tested also in combination with 10 mM glutamine as described above.

3.7. Gel filtration chromatography of purified PII proteins

Protein samples were diluted into 25 mM Tris/Cl, pH 7.5, 150 mM NaCl, 10% (v/v) glycerol buffer to a final concentration of 0.6 μg/ μl, and a 250 μl sample was then chromatographed on a ENrich™ SEC 650 10 x 300 Column (BioRad) equilibrated at room temperature in Tris/HCl buffer. The fast

protein liquid chromatography flow rate was 0.75 ml/min, and 100 μl fractions were collected.

The column was previously calibrated from a standard plot of Kav versus molecular mass for ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), ß-amylase (200 kDa), apoferritin (440 kDa) and blue dextran (void).

3.8. Western blot and antibody production

The presence of proteins in the elution fractions was detected by SDS-PAGE and western blot analysis as described previously (Cánovas *et al*. 1991). Antibodies against PpPIIa were raised in rabbits by immunization with PpPIIa recombinant protein overexpressed in *E. coli* as described by Cantón et al. (1996). The primary antibodies anti-PpPIIa were used at 1:10000 dilution and a 1:10000 dilution of horseradish peroxidase conjugated antirabbit serum (Sigma) was used as a secondary antibody.

4. Results

Recent developments in next generation sequencing have facilitated the generation of a range of genomic resources for maritime pine (Canales et al. 2014; Cañas et al. 2017). The *in silico* analysis of the maritime pine transcriptome revealed the existence of two different transcripts encoding the protein sensor PII that were named *PpPIIa* and *PpPIIb* (Figure 1). A striking characteristic of these transcripts is that their sequences differed in the 5'-unstranslated region and the open reading frame (ORF) but were identical in the 3'-unstranslated region (Figure 1A).

The ORFs of *PpPIIa* and *PpPIIb* encoded polypeptides of similar molecular sizes, differing only in the lack of three aminoacids in PpPIIb compared to PpPIIa. Both contain a transit peptide for plastid targeting which hydrolysis produce matures polypeptides of identical size (Table 1). Maritime pine PpPIIa and PpPIIb proteins are larger than the single PII counterpart in Arabidopsis (Table 1).

Figure 1. PpPIIa and PpPIIb maritime pine sequences. A) Primary structure of full-length cDNAs. B) Predicted protein sequences indicating main domains and Q-Loop positions. Black arrow indicates mature polypeptides. Red asterisk indicate main aminoacid change in the position 18 between mature proteins.

Table 1. Characteristics of PpPIIa and PpPIIb polypeptides and processed proteins. cTP: Chloroplast transit peptide.

	PpPIIa	PpPIIa-cTP	PpPIIb	PpPIIb-cTP
Molecular size (Da)	25,467	16,579	25,009	16,482
Aminoacids	237	154	234	154
Isoelectric point	9.38	5.42	8.91	5.64

Although PpPIIa and PpPIIb mature polypeptides showed identical size they differs in their aminoacid residues. Most changes were conservative, but interestingly, a substantial change of a glutamate for a lysine in the position 18 of the mature protein, was observed (Figure 2A).

As only a single form of PII has been previously described in angiosperms, the occurrence of PII variants was investigated in conifers for those species that transcriptomic data are available. In addition to *P. pinaster*, several pine species examined, including *P. taeda*, *P. contorta*, and *P. lambertiana,* contain both PII variants meanwhile only one form of PII was found in spruce species. Furthermore, their amino acid sequences were highly conserved in pine including the non-conservative change of the residue in the position 18. The phylogenetic analysis of PII protein variants clearly classified the pine proteins in two separate clusters being the PIIb cluster closer to PII proteins from other conifer species whose represent an intermediate branch between pines and herbaceous plants (Figure 2B).

To further understand the biological roles of these proteins their subcellular localization was determined. The ORF of PpPIIa and PpPIIb were

PCR-amplified using specific primers and cloned into the gateway vector pGWB5. GFP fusions were transiently expressed in *N. benthamiana* leaves via agroinfiltration to analyse the subcellular localization of the expression products. Transiently expressed constructs for PpPIIa and PpPIIb exhibited GFP fluorescence associated exclusively to chloroplast organelles, which were visualized by the red autofluorescence of chlorophyll (Figure 3A).

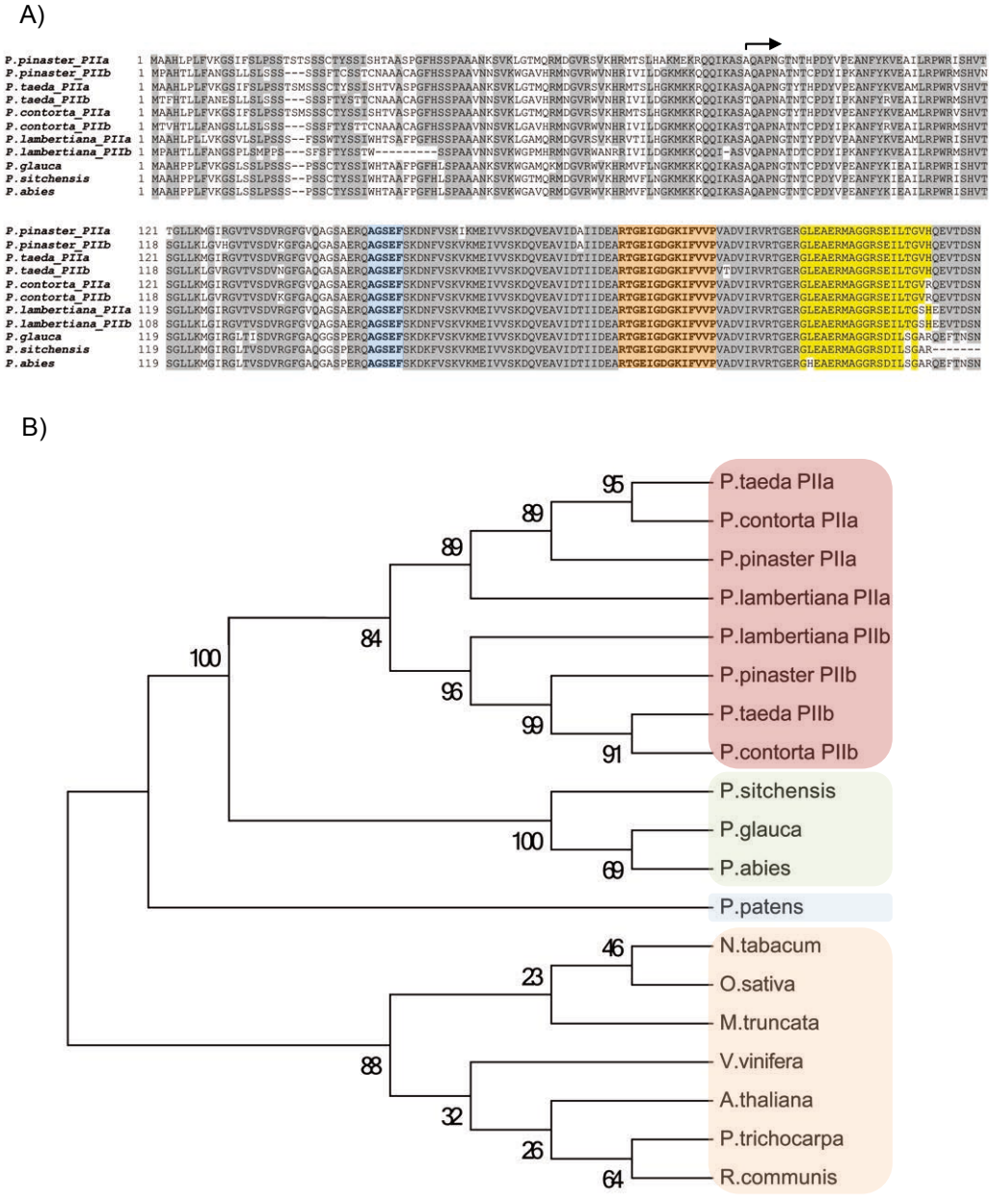


Figure 2. Comparative analysis of PII proteins in other conifers.

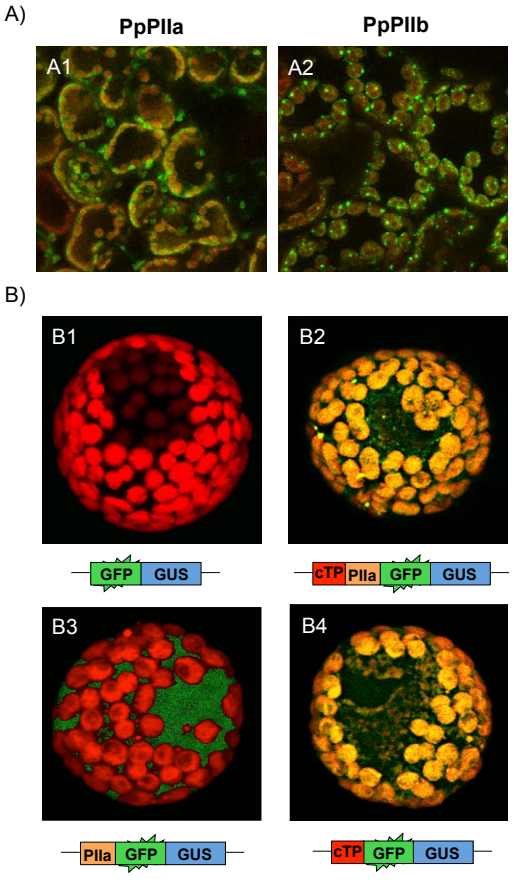


Figure 3. Subcellular localization of maritime pine PpPIIa and PpPIIb. A) Localization in *N. benthamiana* leaves. A1: GFP-Chlorophyll merged images of PpPIIa chloroplast localization; A2: GFP-Chlorophyll merged images of PpPIIb chloroplast localization. **B)** Localization in pine protoplasts. B1: GFP-control construct; B2: GFP-PpPIIa construct containing the transit peptide; B3: GFP-PpPIIa construct without the transit peptide, B4: GFP-control construct without the PpPIIa protein but containing the transit peptide.

The merged images clearly showed co-localization of chlorophyll with both GFP-PpPIIa and GFP-PpPIIb in the chloroplasts visualized as yellow colour, confirming therefore plastid localization. To further confirm the chloroplast localization of PpPIIa protein, GFP fusions were constructed with and without the transit peptide and were subsequently electroporated into isolated maritime pine protoplasts. Figure 3B shows that GFP-PpPIIa constructs containing the transit peptide were appropriately targeted to the

chloroplasts as also observed for control constructs without the PpPIIa protein but containing the transit peptide. However, for the constructs lacking the transit peptide the GFP signals were visualized exclusively in the cytosol confirming indisputably the plastid localization of PpPIIa.

To further understand the structural and functional properties of PpPIIa and PpPIIb the three-dimensional structure of both proteins was modelled using the x-ray structure of PII from *Arabidopsis thaliana* (Mizuno et al. 2007). The resulting structure form the model of PpPIIa is similar to the one described for other PII proteins in plants (Chellamuthu et al. 2014; Forchhammer and Lüddecke, 2016) and contains a well-defined T-loop involved in the binding of ATP and 2-oxoglutarate (Figure 4A). In addition, the three-dimensional structure for PpPIIa contains a Q-loop defined by the C-terminal extension that is absent in the *Brassicaceae* family but characteristic of plant PII proteins when compared to their bacterial counterparts (Figure 4A). This Q-loop domain is stabilized by the binding of glutamine making contact with the basal part of the T loop and transmitting the glutamine signal through a conformational change in the protein that favours the interaction with NAGK (Forchhammer and Lüddecke, 2016). A comparison of the threedimensional structures of PpPIIa and PpPIIb shows that they are almost identical proteins (Figure 4B).

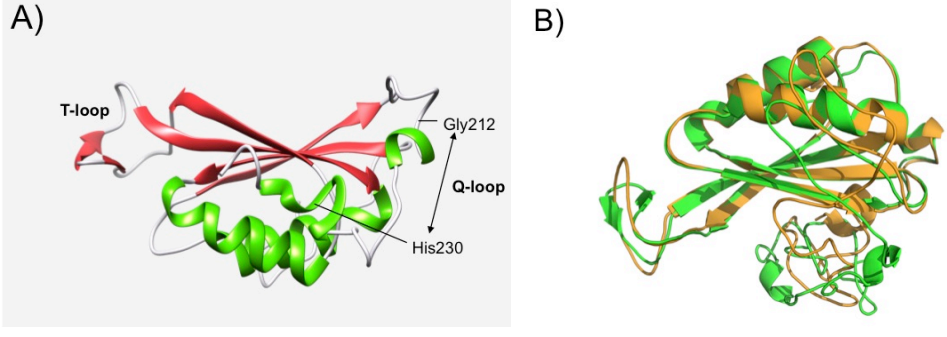


Figure 4. Molecular modelling of maritime pine PpPIIa and PpPIIa. A) Three-dimensional model of PpPIIa protein. B) Overlapping models of PpPIIa and PpPIIb.

The PII variants from maritime pine were recombinantly expressed in *E. coli* to raise monospecific antibodies and to further characterize their molecular characteristics. Furthermore, sufficient amounts of recombinant native proteins were necessary to perform functional studies on the interaction of PpPIIa and PpPIIb with pine N-acetylglutamate kinase (PpNAGK) and to investigate how this interaction is regulated by the availability of glutamine and 2-OG.

First, the native molecular sizes of PpPIIa and PpPIIb were determined by gel filtration chromatography through a FPLC calibrated column with proteins standards (Figure 5). The size of both holoproteins was estimated to be 57 kDa, and considering the molecular size of the corresponding polypeptides revealed by western blotting, the results are compatible with a trimeric structure of the native proteins.

The target enzyme of maritime pine PpPII proteins in arginine metabolism is PpNAGK which is encoded by a full-length cDNA of 1824 bp that contains an ORF for a predicted full-length protein of 352 amino acids (Llebrés et al. 2017). The N-terminus of the predicted polypeptide targets the protein to the plastid where the mature enzyme has been recently localized (Llebrés et al. 2017). To examine the potential differences in the interaction of PpPIIa and PpPIIb with their target enzyme, a full-length cDNA for PpNAGK was amplified using appropriate primers. A construct containing the mature PpNAGK enzyme (301 residues) and lacking the predicted transit peptide was fused to a His-tag and overexpressed in *E. coli.* The resulting recombinant protein was purified by affinity chromatography and used for interaction studies. Figure 6 shows that saturation curves of PpNAGK activation were substantially different for PpPIIa and PpPIIb. Lower amounts of PpPIIb (about 10 µg) than PpPIIa (more than 20 µg) were necessary to reach similar saturation levels of PpNAGK activity. Subsequent experiments were addressed to study whether PpNAGK activity at increasing levels of 2-OG could be affected by the interaction of both PpPIIa and PpPIIb with PpNAGK. As shown in Figure 7 the addition of similar amounts of PpPIIa and PpPIIb enhanced the PpNAGK activity levels at varied concentrations of 2-OG.

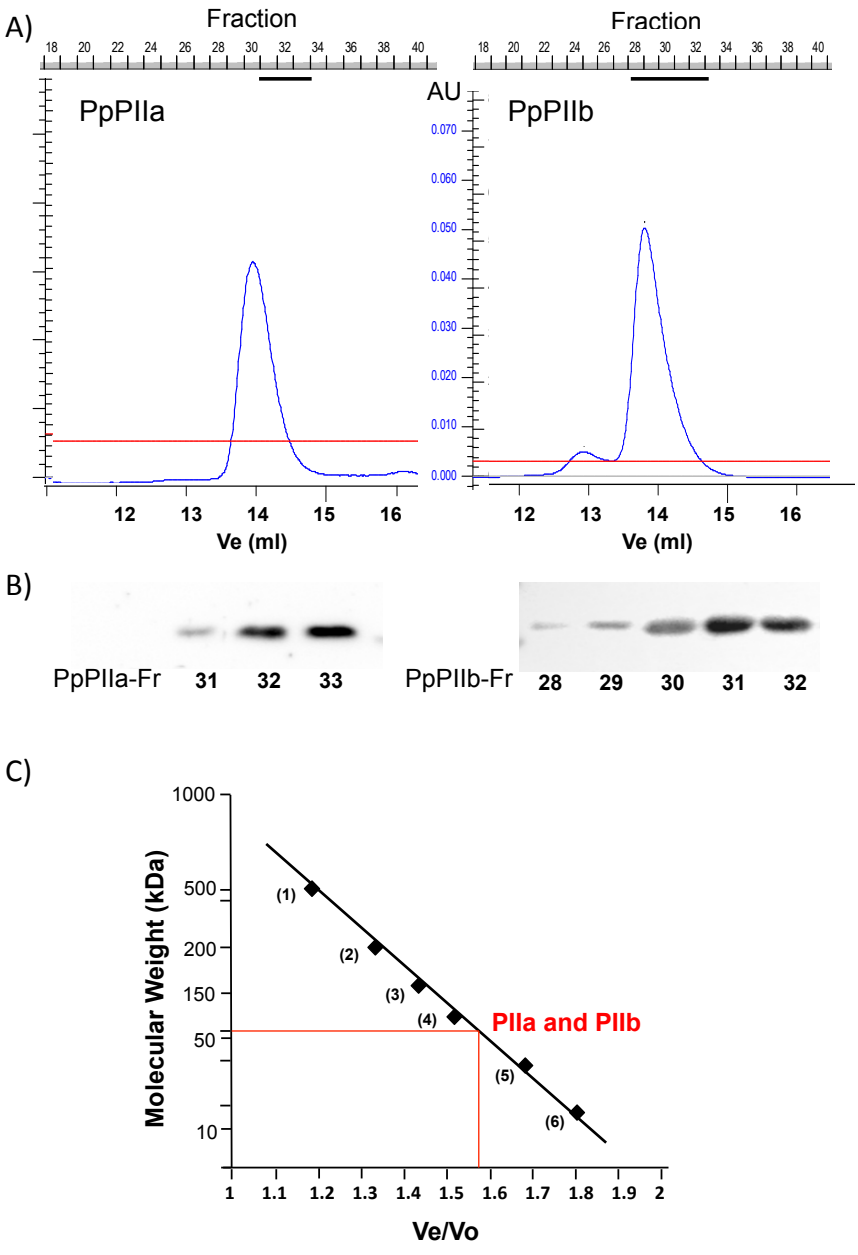


Figure 5. Estimation of native PpPII proteins molecular size by gel filtration chromatography. A) Elution profile of recombinant PpPIIa and PpPIIb proteins using gel filtration in an FPLC system. B) Fractions containing proteins were tested by western blotting using monospecific antibodies against PpPIIa. C) Calculation of PpPIIa and PpPIIb molecular size using protein markers of known molecular size.

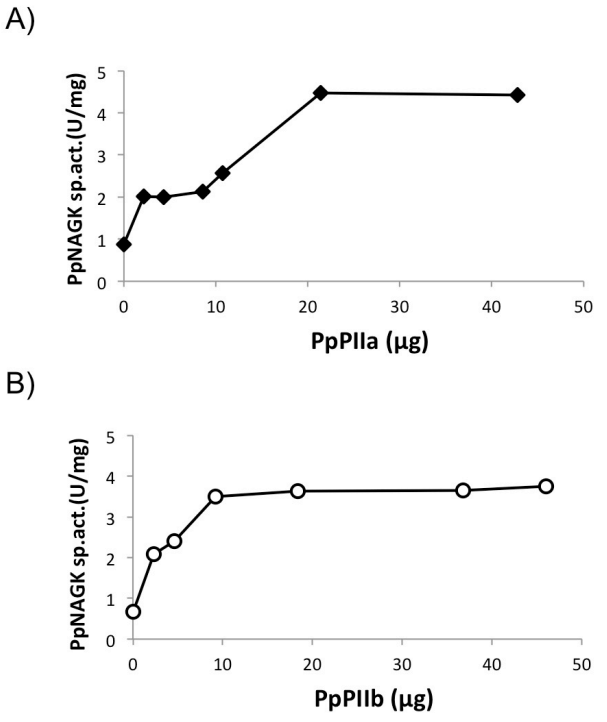


Figure 6. Effects of PpPIIa and PpPIIb on the PpNAGK specific activity. A) Effect of increasing amount of PpPIIa protein on PpNAGK activity (black squares). B) Effect of increasing amount of PpPIIb protein on PpNAGK activity (open circles).

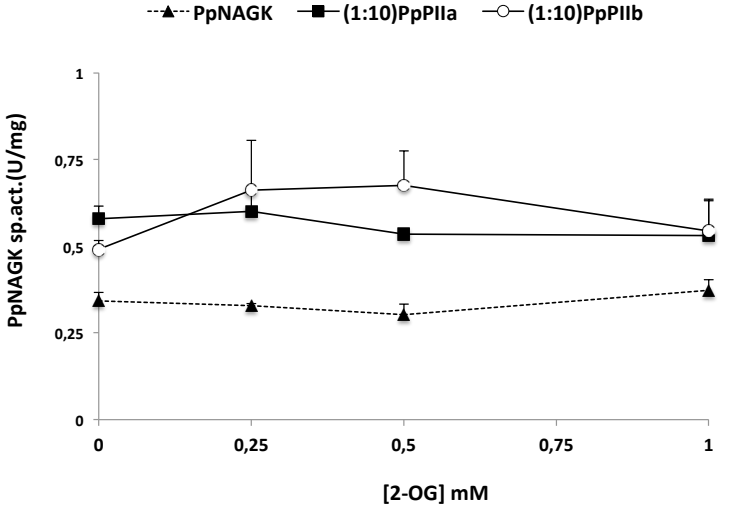


Figure 7. Effect of 2-OG increasing concentration in PpNAGK activity. In the presence of 2-OG (black triangles) without PII protein, or PpNAGK-PpPIIa complex activity at ratio (1:10) (black squares) or PpNAGK-PpPIIb complex activity at ratio (1:10) (open circles).

To further characterize the sensor properties of PpPIIa and PpPIIb the effect of glutamine in the interaction with PpNAGK was studied. The addition of 10 mM glutamine did not affected in a great extent the inhibition of PpNAGK activity observed at increasing levels of arginine, the end product of the arginine metabolic pathway (Figure 8A). The addition of recombinant PpPIIa protein had a limited effect antagonizing the feedback regulation of PpNAGK by arginine. Only a peak point activation of PpNAGK was observed at 0.25 mM arginine and similar levels of activity were observed at increased levels of arginine independently whether glutamine was present or not (Figure 8B). In contrast, when the recombinant PpPIIb protein was added to PpNAGK, the availability of glutamine had a positive effect in the relief of the arginine inhibition of NAGK catalytic activity (Figure 8C). Figure 8C also shows that the antagonism effect of glutamine in the inhibition by arginine was particularly effective at low concentrations of arginine between 0.25 and 3 mM.

5. Discussion

A bioinformatics survey of the maritime pine transcriptome revealed the occurrence of two transcripts (*PpPIIa and PpPIIb*) encoding PII proteins (Figure 1). These transcripts contain ORF encoding PII polypeptides of similar molecular size but differing in their amino acid sequence. These findings are striking because only one form of PII has been reported to exist in other plant species (Chen et al. 2006). Interestingly, two PII homologues have also been identified in several pine species for which large transcriptomics data are available. In contrast, only one PII gene product has been found in spruce species (Figure 2). Taking together, the results found in the present study suggest that the existence of PII variants is not a general characteristic of conifers and appears to be restricted to pines.

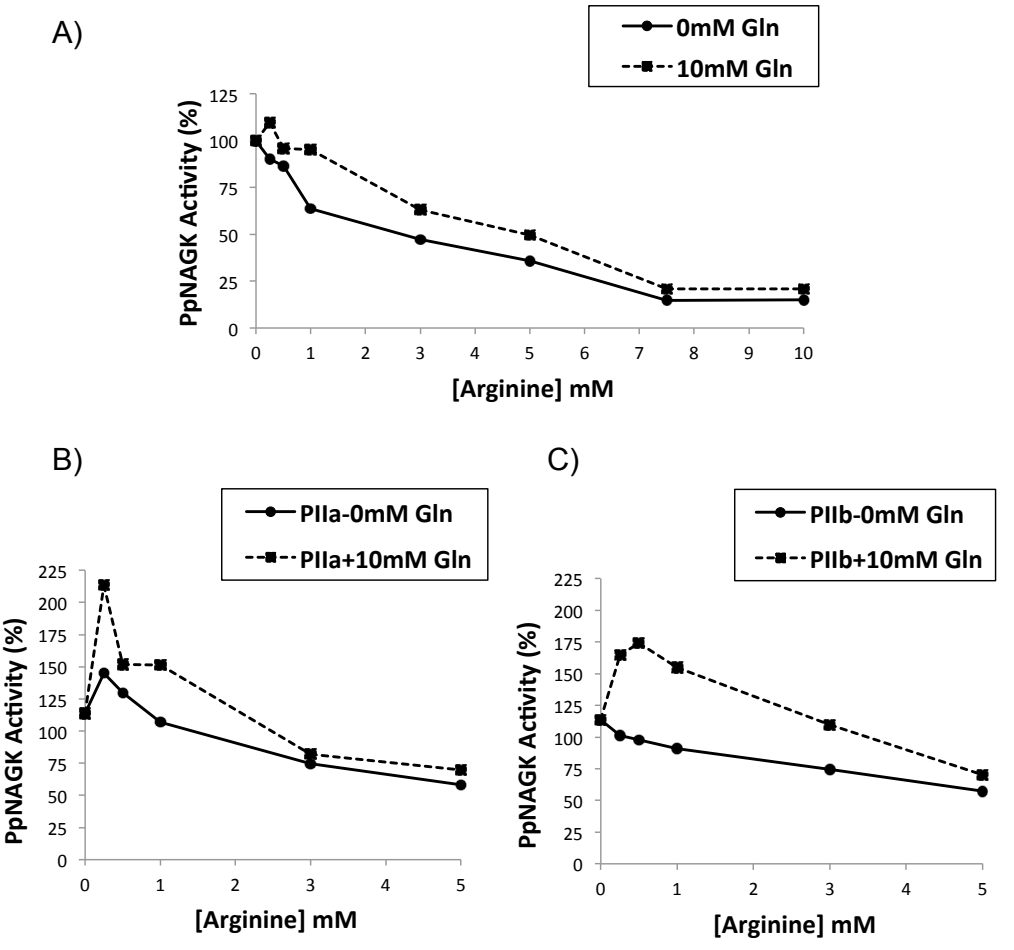


Figure 8. Effect of arginine in PpNAGK activity. A) PpNAGK activity inhibition by arginine with or without glutamine. B) PpNAGK-PpPIIa complex activity inhibition by arginine with or without glutamine. C) PpNAGK-PpPIIb complex activity inhibition by arginine with or without glutamine.

As the sequence of the maritime pine genome is still not available, homologues of PII were searched in the loblolly pine (*Pinus taeda*) genome (Neale et al. 2014), to figure out whether these transcripts were the products of different genes, or the result of a differential splicing of primary transcripts derived from the same gene. Sequences for *PtPIIa* and *PtPIIb* were found to be located in the same scaffold but unfortunately were interrupted by large

gaps that prevented to retrieve a complete gene structure from the *P. taeda* genome draft.

Considering the primary structure of their cDNAs the most plausible explanation is that *PtPIIa* and *PtPIIb* have been generated by alternative splicing of a single gene. No different 3'-untranslated regions were amplified from a range of different pine tissues by rapid amplification of cDNA ends (RACE). Furthermore, the above conclusion is reinforced by the identification of single hybridizing bands in a Southern blot of pine genomic DNA digested with several restriction enzymes (Supplemental Figure 1).

The sequence analysis of the polypeptides encoded by *PtPIIa* and *PtPIIb* predicted a transit peptide for plastid targeting in the N-terminal end. Nevertheless, the existence of two forms of PII could be related with a distinct subcellular localization in pine cells. To investigate this possibility subcellular localization studies were conducted and the results shown in Figure 3 confirmed that PpPIIa and PpPIIb proteins are indisputably located in the chloroplast stroma as previously reported for the PII protein in *Arabidopsis* Chen et al. (2006). Consequently, these data suggest that both PpPIIa and PpPIIb are able to physically interact *in vivo* with the target enzyme, PpNAGK, which is also located in the same cellular compartment of pine cells (Llebrés et al. 2017).

The modelling of PpPIIa and PpPIIb revealed that they are almost identical proteins and their three-dimensional structures are highly similar to the previously described crystal structures of PII proteins from *Synechococcus* and *Arabidopsis* (Mizuno et al 2007; Llacer et al. 2008). However, a conserved C-terminal extension sequence have been identified in almost all PII proteins in plants with the exception of members of the *Brassicaceae* family (Chellamuthu et al. 2014). This C-terminal segment is able to bind glutamine and is not found in the cyanobacteria and *Arabidopsis* counterparts (Chellamuthu et al. 2014). The C-terminal sequence of the plant PII is disordered in the absence of glutamine but when the amino acid is present, the characteristic Q-loop in the three-dimensional structure is stabilized making contact with the T-loop adopting the appropriate conformation that is required for interaction with NAGK (Chellamuthu et al.

2014; Forchhammer and Lubbeck, 2016). The change of a glutamate for a lysine in position 18 is close to the T-loop in the three-dimensional structure of the PpPII proteins and therefore could affect the contact with the glutamine-stabilized Q-loop. Accordingly, the single change of a glutamate in PpPIIa for a lysine in PpPIIb suggest that these N sensor proteins in maritime pine could have a differential response to glutamine availability, the interaction with NAGK could be modulated in a different way and consequently the regulation of arginine biosynthesis. Additional studies are necessary to model the glutamine binding to PpPIIa and PpPIIb and determine whether the change of charge at position 18 of the mature protein has an effect in the contact of T-loop and Q-loop.

The glutamine-binding site exhibits low affinity for this amino acid inducing the stabilization of the Q-loop at physiological concentrations in the range of millimolar, which are present in plant cells. The ability to sense glutamine levels, and therefore the cellular N status, was acquired by a C-terminal extension of the PII polypeptide that is exclusively present in the plant proteins (Chellamuthu et al. 2014). It has been recently suggested that the glutamine-binding site appeared when the gene encoding PII (*glnB*) of an ancestral plant was transferred from the chloroplast genome to the nuclear genome (Forchhammer and Lubbeck, 2016). In fact, a number of genes involved in N metabolism including those encoding PII and Fd-GOGAT are encoded in the plastid genome of *Porphyra* red algae.

The overexpression in *E. coli* of recombinant PpPIIa and PpPIIb allowed to determine that these two variants are trimeric proteins (Figure 5) as previously described for other PII proteins form plants and microorganisms (Chen et al. 2006; Forchhammer and Lubbeck, 2016). Moreover, the overproduction in *E. coli* of PpNAGK, the target enzyme of PII, enables to study the formation of PpNAGK-PpPIIa and PpNAGK-PpPIIb complexes and how glutamine and 2-OG affect the formation of such regulatory complexes.

The binding of both PpPIIa and PpPIIb to PpNAGK increased enzyme activity as previously described in *Arabidopsis* Chen et al. (2006). However, maximal levels of enzyme activity were reached with different amounts of PpPIIa and PpPIIb recombinant proteins (Figures 6A and 6B). These results

strongly suggest a differential affinity of PpPIIa and PpPIIb for PpNAGK in the formation of the regulatory complexes.

The addition of increasing amounts of 2-OG did not affect PpNAGK activity when the enzyme was incubated in the absence of PpPIIa and PpPIIb. Equally, the addition of 2-OG did not affect substantially the enhancement of PpNAGK activity observed in the presence either of PpPIIa or PpPIIb. These results suggest that 2-OG has a limited effect in the activation of PpNAGK.

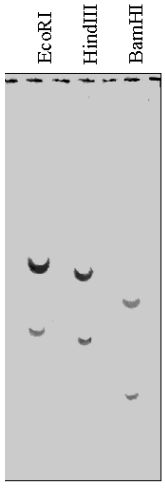
It is well documented that the more sensitive assay for the NAGK-PII interaction is the relief of the observed inhibition of the enzyme when increasing amounts of arginine are present (Llacer et al. 2008). As observed for 2-OG, the addition of increased amounts of glutamine did not affect PpNAGK activity in a great extent (Figure 8A). Chellamuthu et al. 2014 reported that glutamine is able to further relief the feed-back inhibition of NAGK by arginine when the enzyme is bound to PII. Similarly, enhanced PpNAGK activity was observed in the presence of PpPIIb when glutamine was present (Figure 8B). In contrast, the same activation effect was not observed in the presence of PpPIIa, except at low concentration of arginine. These results suggest that PpPIIb will be more effective than PpPIIa in relieving the inhibition of PpNAGK by arginine. This assumption is consistent with the superior activation of PpNAGK activity observed in the presence of PpPIIb. Moreover, this effect was particularly important at concentrations of arginine in the range of 0.5 to 3 mM, which are in the range of those found in the chloroplasts when growing in conditions of high N availability. Therefore, PpPIIa will require other conditions for sensing N availability non-yet identified. One speculate that PpPIIa may have higher affinity for glutamine than PpPIIb and it would be able to sense low to intermediate concentrations of nitrogen. Whether the change of a glutamate for a lysine in position 18 could be involved in the differential response of PpPIIa and PpPIIb to glutamine availability remains to be determined.

In summary, two different forms of PII have been identified in maritime pine named PpPIIa and PpPIIb. These PII variants are most likely the splicing products of a single gene in the pine genome and their existence appears to be restricted to pines because a single PII was identified in other conifer

species and in angiosperms. PpPIIa and PpPIIb are trimeric proteins of similar three-dimensional structures and located in the chloroplasts where the PII-target enzyme PpNAGK is also located. The interaction of recombinants PpPIIb and PpNAGK is regulated by glutamine relieving the feedback inhibition of PpNAGK activity by arginine. In contrast, glutamine had a limited effect antagonizing the inhibition by arginine of the PpNAGK-PpPIIa complex.

6. Supplemental material

Figure S1 - Southern blot of genomic DNA from *P. sylvestris* digested with the restriction enzymes *Eco*RI, *Hind*III y *Bam*H and hybridized with a PpPIIa cDNA probe.



7. References

Arcondeguy T, Jack R, Merrick M (2001). PII signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiology and Molecular Biology Reviews* 65: 80-105.

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.

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

Canales J, Avila C, Cantón FR, Pacheco-Villalobos D, Díaz-Moreno S et al. (2012). Gene expression profiling in the stem of Young maritime pine tres: detection of ammonium stress-responsive genes in the ápex. *Trees – Structure and Function* 26: 609-619.

Cañas RA, Li Z, Pascual MB, Castro-Rodríguez V, Ávila C, Sterck L, Van de Peer Y, Cánovas FM (2017). The gene expression landscape of pine seedling tissue. *The Plant Journal* 91: 1064-1087.

Cánovas FM, Cantón FR, Gallardo F, García-Gutiérrez A, de Vicente A (1991). Accumulation of glutamine synthetase during early development of maritime pine seedlings (*Pinus pinaster*). *Planta* 185: 372-378.

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

Cantón F.R, García-Gutiérrez A, Crespillo R and Cánovas FM. (1996) High-level expression of *Pinus sylvestris* glutamine synthetase in *Escherichia coli*. Production of polyclonal antibodies against the recombinant protein and expression studies in pine seedlings. *FEBS Letters* 393: 205-210.

Chellamuthu VR, Ermilova E, Lapina T, Luddecke J, Minaeva E, Herrmann C, Hartmann MD, Forchhammer K (2014) A widespread glutamine sensing mechanism in the plant kingdom. *Cell* 159: 1188–1199.

Chen YM, Ferrar TS, Lohmeier-Vogel EM, Morrice N,cMizuno Y, Berenger B, Ng KKS, Muench DG, Moorhead GBG (2006) The PII signal transduction protein of Arabidopsis thaliana forms an arginine regulated complex with plastid N-acetyl glutamate kinase. *Journal of Biological Chemistry* 281: 5726-5733.

Coruzzi GM, Zhou L (2001). Carbon and nitrogen sensing and signaling in plants: emerging `matrix effects´. *Current Opinion in Plant Biology* 4: 247-253.

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

Ferrario-Méry S, Besin E, Pichon O, Meyer C, Hodges M (2006). The regulatory PII protein controls arginine biosynthesis in *Arabidopsis. FEBS Letters* 580: 2015-2020.

Ferrario-Méry S, Meyer C, Hodges M (2008). Chloroplast nitrite uptake is enhanced in *Arabidopsis* PII mutants. *FEBS Letters* 582: 1061-1066.

Forchhammer K, Lüddecke J (2016). Sensory properties of the PII signalling protein family. *FEBS Journal* 283: 425-437.

Gómez-Maldonado J, Cánovas FM, Avila C (2004). Molecular analysis of the 5-upstream region of a gibberellin-inducible cytosolic glutamine synthetase gene (GS1b) expressed in pine vascular tissue. *Planta* 218: 1036-1045.

Heinrich A, Maheswaran M, Ruppert U, Forchhammer K (2004). The *Synechococcus elongatus* P_{II} signal transduction protein controls arginine synthesis by complex formation with *N*-acetyl-l-glutamate kinase. *Molecular Microbiology* 52: 1303-1314.

Jiang P, Pelinska JA, Ninfa AJ (1998). Reconstitution of the signal-transduction bicyclic cascade responsible for the regulation of Ntr gene transcription in *Escherichia coli*. *Biochemistry* 37: 12795-12801.

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

Llácer JL, Contreras A, Forchhammer K, Marco-Marín C, Gil-Ortiz F, Maldonado R, Fita I, Rubio V (2007). The crystal structure of the complex of PII and acetylglutamate kinase revelas how PII controls the storage of nitrogen as arginine. *Proceedings of the National Academy of Sciences USA* 104: 17644-17649.

Llacer JL, Fita I, Rubio V (2008). Arginine and nitrogen storage. *Current Opinion Structural Biology* 18: 673–681.

Llebrés MT, Pascual MB, Debille S, Trontin JF, Avila C, Cánovas FM (2017) The role of arginine metabolic pathway during embryogenesis and germination in maritime pine (*Pinus pinaster* Ait.). *Tree Physiol*ogy 3:1-14. doi: 10.1093/treephys/tpx133

Mizuno Y, Moorhead GB and Ng KK (2007). Structural basis for the regulation of N-acetylglutamate kinase by PII in *Arabidopsis thaliana*. *Journal of Biological Chemistry* 282: 35733–35740.

Neale DB, Wegrzyn JL, Stevens KA et al. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology* 15: R59. Available at: https://doi.org/10.1186/gb-2014-15-3-r59.

Ninfa AJ and Atkinson MR (2000). PII signal transduction proteins. *Trends in Microbiology* 8: 172-179

Senior PJ (1975). Regulation of nitrogen metabolism in *Escherichia coli* and *Klebsiella aerogenes*: studies with continuous-culture technique. *Journal of Bacteriol*ogy 123: 407-418.

Todd CD, Cooke JE, Mullen RT, Gifford DJ (2001). Regulation of loblolly pine (*Pinus taeda* L.) arginase in developing seedling tissue during germination and post-germinative growth. *Plant Molecular Biology* 45*:* 555-565.

Todd CD, Gifford DJ (2003). Loblolly pine arginase responds to arginine in vitro. *Planta* 217: 610-615.

ARTÍCULO 5

PpAAP1, a novel amino acid permease involved in arginine uptake in Pinus pinaster

Manuscrito en preparación (2018)

1. Abstract

Plants have access to a numerous nitrogenous compounds in the soil such as ammonium, nitrate and amino acids. A series of transporters that facilitate the uptake of nitrate and ammonium have been identified and characterized in plants. Recent molecular work, using cloning and functional complementation in yeast, has revealed also the existence of numerous transporters in plants with affinity by different amino acids.

Based on the knowledge that amino acid uptake from the soil is dependent on transporters, and to understand plant amino acid uptake, it is critical the identification and characterization of the membrane proteins involved. To study amino acid transport in *Pinus pinaster* at a molecular level, we have identified 10 cDNAs encoding putative amino acid permeases of the AAP family. One member of the family, PpAAP1, exhibited enhanced expression levels in pine roots when arginine was externally supplied in hydroponic media. *PpAAP1* encodes a predicted polypeptide of 54 kDa highly hydrophobic with 9 putative membrane-spanning regions. *PpAAP1* was functionally characterized by complementation of a yeast mutant defective in the transport of some amino acids such as arginine. Complementation with PpAAP1 allowed yeast mutants to take up L- [^{14}C] arginine efficiently with a Km of 83 μM and Vmax of 5 nmol min^{-1} 10^{8} cells^{-1}.

To further study the function of PpAAP1, it was investigated its ability to rescue the phenotype of *Arabidopsis thaliana* AAP5 T-DNA mutants that were severely affected in the uptake of arginine. AAP5 mutants showed enhanced root biomass when grown under 1mM of arginine compared with wild-type plants. In contrast, PpAAP1 transgenic plants exhibited a dramatic reduction in root biomass when growing at such high arginine concentration suggesting that an efficient transport of this amino acid could be the cause of the observed phenotypical change.

2. Introduction

Plant nitrogen uptake is a key process for plant growth and productivity. Nitrogen (N) is available to plants in the soil, as inorganic forms, such as ammonium and nitrate, and as organic compounds such as amino acids and peptides (Öhlund et al. 2004). Due to its importance, plants have evolved different ways and mechanisms to acquire, transport and store N in different organs and tissues (Xu et al. 2012).

Plant species differ in their preferred N source, absorbing it by the roots primarily in inorganic form. Conifers from boreal forests have a marked preference for ammonium over nitrate due to their natural distribution in ammonium rich soils (Kronzucker et al. 1996; Cañas et al. 2016). In some forest ecosystems, organic N in the form of amino acids and peptides is also quantitatively important and represent a major source of N for tree nutrition (Lipson and Näsholm, 2001; Näsholm et al. 2009). Amino acid uptake rates are highly dependent on substrate concentration (Wright, 1962). The natural concentrations of amino acids in soils vary considerably between ecosystems, and are usually found in the lower μmolar range (0.01-10 μM) (Raab et al. 1996; Öhlund, 2004). Furthermore, although amino acid concentrations are low, their high turnover rate suggests that they are an important N source in some ecosystems (Kielland 1995; Jones & Kielland 2002).

The first evidence that amino acids are taken up by plants started to appear at the beginning of 20th century (Hutchinson & Miller, 1912). Since then, molecular cloning and functional complementation in yeast have revealed that plants express a multitude of different amino acid transporters. Amino acid transporters in plants can be classified into five gene families: amino acid transporter family (ATF), amino acid-polyamine-choline transporter superfamily (APC), mitochondrial carrier family (MCF), preprotein and amino acid transport (OEP16 or PRAT^{d}) and divalent anion: Na^{+} symporter (DASS). In *Arabidopsis thaliana*, *Populus trichocarpa* and *Oryza sativa,* at least 100, 134 and 96 genes respectively were annotated as putative amino acid transporters (Pratelli et al. 2014; Tuskan et al. 2006; Yuan et al. 2005; Ouyang et al. 2007).

In *Arabidopsis* the amino acid transporter family (ATF), also called the amino acid/auxin permease family (AAAP) is the largest family, consisting of 46 members. The ATFs can be divided in six subfamilies, being the amino acid permease family (AAP) the best characterized. AAP family was discovered after the identification of the *Arabidopsis* AAP1 using yeast complementation (Frommer et al. 1993; Hsu et al. 1993) and since then seven additional AAPs have been discovered in this plant species. When examined in heterologous expression systems, AAPs generally display affinity for neutral and acidic amino acids; the exceptions are AAP3 and AAP5, which transport all classes of amino acids, but have the greatest affinity to the cationic amino acids (Fischer et al. 1995; Boorer & Fischer, 1997; Fischer et al., 2002; Okumoto et al. 2004). Screening of *A. thaliana* T-DNA mutants, showed that only the AAP5 mutant displayed a clear phenotypic divergence on high levels of arginine in the growth media, and the root uptake of basic amino acids was also strongly affected in AAP5 mutant lines (Svennerstam et al. 2008).

Organic nitrogen and particularly arginine is a relevant source of external nitrogen for conifer nutrition (Gruffman et al. 2014). However, in spite of recent advances in the characterization of plants AAPs little is known about the AAPs family members in conifers. In this study, 10 cDNA sequences of putative AAPs have been detected using the available information of *P. pinaster* transcriptome (Canales et al. 2014). The deduced proteins exhibit a high sequence similarity to their previously described counterparts in *Arabidopsis, Oryza* and *Populus*. In this research, a *P. pinaster* amino acid permease (*PpAAP1*) has been identified as directly involved in arginine uptake by the roots. We have functionally characterized*PpAAP1* in yeast to determine its kinetics and ability for arginine uptake. In addition *Arabidopsis* transgenic plants overexpressing *PpAAP1* were generated to further study its function during development at both high and low arginine levels.

3. Material and methods

3.1. Plant material, transformation and growth conditions

Seeds of *Pinus pinaster Ait.* were grown in vermiculite in a plant growth chamber with a 16 h light photoperiod and constant temperature of 24°C for two weeks. Seedlings of 2 cm cotyledons were harvested and individually transferred to a 2 L pot with a hydroponic nutritive solution containing macronutrients and micronutrients for optimal growth as described previously in Canales et al. (2010) without nitrogen source as control group. The hydroponic solution was supplemented with 2.25 mM of arginine in an additional group. The pH of all nutrient solutions was adjusted to 5.8 using hydrochloric acid and changed every day to avoid contamination. Root samples were harvested after 2 hours, 48 hours and 5 days after the first irrigation with the nutrient solution, frozen in liquid nitrogen and stored at −80 °C until further analysis.

Full-length *PpAAPs* cDNA sequences were obtained in the SustainPine Database v.3.0 (Canales et al. 2014). PCR amplification was carried out with 100 ng of *P. pinaster* cDNA from cotyledons using *PpAAP1* specific primers; forward 5′-AAAGCAGGCTTCATGGATCGGGAAAACGC TCA-3′ and reverse 5′-GAAAGCTGGGTCTCAGTAAGCGGTATGGAAGG-3′. PCR product was recombined into pDONR207 (Invitrogen, Germany) and cloned into the destination vector pGWB2, which contains a hygromycin selectable marker gene and the CaMV 35S promoter to drive the expression of the transgene. The construct was verified by sequencing.

Arabidopsis thaliana amino acid transporter AAP5 mutants (SALK_041999, *aap5-1*) and (SALK_099586, *aap5-2*) from the SALK institute (Nottingham Arabidopsis Stock Centre, Nottingham, UK) were grown in soil in a controlled chamber at 24°C and 16h light photoperiod for 8 weeks. Homozygous mutant plants were transformed by the floral dipping method (Clough and Bent, 1998) with *A. tumefaciens* strain C58C1 carrying the construct containing the *PpAAP1* cDNA into gateway vector pGWB2. Seeds were collected and screened separately on half-strength Murashige and

Skoog (MS) medium with 20 µg mL^{-1} Hygromycin B to ensure that lines originating from individual transformation events were isolated.

For arginine growth experiments, *Arabidopsis* transgenic seeds were germinated on sterile agar plates containing N free half-strength Murashige and Skoog (MS) medium (Murashige & Skoog, 1962), with 0.65% w/v agar (plant agar, Duchefa Biochemie, Haarlem, the Netherlands), 0.5% w/v sucrose, 3 mM nitrate and several concentrations of arginine (10 µM, 50 µM, 100 µM or 1mM). The media was buffered to pH 5.8 with 3.6 mM MES. Plants were grown at 24°C and 16h light photoperiod for 15 days.

Wild-type *Arabidopsis thaliana* Columbia ecotype plants were used as control in all experiments described.

3.2 RNA isolation and qRT-PCR analysis

RNA extraction was performed as described in Canales et al. (2012). Samples were ground in liquid nitrogen and 100 mg of each were extracted with 650 µL extraction buffer containing 3 % (*w*/*v*) cetyltrimethylammonium bromide (CTAB), 100 mM Tris pH 8.0, 2 M NaCl, 2 % (*w*/v) PVP-40 and 30 mM ethylenediaminetetraacetic acid (EDTA). Residual genomic DNA was removed by a treatment with RQ1 RNase-Free DNase (Promega, Madison, WI, USA) and then purified using Nucleospin® Gel and PCR Clean-Up (Macherey-Nagel, Dürem, Germany). Total RNA was quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, EEUU). Reverse transcription was carried out with 500 ng of total RNA in a 10 µL reaction volume with iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA) under the following manufacturer conditions. Real-time PCR was carried out on a CFX384 Real-Time System C1000 Thermal Cycler (Bio-Rad, Hercules, CA, USA) using RT-qPCR SsoFast™ EvaGreen® Supermix (Bio-Rad, Hercules, CA, USA) under the following conditions: 95 °C for 2 min (1 cycle), 95 °C for 1 s and 60 °C for 5 s (40 cycles). Ten nanograms of reverse transcribed cDNA were used as the template for each reaction, primer pair concentration was 0.5 µM and final reaction volume was 10 µL. Three biological and three technical replicates were performed per sample type.

Elongation factor alpha 1 and *Actin* were used as reference genes. Gene specific primer sequences used are listed in Table S1.

3.3. Yeast growth and transformation

The yeast strain *22Δ8AA* was a mutant lacking multiple amino acid uptake systems, (Mat*α* gap1-1, put4-1, uga4-1, Δcan1, Δapl1, Δlyp1, Δhip1, Δdip5, ura3-1; Fischer et al. 2002). The coding sequence corresponding to *PpAAP1* was subcloned into the pDR196 yeast expression vector. Strain *22Δ8AA* was transformed with the resulting plasmid or with the empty pDR196 vector as negative control. Transformation was carried out by heat shock according to Dohmen et al. (1991). Colonies carrying recombinant plasmids were screened using a selective uracil-free medium. Growth assays were performed in N-free media containing 3 mM aspartic acid, glutamic acid, proline or arginine, as sole nitrogen source.

Transformants were grown in liquid N-free minimal medium. At a final OD_{600} of 2, the cultures were washed with sterile water and serially diluted (1,1/10,1/100,1/1000,1/10000). Then, 10 µL of each dilution was spotted on N-free media containing 0.1 mM, 0.2 mM or 0.5 mM Arginine. Plates were incubated at 30 °C and images were taken after 3 days.

3.4. Transport measurement of L- [^{14}C]-arginine

Yeast cells were grown till logarithmic phase, harvested at OD_{600} of 0.6, washed, and resuspended in ice-cold buffer (50 mM KH_2PO_4, pH 5.0), to a final OD_{600} of 5. Before the uptake measurements an aliquot of yeast cells was supplemented with 20 mM glucose, and incubated at 30º C for 5 min. Uptake was initiated by the addition of 100 µL of radioactive substrate mixture, containing L-[^{14}C]-arginine at several concentrations ranged between 2.5 µM to 150 µM. Uptake rates were determined over 4 min. Samples were removed after 30s, 1 min, 2 min and 4 min, transferred to 5 mL of 100 mM unlabelled arginine ice-cold buffer, filtered on glass fibre filters (Whatman), and washed twice with 4 mL of buffer. Radioactivity was determined by liquid scintillation spectrometry. Transport measurements were performed in triplicate.

3.5. Sequence analysis of *P. pinaster* AAP family

Full-length amino acid sequences of plants AAPs were retrieved using PLAZA 3.0 database (http://bioinformatics.psb.ugent.be/plaza/). Protein alignment was performed with the ClustalW2 package. For phylogenetic analyses, the alignments were imported into the Molecular Evolutionary Genetics Analyses software (MEGA) version 6.06, Neighbour-joining (NJ) method was applied with the Poisson correction model, the pairwise deletion option and bootstrap test with 1000 replicates (Courturier et al. 2010). Transmembrane domains were predicted by Protter v1.0 program THMM (Omasits et al. 2014).

Transcriptome sequences and localization of transcript expression were available at ConGenIE.org (http://v22.popgenie.org/microdisection/).

4. Results

4.1. *In silico* analysis of *P. pinaster* AAP family

A total of 10 members of AAP family were identified in the *P. pinaster* transcriptome (Canales et al. 2014) by BLAST search of known sequences from *Arabidopsis*. These membrane proteins were further confirmed, using Protter transmembrane prediction program (THMM), by the identification of 9 to 11 transmembrane domains in each sequence (Table 1). From them, PpAAP1, PpAAP2, PpAAP3, PpAAP4, PpAAP5 and PpAAP6 presented a N-terminal of intracellular localization and a C-terminal of extracellular localization. By contrast, PpAAP7, PpAAP8, PpAAP9 and PpAAP10 presented N and C terminals of extracellular localization (Figure S1).

The cDNAs identified in *P. pinaster* encode proteins with similar number of amino acid residues, from 466 to 500, and with a molecular mass ranging 51 to 58.5 kDa (Table 1). These results are quite similar to those described for the AAPs family members characterized in other plants such as *O. sativa*, *P. trichocarpa* and *A. thaliana* (Yuan et al. 2005; Tuskan et al. 2006; Rentsch et al. 2007; Ouyang et al. 2007; Taylor et al. 2015).

Table 1. Genes of *P. pinaster* AAP family.

Name	Gene ID	FL cDNA (bp)	Protein (aa)	MW (Da)	Transmembrane domains
PpAAP1	isotig28092	2267	489	54,012	9
PpAAP2	isotig32920	1791	492	54,688	9
PpAAP3	isotig44810	2040	497	54.349	9
PpAAP4	isotig44984	2009	500	54,643	9
PpAAP5	unigene165	1995	499	54,543	11
PpAAP6	unigene2046	2195	533	58,550	11
PpAAP7	unigene41903	2083	482	52,835	10
PpAAP8	isotig44509	2103	483	53,526	10
PpAAP9	isotig22404	2412	466	51,331	10
PpAAP10	isotig43366	1797	466	51,530	10

A phylogenetic tree, based on a multiple protein sequence alignment of 48 AAPs from *P. pinaster* (10), *O. sativa* (16), *P. trichocarpa* (14) and *A. thaliana* (8), revealed that plant AAPs are grouped into four main subgroups as previously described by Couturier et al. (2010) (Figure 1). In the subgroup-I, together with AtAAP2, AtAAP3, AtAAP4 and AtAAP5 4 putative *P. pinaster* AAPs (PpAAP1, PpAAP2, PpAAP3 and PpAAP4) were included. These proteins present high similarity between them, from 60 to 69 %. The subgroup-II included, three PpAAP proteins (PpAAP5, PpAAP6 and PpAAP7) together with AtAAP1, AtAAP6 and AtAAP8 of *Arabidopsis*, sharing identities of 52-57%. As previously described by Okumoto et al. (2002), AtAAP7 was the most distal member of the *Arabidopsis* AAP family and grouped together with three PpAAP proteins included in subgroup-IV (PpAAP8, PpAAP9 and PpAAP10), with 50-52% identity. On the other hand, only *O. sativa* proteins are present in subgroup-III. The percentages of identities between *P. pinaster* and *A. thaliana* AAPs are shown in Table 2. The closest *P. pinaster* homologs of AtAAP5 are PpAAP1, PpAAP2, PpAAP3 and PpAAP4, which share 60 to 68% identity at protein level and could be involved in the amino acids root uptake.

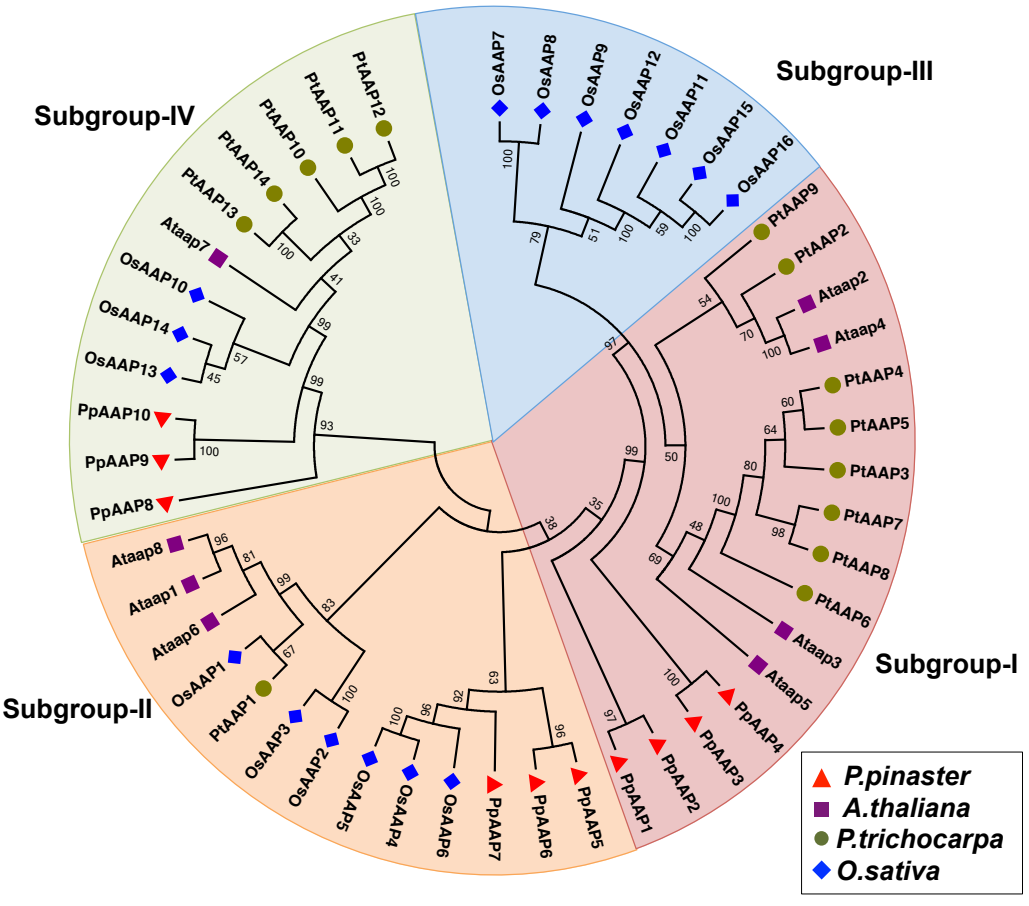


Figure 1. A circular, Neighbour-joining (NJ)-based tree of the amino acid permease (AAP) family in plants. The analysis was performed as described in the 'Materials and methods' section and the tree was generated using MEGA version 6.06. Bootstrap values are indicated (1000 replicates). Names of the species are abbreviated with a two-letter code *P. pinaster* (Pp), *A. thaliana* (At), *P. trichocarpa* (Pt) and *O. sativa* (Os). Accession numbers are provided as supplementary data inTable S2.

Relative expression levels of *P. pinaster AAP* genes were investigated at the ConGenIE database (Cañas et al. 2017) and revealed that among AAPs included in Subgroup-I, only PpAAP1 and PpAAP2 presented high levels of expression in developing roots (Figure 2). PpAAP1 presents high expression in developing root vascular tissue (DRV) and PpAAP2 in root meristem tissue (RM). These two AAPs were therefore selected as potentially involved in the primary uptake of amino acids from soi

Table 2. Identity percentages among proteins of the *P. pinaster* and *A. thaliana* amino acid permease (AAP) family. Highest identities shaded in grey colour are indicated.

	AtAAP1	AtAAP2	AtAAP3	AtAAP4	AtAAP5	AtAAP6	AtAAP7	AtAAP8
PpAAP1	57	62.1	63.4	64	62	60	50.3	58.7
PpAAP2	56.6	60.7	61.4	64.6	60	58.4	49.9	55.2
PpAAP3	58.8	67.4	68.8	69.4	68.2	59.5	50.3	59.1
PpAAP4	56.4	65.1	67.4	67.1	65.1	57.7	48.8	57.5
PpAAP5	54.9	58.5	59.6	60.1	58.7	56.2	48.3	55.8
PpAAP6	52.7	55.7	58.1	58.4	58	54.5	48.3	56.5
PpAAP7	51.8	51.5	52.6	53.2	52	54.2	46.6	52.6
PpAAP8	55.1	53.6	57.2	55.5	53.6	54.8	51.4	55.9
PpAAP9	50.6	54.1	55.8	54.4	52.2	51.1	52.4	51.4
PpAAP10	51.9	51.3	51.7	51.3	49	51.1	50.5	50.1

4.2. Effect of arginine availability on the expression of PpAAP1 and PpAAP2

To study whether *PpAAP1* and *PpAAP2* expression is affected by exogenous supply of arginine, 2 weeks-old maritime pine plantlets were cultured in hydroponic media at a concentration of 2.25 mM of arginine during two days. Time-course analysis of the expression of these genes was performed by qPCR in three sections of the root: upper root (1), medium root (2) and root tip (3) (Figure 3A).

The expression of *PpAAP1* was upregulated in response to arginine availability in the root tip and in the medium part of the root, reaching maximal values after 48 hours of culture (Figure 3B). By contrast, gene expression of *PpAAP2* does not show any significant difference respect to control plants after two days of arginine supply (Figure 3C). These results strongly suggest that only *PpAAP1* gene seems to be related with arginine transport in pine roots and it was selected for functional analysis.

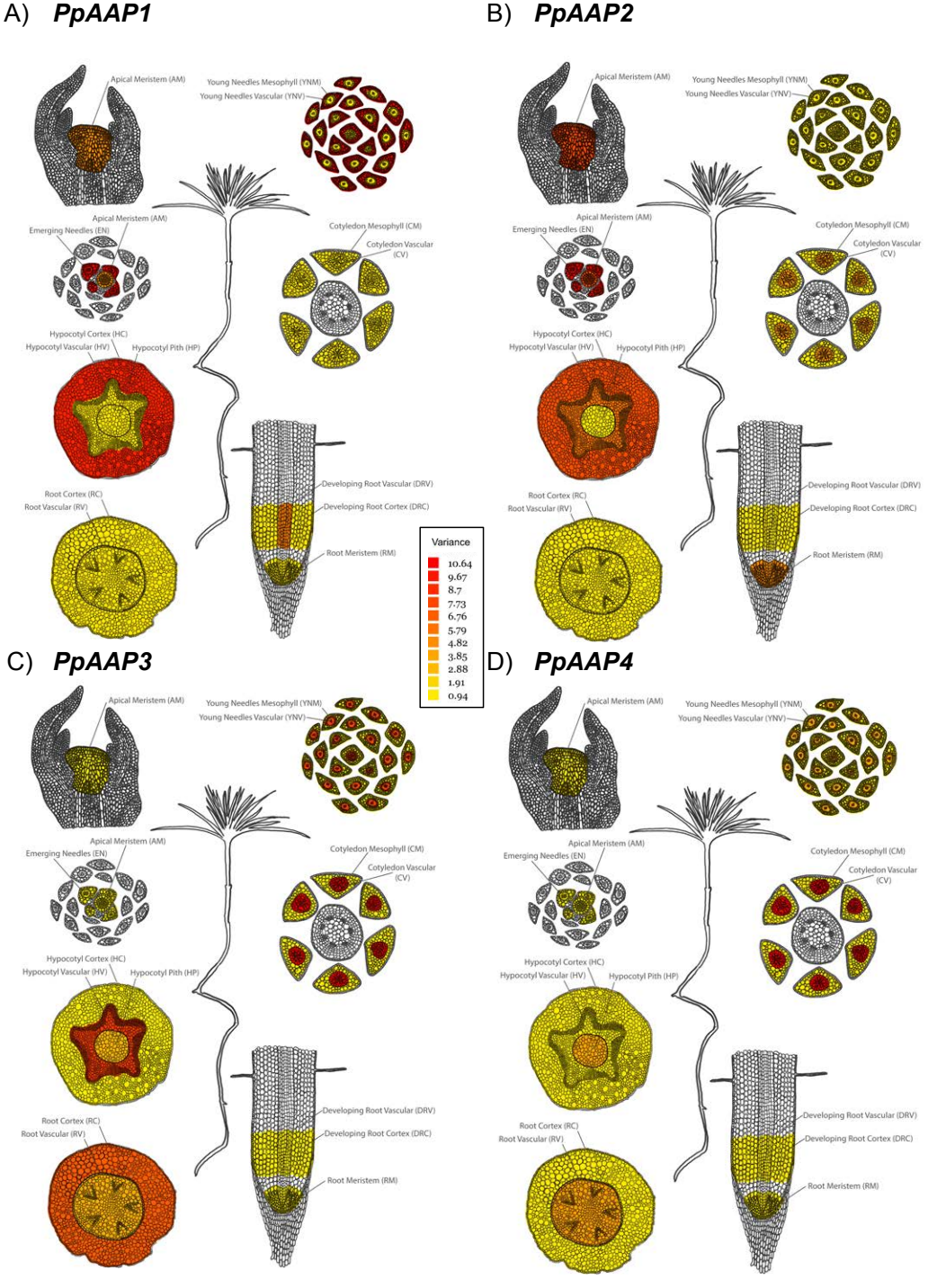


Figure 2. Relative expression levels of *P. pinaster* AAPs from subgroup-I. Image capture using the exImage tool at ConGenIE.org (http://v22.popgenie.org/microdisection/) indicating the expression and tissue localization for: A) *PpAAP1*, B) *PpAAP2*, C) *PpAAP3* and D) *PpAAP4*. Red colour indicates high level of expression and yellow colour low expression level.

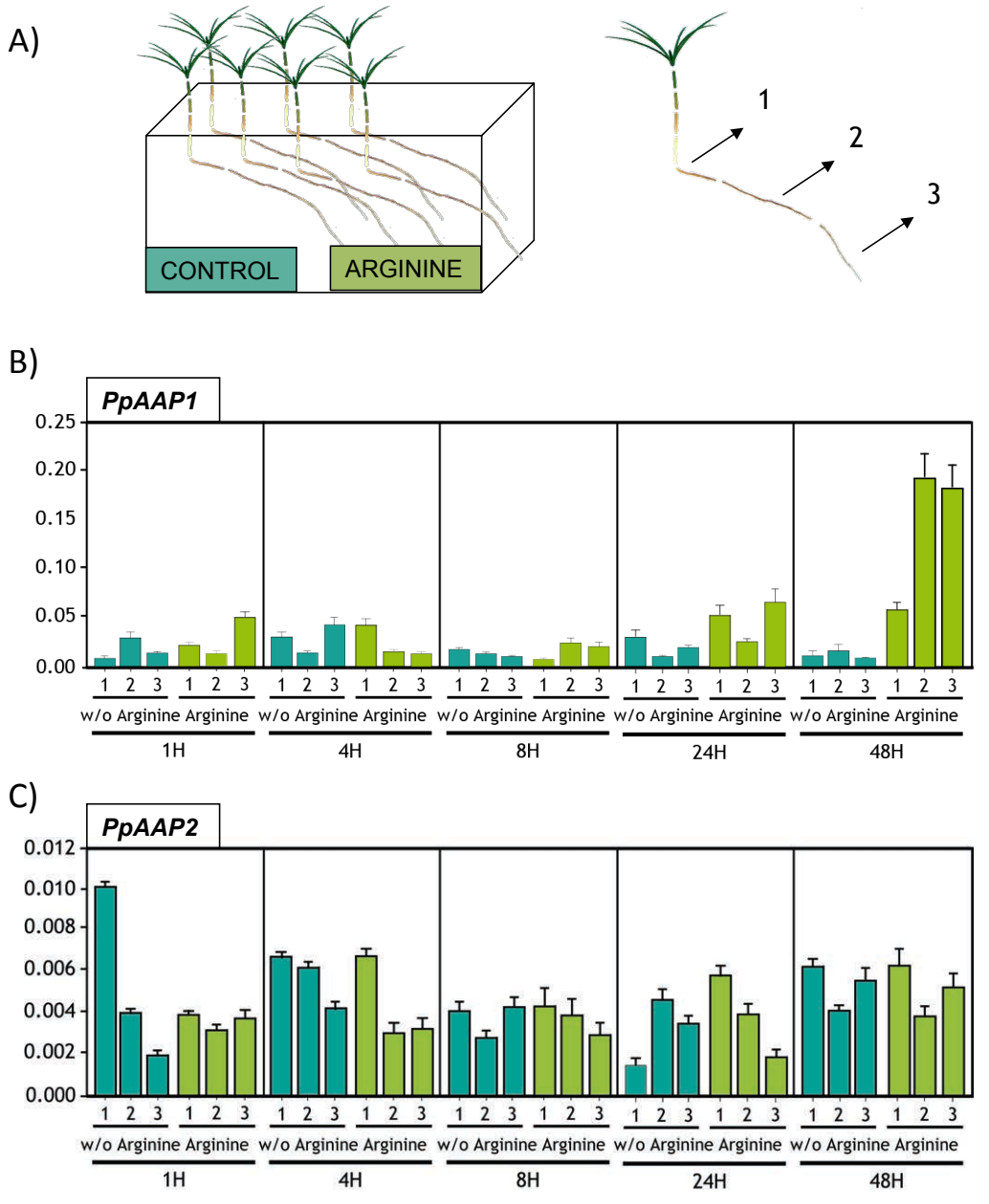


Figure 3. Relative expression levels of *PpAAP1* and *PpAAP2* in maritime pine roots cultured with arginine. Two-weeks old *P. pinaster* seedlings were submerged in hydroponic solution supplemented with arginine 2.25 mM. Root samples were collected after 1 hour (1H), 4 hours (4H), 8 hours (8H), 24 hours (24H) and 48 hours (48H). RNA was prepared from three sections of pine roots from the upper part to the tip (1, 2, 3), and the relative expression levels were determined by quantitative PCR (qPCR). Each value represents the mean +SE of three biological replicates in triplicated. Significant differences were calculated using Student's t-test ($P < 0.01$).

4.3. Functional characterization of PpAAP1 in yeast

To confirm that *PpAAP1* gene encodes a functional amino acid permease, a yeast complementation test was performed by transforming the mutant strain 22Δ8AA with the *PpAAP1* coding region under the control of the *PMA1* promoter (plasma membrane *ATPase* from yeast expression vector pDR196, Wipf et al. 2003). The mutant yeast strain 22Δ8AA is unable to use efficiently arginine, aspartate, glutamate and proline as sole N source (Fischer et al. 2002). As a control, the strain 22Δ8AA was transformed with the empty vector pDR196. Results were obtained after growing the yeasts in a media supplemented with only 3mM aspartic acid, glutamic acid, proline or arginine as sole nitrogen source for three days. These results indicated that the transformation of yeast with pDR196 bearing the *PpAAP1* coding sequence conferred the ability of strain 22Δ8AA to grow with arginine and proline at high concentrations (Figure 4A). On the contrary, the observed growth in glutamic and aspartic acid did not differ from the growth of the yeast transformed with the empty pDR196 vector.

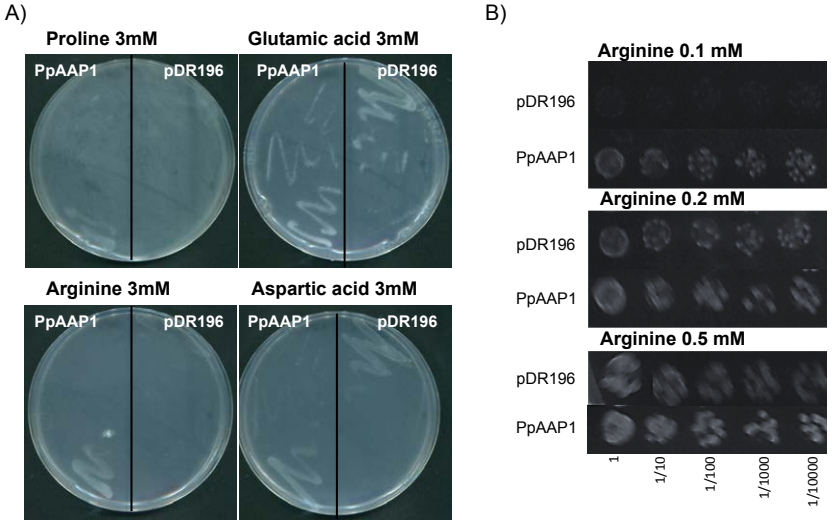
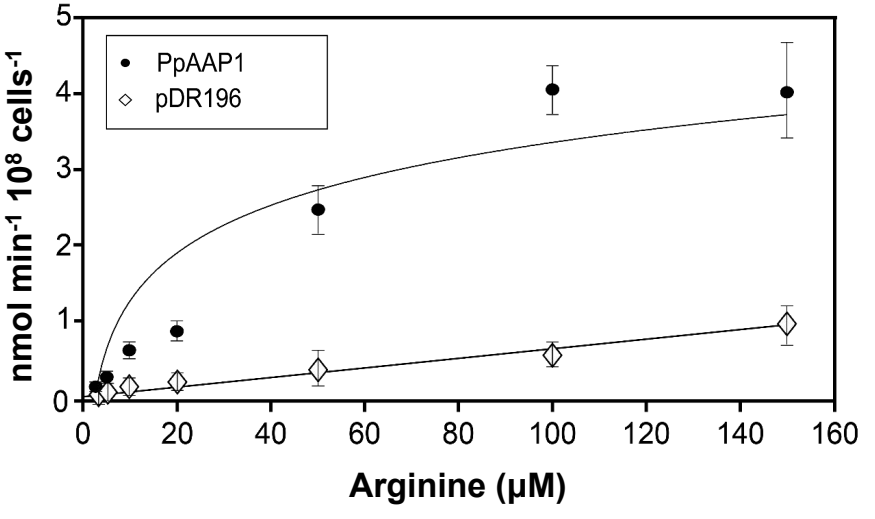


Figure 4. Functional complementation of yeast mutant strain 22Δ8AA by *Pinus pinaster* permease PpAAP1. Yeast strain *22Δ8AA* was transformed with the yeast expression vector pDR196 harbouring the coding sequence of *PpAAP1*, and empty pDR196 as control. A) Growth was assayed on N-free medium containing 3 mM proline, glutamic acid, arginine and aspartic acid as sole N source. B) Complementation was assayed at arginine concentrations of 0.1, 0.2 and 0.5 mM as sole N source. Each spot contains 10 µL of cells at five different dilutions (1,1:10,1:100,1:1000, 1:10000). All results were scored after 3 d of growth at 30 °C.

We also used a yeast-based functional complementation assay to assess if the defective yeast strain transformed with PpAAP1 could growth at low arginine concentrations. In this assay, the transformed yeast cells were grown at 0.1, 0.2 and 0.5 mM arginine, as sole nitrogen source (Figure 4B). The yeast strain was able to grow differentially compared with the yeast cells transformed with the empty vector (Figure 4B).

To determine the transport properties of *PpAAP1,* the uptake of L-[^{14}C]-arginine by yeast cells that expressed *PpAAP1* was quantified (Figure 5A). Expression of *PpAAP1* in the yeast mutant 22Δ8AA gave an increase of L-[^{14}C]-arginine over time compared to cells transformed with the empty vector. *PpAAP1*-mediated L-[^{14}C]-arginine uptake was concentration dependent and showed saturated kinetics with an apparent Km value of 83 μM and a capacity transport Vmax of 5 nmol min^{-1} 10^{8} cells^{-1} (Figure 5B).

A)



B)

PpAAP1	Km (μM)	Vmax(nmol min^{-1} 10^{8} cells^{-1})
	83	5

Figure 5. Kinetic analysis of PpAAP1 in transformed yeast. A) Concentration-dependent kinetics of L-[^{14}C]-arginine uptake by the yeast strain 22Δ8AA transformed with either empty pDR196 or with pDR196 harbouring the PpAAP1 construct. The values are expressed as the means ± standard deviation (SD) of three replicates. B) Kinetic parameters of PpAAP1 amino acid permease.

4.4. Growth of *Arabidopsis* mutants grown under high concentration of arginine

Two amino acid transporter AtAAP5 mutants (*aap5-1* and *aap5-2*) were grown in high arginine concentration of 1mM, phenotypic differences were determined and compared with wild-type plants. Both mutants had a T-DNA insertion located in the AtAAP5 fifth exon (Figure 6A). *Arabidopsis* WT, *aap5-1* and *aap5-2* plants display similar root development after 15 days growing with 3mM nitrate as the sole N source in the culture media. However, phenotypic differences were noticeable at 1mM of arginine (Figure 6B). The presence of high arginine concentration on the media caused shorter primary roots in wild-type plants and enhanced secondary root development (Figure 6C and 6D). In contrast, the growth of mutant plants was not affected by arginine displaying similar phenotype to the plants growing on nitrate as the sole nitrogen source.

The length of primary root was recorded for the *aap5-1* and *aap5-2* mutants and the wild type plants. Primary roots of aap5-1 and aap5-2 mutant's seedlings were significantly longer reaching an average length of 6.2 cm and 5.6 cm respectively, whereas WT roots reach an average length of 4 cm (Figure 6C). These differences suppose an increase of up 70% in the root growth of aap5-1 plants and 52% in aap5-2 plants, compared to WT plants. In the secondary root development no significant differences were observed in the number, but secondary roots were longer and thicker in WT plants compared with *aap5-1* and *aap5-2* mutants plants (Figure 6B and 6D). The relative expression level of *AtAAP5* gene was measured in WT, aap5-1, and aap5-2 mutant plants grown in high arginine levels (Figure 6E). Results indicate that *AtAAP5* expression was completely suppressed in both mutant lines.

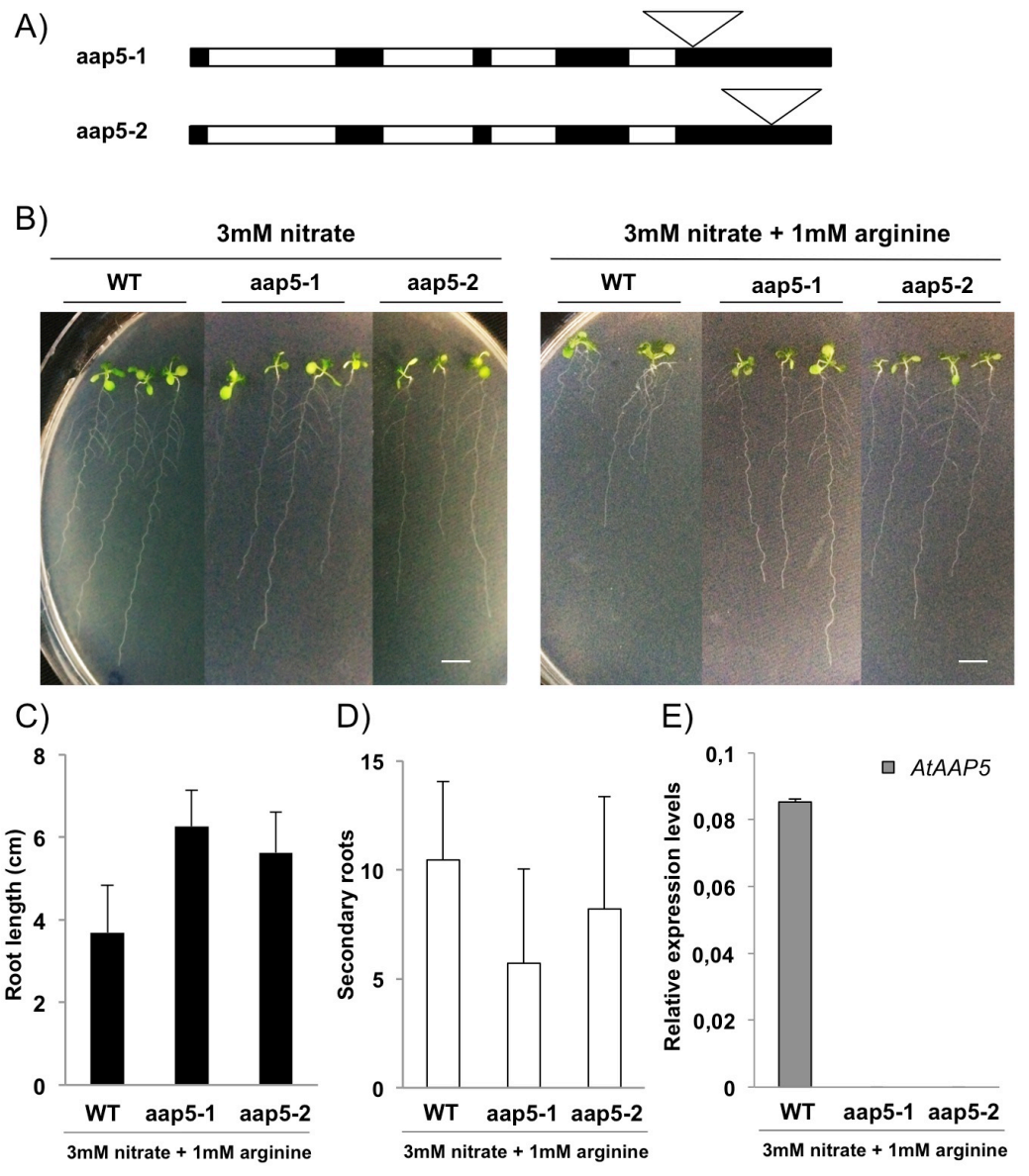


Figure 6. *Arabidopsis* mutant's plants grown under 1 mM arginine. **A)** Positions of the T-DNA insertions in aap5-1 (SALK_041999) and aap5-2 (SALK_099586) mutants are marked. Black and white boxes represent exon and intron regions, respectively. Inverted triangles indicate the sites of T-DNA insertions. **B)** *Arabidopsis* wild-type (WT), and T-DNA mutants (aap5-1 and aap5-2), grown on free nitrogen MS medium supplemented with only 3 mM nitrate or together with 1mM arginine for 15 days (Bar 1cm). **C)** Root length of plants grown on 3 mM nitrate and 1mM arginine. **D)** Number of secondary roots developed on plants grown on 3mM nitrate and 1mM arginine. **E)** Relative expression levels of *AtAAP5* on WT, aap5-1 and aap5-2 plants.

4.5. Expression of *PpAAP1* inhibits the growth under low and high arginine in the *Arabidopsis* aap5-1 mutants

To investigate whether PpAAP1 is a functional orthologous of *Arabidopsis* AAP5, the full-length cDNA of *PpAAP1* was used to transform the *Arabidopsis* aap5-1 and aap5-2 mutants. We studied the growth in aap5-1 and aap5-2 mutants and in the independent transgenic lines reached using different arginine concentrations, from the micromolar range at 10 µM, 50 µM and 100 µM, to 1mM, but maintaining the control nitrate concentration (3 mM) on the medium. After 15 days growth, visible phenotypic differences in development were observed between the mutant plants and transgenic lines. Under high arginine (1mM arginine), the growth of the transgenic lines was dramatically affected, exhibiting an inhibition of root growth compared to mutant plants (Figure 7A). This suggests that the arginine root uptake affected by the mutation in the *AtAAP5* gene was restored, and it had a direct influence on root development. The reduction in the root growth of the transgenic plants was also detected at concentrations of 50 µM and 100 µM of arginine, whereas no clear phenotypic differences between the mutants and the transgenic lines were detected at 10 µM of arginine (Figure S2). No phenotypic changes between mutant plants and mutant plants transformed with empty vector were observed (Figure 7A). The relative expression level of *PpAAP1* and *AtAAP5* genes were measured in WT, aap5-1 mutant, transgenic aap5-1 containing pGWB2 empty vector (EV) and transgenic aap5-1 overexpressing *PpAAP1* (L1#PpAAP1) plants. (Figure 7B).

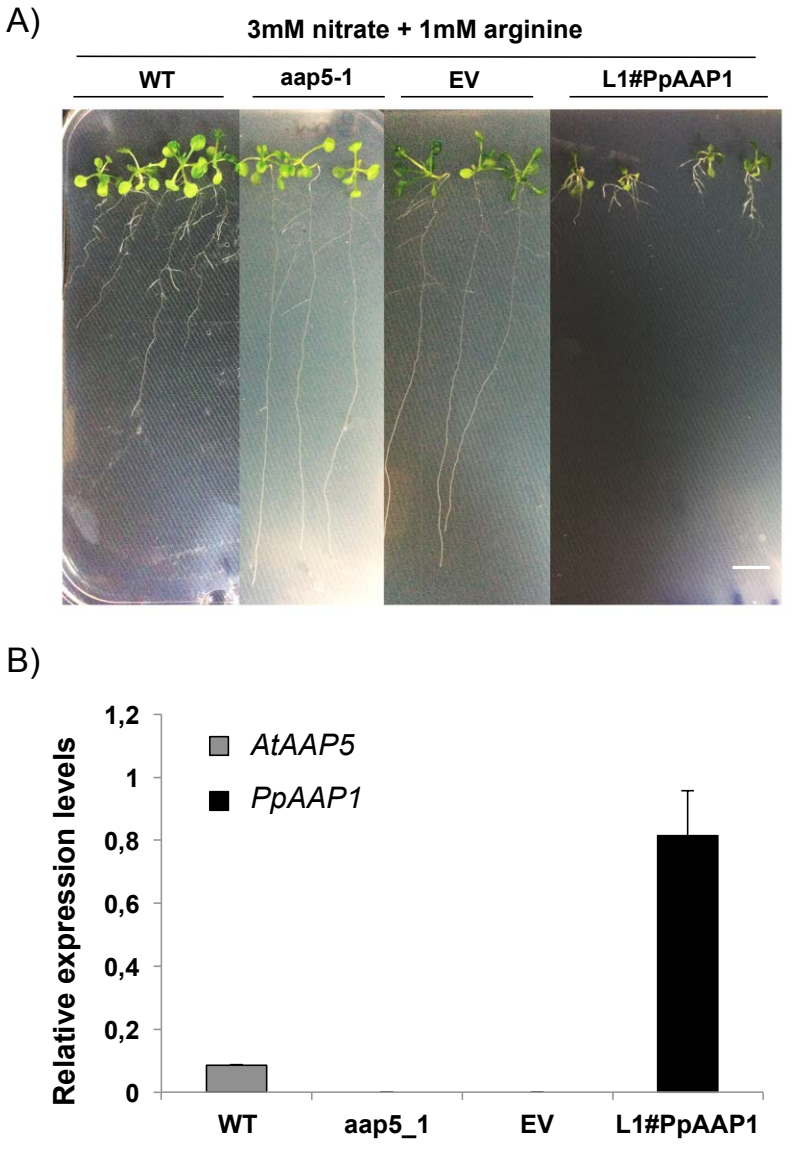


Figure 7. *Arabidopsis* mutant's plants overexpressing PpAAP1 grown under 1 mM arginine. A) Wild-type, EV (empty vector) and transgenic L1#PpAAP1 plants grown on 3 mM nitrate and 1 mM arginine for 15 days (Bar 1 cm). B) Relative expression levels of *AtAAP5* and *PpAAP1* genes in wild-type, aap5-1 mutant, EV and transgenic L1#PpAAP1 plants. EV; Transgenic aap5-1 line used as control, containing pGWB2 empty vector. L1#PpAAP1; Transgenic aap5-1 line overexpressing *PpAAP1.*

5. Discussion

Nitrogen (N) is one of the most limited resources for plant growth because large amounts of N are needed to synthesize nucleic acids and proteins. Among the 21 proteinogenic amino acids, arginine has the highest nitrogen to carbon ratio, which makes it especially suitable as a storage form of organic N. Long distance transport of arginine, to N storing organs or seeds, occurs through the vascular tissue and is presumably dependent on amino acid transporters of the AAP family (Winter et al. 2015). Especially important for long distant arginine transport in *Arabidopsis* seems to be AtAAP5, which transports arginine and lysine with high affinity (Svennerstam et al. 2008) and have an important role in the uptake of basic amino acids by roots (Svennerstam et al. 2011).

The main research on amino-N transporters in plants has been performed on *Arabidopsis,* where a number of amino acid transporters have been characterized, but little is known about the amino acid transport in other plant species (Fisher et al. 1998). In conifers, arginine is a key amino acid for N storage and mobilization (Cañas et al. 2016) that constitutes a large proportion of the amino acid content in the seed storage proteins. In the soil of boreal forests organic nitrogen in the form of amino acids such as arginine is an important fraction of the total nitrogen that is available for tree nutrition (Inselsbacher and Näsholm 2012). Increased availability of nitrogen triggers the transient accumulation of arginine in needles and wood of conifers possibly reflecting the internal nitrogen status (Nordin et al. 2001). However, in spite of the relevance of arginine in conifer nitrogen nutrition and metabolism no specific transporter for this amino acid has yet been identified.

Taking advantage of the current available genomic resources in *P. pinaster,* the members of the amino acid permease AAP family have been investigated in this gymnosperm species. The AAP family in *P. pinaster* is composed for at least 10 members, 4 of them included into the Subgroup-I, 3 in the Subgroup-II and other 3 in the Subgroup-IV (Figure 1). Pine transporters of the first Subgroup were selected to study the possible relation with arginine root uptake; mainly due to their higher structural similarities with

AtAAP5 (Table 2). The pattern of expression indicates that relative expression levels of PpAAP1, PpAAP2, PpAAP3 and PpAAP4 are detected in almost all plant tissues (Figure 2). Nevertheless, only PpAAP1 and PpAAP2 present high expression levels in root developing tissues (Figure 2A and 2B), whereas PpAAP3 expression is mainly located in vascular tissues (Figure 2C) and PpAAP4 in cotyledons (Figure 2D). These wide expression distributions agree with the hypothesis that many transporters are expressed in different tissues at different developmental stages and therefore have multiple functions in plant nitrogen metabolism (Liu & Bush, 2006).

The first evidence of the PpAAP1 relation with arginine root uptake was observed when a 10-fold increase in the expression level was detected in roots of plants submerged in a solution containing high concentration of arginine (Figure 3). These results suggest a related role of PpAAP1 in this amino acid transport and in its mobilization during plant development. Moreover, the higher expression levels observed in the root tip in response to arginine availability further support a role of PpAAP1 in the uptake of external arginine by pine roots.

Previous studies on AAP transporters have been performed using several organisms, as yeast and *Arabidopsis*, varying concentrations of externally supplied amino acids (Forsum et al. 2008; Svennerstam et al. 2008; Svennerstam et al. 2007; Hirner et al. 2006; Fischer et al. 1995; Frommer et al. 1993). AAP transporters family recognize a wide spectrum of amino acids as determined by competition studies derived from gene expression studies in yeast. Overall, the results obtained in the yeast complementation test identified PpAAP1 as a participant in the transport of arginine (Figure 4), and kinetic analysis indicates that PpAAP1 is able also to function as a transporter of arginine at micromolar ranges (Figure 5).

Another line of evidence showed that root uptake of arginine in *Arabidopsis* was severally affected by loss of AtAAP5, and that T-DNA mutant plants had substantially larger roots after growing on 3mM nitrate and 1mM arginine (Svennerstam et al. 2008). In these mutant plants, the uptake of lysine and arginine was 68 to 88% lower than in the wild-type plants, while the uptake rates of acidic and neutral amino acid remained unchanged. These

plants were not affected by the arginine millimolar concentration range a, probably due to the fact that they are not able to transport it. This would explain the normal rates of development in these conditions. On the contrary, wild-type plants at this high arginine level displayed a considerable decrease in the root length (Figure 6). The phenotypic behaviour of *Arabidopsis* AAP5 mutant plants overexpressing *PpAAP1* at high arginine concentration is in agreement with an efficient transport of this amino acid. The possible recovery of the arginine transport capacity in these plants originates a dramatic inhibition of the root growth, even greater than that observed for wild-type plants, affecting even to the aerial part with smaller rosettes (Figure 7). Taken together, these findings suggest a function for PpAAP1 to transport efficiently arginine. The free amino acid pool *in plant* could be responsible of the limited growth of the transgenic *Arabidopsis*.

Typically, amino acid concentrations in soils of agricultural systems, and temperate and boreal forests are in the micromolar range (Jämtgård et al. 2010; Raab et al. 1999; Kielland, 1994), being necessary the study and characterization of plant amino acids uptake at naturally occurring soil concentrations. When root amino acid uptake has been studied at a soil-relevant concentration range, the affinity constant (Km) for different compounds has been found to fall within the range from 10 μM (arginine; Soldal & Nissen, 1978) to 300 μM (glutamine; Kielland, 1994). The AtAAP5 amino acid transporter mutation also caused 10-fold decreased rates of arginine uptake at 10 μM concentration, suggesting that AtAAP5 is the most important amino acid transporter for root uptake in the μM range. To elucidate whether PpAAP1 was important for root uptake of arginine at field-relevant concentrations, we have studied the effect of arginine concentrations ranging from 10 μM to 100 μM in *Arabidopsis* plant growth. Our results indicate that mutant *Arabidopsis* plants overexpressing PpAAP1 from pine were able to transport usual arginine soil concentrations and the inhibitory root growth effect were easily distinguishable at 50 μM arginine.

Early studies by Lee et al. (2007) identified AtAAP1 as an important amino acid transporter for plant roots although the characterized phenotype of the mutants was only displayed when plants were exposed to relatively high

amino acid concentrations (0.15–10 mM). Kinetic studies of the root amino acid AtAAP1, in the concentration range 2–50 μM, confirm the function of AtAAP1 to mediate amino acid uptake at these external concentrations.

In *Arabidopsis*, the assignment of physiological function to individual AAPs has not been reported so far (Tegeder, 2014). To unravel the role of each AAP transporter in amino acid uptake is still needed more research in their individual physiology, biochemistry and molecular biology. The identification of this multitude of different transporters in plants and the clarification of their specific roles and potential redundancy in function is currently the objective of numerous studies. The possible functional redundancy among some transporters could be the explanation of why the *Arabidopsis* mutants of the AtAAP3 amino acid transporter cannot be distinguished from wild-type plants (Okumoto et al. 2004).

To assess a specific physiological role to PpAAP1 in nitrogen transport in conifers is still a challenge, being necessary to determine if is the unique pine transporter involved in arginine uptake or it could also be involved in the uptake of other amino acids. Additional studies in the *Pinus pinaster* AAP family are underway to understand the mechanisms of how pines are able to incorporate amino acids from soil and mobilize them through the whole plant.

6. Supplemental material

Table S1. Gene specific primer sequences used.

Gene	Forward 5´-3	Reverse 5´-3
PpAAP1	TCGATCCATTGTATTTCTCAGTCAA	AGAACATCCTGTAGTGCATTAAG
PpAAP2	TGTTGTGCAAGATTTGAAAATGTAC	GGCTAAATTCCAAGAAACAACTTGCAG
aap5-1	AAAGAAAGAACGGCTCCGAATT	TTGTTGTCATCACGAAGAACG
aap5-2	TTGGGACAGTGACACTGAGTG	AACAATGCCAATAACAGATCCC
pGWB2	CAGTGACAGTTGACAGCGACA	AATTCGAGCTCTAAGCGCTG
AtAAP5_qPCR (Fisher et al. 1995)	GCCTGCCGAAGTAAACACGATGAG	ACCCTCCA TGAGCGAGGAGA TTTC
PpAAP1_qPCR	ACAGGAAGATCCCGCGTTACT	AGAACATCCTGTAGTGCATTAAG

Table S2. Accession numbers of AAPs from *P. pinaster, A. thaliana, P. trichocarpa* and *O. sativa.*

P.pinaster		
PpAAP1(isotig28092) PpAAP4 (isotig44984) PpAAP7 (unigene41903) PpAAP10 (isotig43366)	PpAAP2(isotig32920) PpAAP5 (unigene165) PpAAP8 (isotig44509)	PpAAP3(isotig44810) PpAAP6 (unigene2046) PpAAP9 (isotig22404)
A.thaliana		
AtAAP1 (At1g58360) AtAAP4 (At5g63850) AtAAP7 (At5g23810)	AtAAP2 (At5g09220) AtAAP5 (At1g44100) AtAAP8 (At1g10010)	AtAAP3 (At1g77380) AtAAP6 (At5g49630)
P.trichocarpa		
PtAAP1 (Poptr1_1:202159) PtAAP3 (Poptr1_1:818545) PtAAP5 (Poptr1_1:551420) PtAAP7 (Poptr1_1:551421) PtAAP9 (Poptr1_1:200759) PtAAP11 (Poptr1_1:270039) PtAAP13 (Poptr1_1:285212)	PtAAP2 (Poptr1_1:207199) PtAAP4 (Poptr1_1:595125) PtAAP6 (Poptr1_1:595874) PtAAP8 (Poptr1_1:674664) PtAAP10 (Poptr1_1:582548) PtAAP12 (Poptr1_1:642685) PtAAP14 (Poptr1_1:582546)	
O.sativa		
OsAAP1 (Os07g0134000) OsAAP3 (Os06g0556000) OsAAP5 (Os01g0878400) OsAAP7 (Os05g0424000) OsAAP9 (Os02g0102200) OsAAP11 (Os11g0195600) OsAAP13 (Os04g0470700) OsAAP15 (Os12g0181600)	OsAAP2 (Os06g0228600) OsAAP4 (Os12g0194900) OsAAP6 (Os01g0878700) OsAAP8 (Os01g0882800) OsAAP10 (Os02g0722400) OsAAP12 (Os12g0195100) OsAAP14 (Os04g0659800) OsAAP16 (Os12g0181500)	

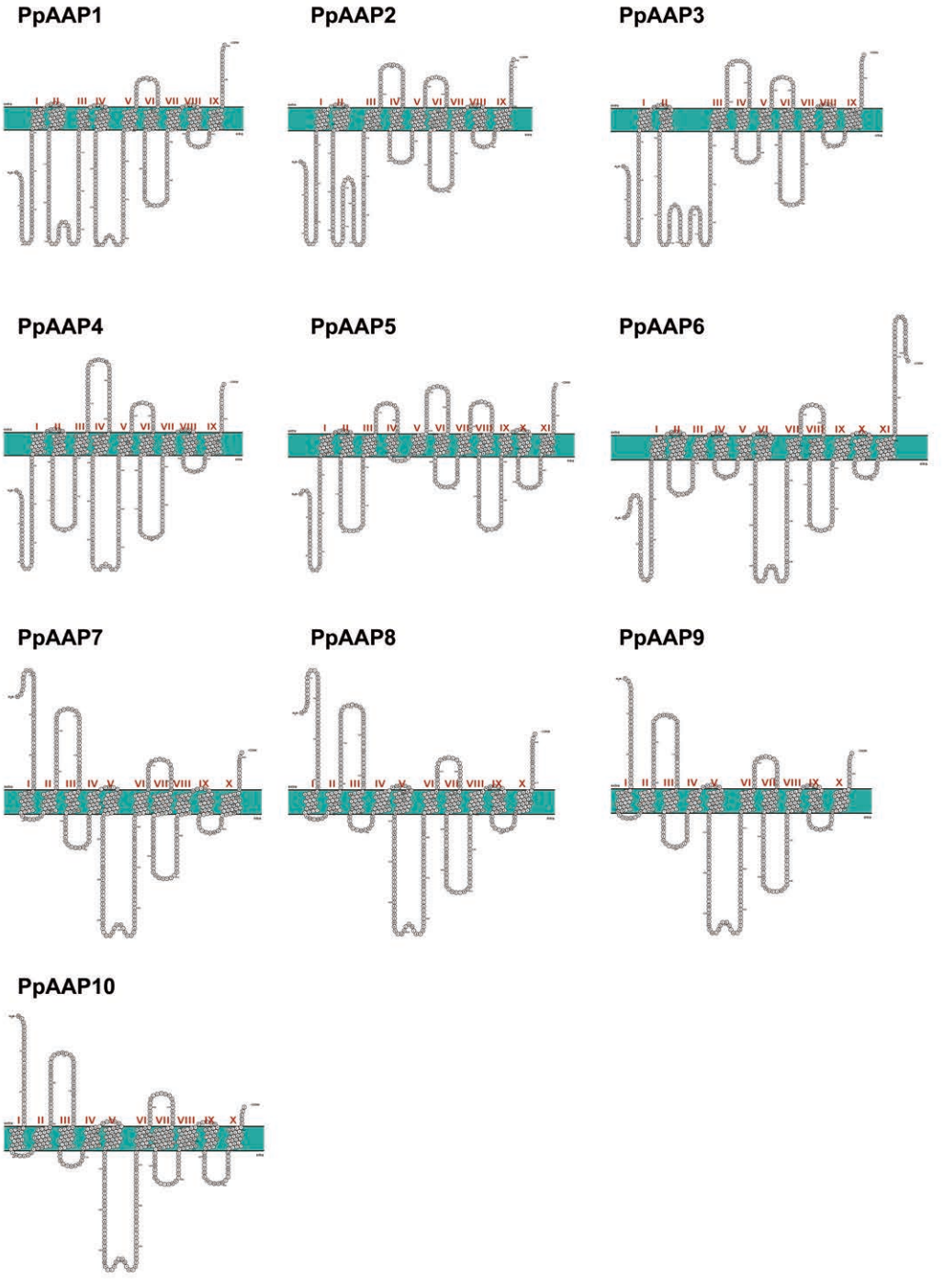


Figure S1. Transmembrane prediction of P.pinaster AAP proteins (THMM).

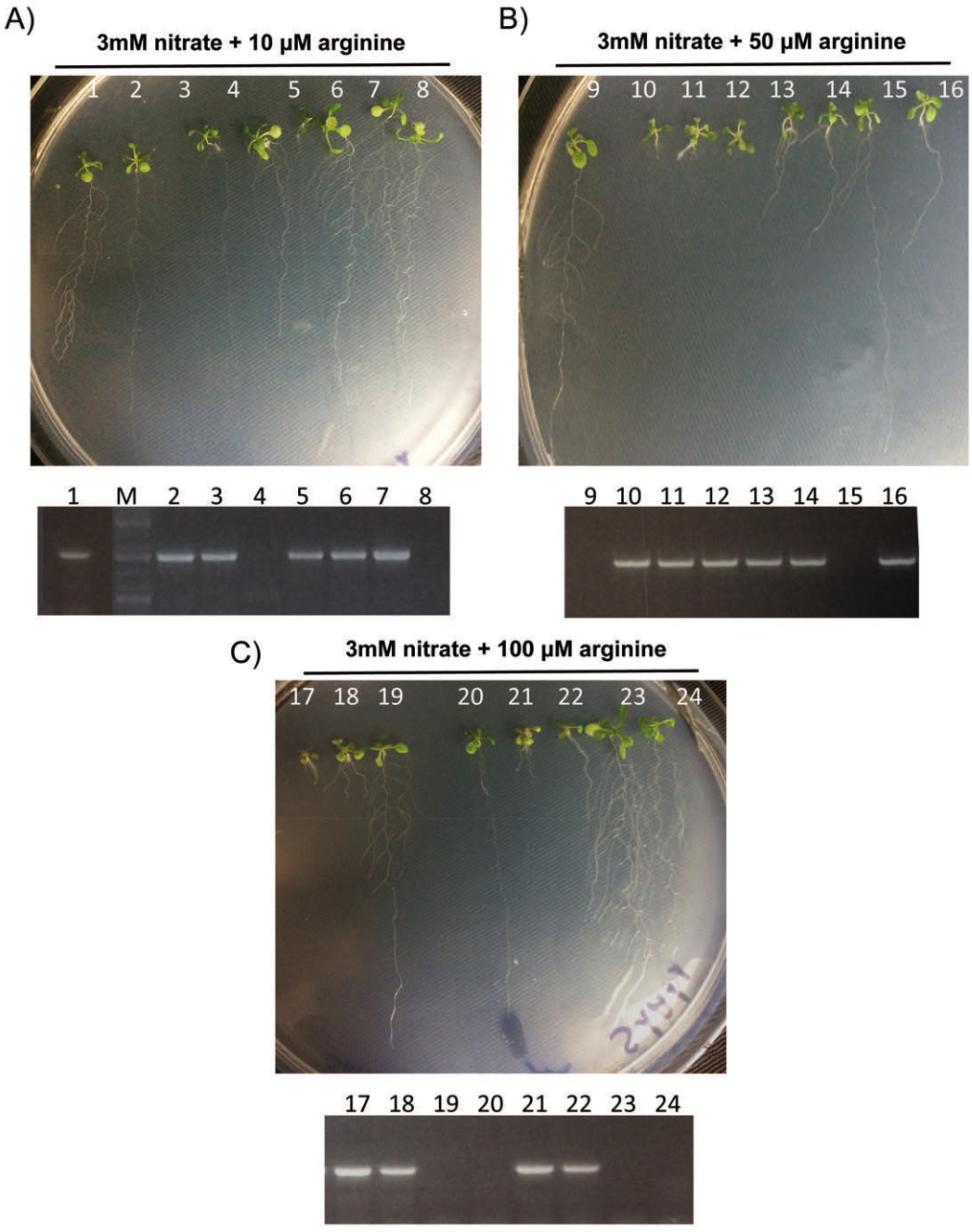


Figure S2. Mutant plants overexpressing PpAAP1 grown on micromolar ranges of arginine. Transgenic plants (L1#PpAAP1) grown on 3 mM nitrate and 10µM (A), 50 µM (B) or 100 µM (C) arginine for 15 d. Presence of the PpAAP1 transgene was detected in each individual plant by PCR with specific primers

7. References

Boorer KJ, Fischer WN (1997) Specificity and stoichiometry of the Arabidopsis H^+ amino acid transporter AAP5. *Journal of Biological Chemistry* 272: 13040-13046

Canales J, Flores-Monterroso A, Rueda-López M, Ávila C & Cánovas FM (2010) Identification of genes regulated by ammonium availability in the roots of maritime pine trees. *Amino Acids* 39: 991-1001.

Canales J, Avila C, Pacheco-Villalobos D, Díaz-Moreno S, Molina-Rueda JJ, Ariza D, Navarro-Cerrillo, Cantón FR, Claros MG, Cánovas FM (2012) Gene expression profiling in the stem of young maritime pine trees: detection of ammonium stress-responsive genes in the apex. *Trees* 26: 609-619.

Canales J, Bautista R, Label P, Gómez-Maldonado J, Lesur I, Fernández-Pozo N, et al. (2014) De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. *Plant Biotechnol Journal 12*: 286-99.

Cañas RA, de la Torre F, Pascual MB, Ávila C, Cánovas FM (2016) Nitrogen economy and nitrogen environmental interactions in conifers. *Agronomy* 6: 26.

Cañas RA, Li Z, Pascual MB, Castro-Rodríguez V, Ávila C, Sterck L, Van de Peer Y and Cánovas FM (2017) The gene expression landscape of pine seedling tissues. *The Plant Journal*, 10.1111/tpj.13617.

Clough SJ and Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana. The Plant Journal* 16: 735-43.

Couturier J, de Fay E, Fitz M, Wipf D, Blaudez D, Chalot M (2010 a) PtAAP11, a high affinity amino acid transporter specifically expressed in differentiating xylem cells of poplar. *Journal of Experimental Botany* 61: 1671-1682.

Fischer WN, Kwart M, Hummel S, Frommer WB (1995) Substrate specificity and expression profile of amino acid transporters (AAPS) in Arabidopsis. *Journal of Biological Chemistry* 270: 16315-16320.

Fischer WN, Andre B, Rentsch D, Krolkiewicz S, Tegeder M, Breitkreuz K, Frommer WB (1998) Amino acid transport in plants. *Trends in Plant Science* 3: 188-195.

Fischer WN, Loo DD, Koch W, Ludewig U, Boorer KJ, Tegeder M, Rentsch D, Wright EM, Frommer WB (2002) Low and high affinity amino acid H^+-cotransporters for cellular import of neutral and charged amino acids. *Plant Journal* 29(6): 717-31.

Forsum O, Svennerstam H, Ganeteg U, Naesholm T (2008) Capacities and constraints of amino acid utilization in Arabidopsis. *New Phytologist*

179: 1058-1069.

Frommer WB, Hummel S, Riesmeier JW (1993) Expression cloning in Yeast of a cDNA encoding a broad-specificity amino acid permease from *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America 90: 5944-5948.

Gruffman L, Jämtgård S, NäsholmT (2014) Plant nitrogen status and co-occurrence of organic and inorganic nitrogen sources influence root uptake by Scots pine seedlings, *Tree Physiology* 34: 205-213.

Hirner A, Ladwig F, Stransky H, Okumoto S, Keinath M, Harms A, Frommer WB, Koch W (2006) *Arabidopsis* LHT1 is a high-affinity transporter for cellular amino acid uptake in both root epidermis and leaf mesophyll. *Plant Cell* 18: 1931-1946.

Hsu LC, Chiou TJ, Chen LS, Bush DR (1993) Cloning a plant amino acid transporter by functional complementation of a Yeast amino acid transport mutant. Proceedings of the National Academy of Sciences of the United States of America 90: 7441-7445.

Hutchinson HB, Miller NHJ (1912) The direct assimilation of inorganic and organic forms of nitrogen by higher plants. *The Journal of Agricultural Science* 4(03): 282-302.

Inselsbacher, E. and Näsholm, T. (2012), The below-ground perspective of forest plants: soil provides mainly organic nitrogen for plants and mycorrhizal fungi. New Phytologist, 195: 329–334.

Jämtgård S, Näsholm T, Huss-Danell K (2010) Nitrogen compounds in soil solutions of agricultural land. *Soil Biology & Biochemistry* 42: 2325-2330.

Jauniaux JC, Vandenbol M, Vissers S, Broman K and Grenson M (1987) Nitrogen catabolite regulation of proline permease in *Saccharomyces cerevisiae.* Cloning of the *PUT4* gene and study of PUT4 RNA levels in wild-type and mutants strains. *Eur. J. Biochem.* 164: 601-606.

Jones DL, Kielland K (2002) Soil amino acid turnover dominates the nitrogen flux in permafrost-dominated taiga forest soils. Soil Biology & Biochemistry 34: 209-219.

Kasajima I, Ide Y, Ohkama-Ohtsu N, Hayashi H, Yoneyama T, Fujiwara T (2004) A protocol for rapid DNA extraction from *Arabidopsis thaliana* for PCR analysis. *Plant Mol. Biol. Rep.* 22: 49-52.

Kielland K (1994) Amino acid absorption by arctic plants. Implication for plant nutrition and nitrogen cycling. *Ecology* 75: 2373-2383.

Kielland K (1995) Landscape patterns of free amino acids in arctic tundra soils. *Biogeochemistry* 31: 85-98.

Kronzucker HJ, Siddiqi MY, Glass ADM (1996) Kinetics of $NH4^+$ influx in spruce. *Plant Physiol.* 110: 773-779.

Lee Y-H, Foster J, Chen J, Voll LM, Weber APM, Tegeder M (2007) AAP1 transports uncharged amino acids into roots of Arabidopsis. *Plant Journal* 50: 305-319.

Lipson D, Näsholm T (2001) The unexpected versatility of plants: organic nitrogen use and availability in terrestrial ecosystems. *Oecologia* 128: 305-316.

Liu X, Bush DR (2006) Expression and transcriptional regulation of amino acid transporters in plants. *Amino Acids* 30: 113-120.

Li Z and Bush DR (1990) ΔpH-dependent amino acid transport into plasma membrane vesicles isolated from sugar beet leaves. *Plant Physiol.* 94: 268- 277.

Llebrés MT, Pascual MB, Debille S, Trontin JF, Avila C, Cánovas FM (2017) The role of arginine metabolic pathway during embryogenesis and germination in maritime pine (*Pinus pinaster* Ait.). *Tree Physiol*ogy, doi: 10.1093/treephys/tpx133.

Murashige T and Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum* 15: 473-497.

Näsholm T, Kielland K, Ganeteg U (2009) Uptake of organic nitrogen by plants. *New Phytologist* 182: 31-48.

Nordin A, Högberg P. and Näsholm T (2001) Soil nitrogen form and plant nitrogen uptake along a boreal forest productivity gradient. Oecologia 129: 125. https://doi.org/10.1007/s004420100698

Öhlund J and Näsholm T (2004) Regulation of organic and inorganic nitrogen uptake in Scots pine (*Pinus sylvestris*) seedlings. *Tree Physiol.* 24:1397-402.

Omasits U, Ahrens H, Müller S, Wollscheid B (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30: 884-886.

Okumoto S, Schmidt R, Tegeder M, Fischer WN, Rentsch D, Frommer WB, Koch W (2002) High affinity amino acid transporters specifically expressed in xylem parenchyma and developing seeds of Arabidopsis. *J. Biol. Chem.* 277 45338-45346.

Okumoto S, Koch W, Tegeder M, Fischer WN, Biehl A, Leister D, Stierhof YD, Frommer WB (2004) Root phloem-specific expression of the plasma membrane amino acid proton co-transporter AAP3. *Journal of Experimental Botany* 55: 2155-2168.

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, Orvis J, Haas B, Wortman J, Buell CR (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res.* 35: 883-7.

Pratelli R, Pilot G (2014) Regulation of amino acid metabolic enzymes and transporters in plants. Journal of Experimental Botany 65, 5535-5556.

Raab TK, Lipson DA, Monson RK (1996) Non-mycorrhizal uptake of amino acids by roots of the alpine sedge *Kobresia myosuroides*: Implications

for the alpine nitrogen cycle. *Oecologia* 108: 488-494.

Raab TK, Lipson DA, Monson RK (1999) Soil amino acid utilization among species of the Cyperaceae: Plant and soil processes. *Ecology* 80: 2408-2419.

Rentsch D, Schmidt S, Tegeder M (2007) Transporters for uptake and allocation of organic nitrogen compounds in plants. FEBS Letters 581: 2281-2289.

Svennerstam H, Ganeteg U, Bellini C, Näsholm T (2007) Comprehensive screening of Arabidopsis mutants suggests the lysine histidine transporter 1 to be involved in plant uptake of amino acids. *Plant Physiology* 143: 1853-1860

Svennerstam H, Ganeteg U, Näsholm T (2008) Root uptake of cationic amino acids by Arabidopsis depends on functional expression of amino acid permease. *New Phytologist* 180: 620-630

Svennerstam H, Jämtgård S, Ahmad I, Huss-Danell K, Näsholm T, Ganeteg U (2011) Transporters in Arabidopsis roots mediating uptake of amino acids at naturally occurring concentrations. *New Phytologist* 191: 459-467.

Taylor MR, Reinders A, Ward JM (2015) Transport Function of Rice Amino Acid Permeases (AAPs). *Plant Cell Physiol.* 56(7): 1355-1363.

Tegeder M (2014) Transporters involved in source to sink partitioning of amino acids and ureides: opportunities for crop improvement. *Journal of Experimental Botany* 65: 1865-1878.

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

Winter G, Todd CD, Trovato M, Forlani G, Funck D (2015) Physiological implications of arginine metabolism in plants. *Front Plant Sci* 6: 534.

Wright DE (1962) Amino acid uptake by plant roots. *Arch Biochem Biophys*. 97:174-180.

Xu G, Fan X, Miller AJ (2012) Plant nitrogen assimilation and use efficiency. *Annu Rev Plant Biol* 63: 153-182.

Yuan Q, Ouyang S, Wang A, Zhu W, Maiti R, Lin H, Hamilton J, Haas B, Sultana R, Cheung F, Wortman J, and Buell CR (2005) The Institute for Genomic Research Osa1 Rice Genome Annotation Database. *Plant Physiol.* 138(1): 18-26.

IV. DISCUSIÓN GENERAL

1. Utilización y mejora de la embriogénesis somática en coníferas

La embriogénesis somática (ES) en coníferas ha sido objeto de estudio desde hace muchos años y son numerosas las especies del género *Pinus* en las que se han desarrollado protocolos viables de regeneración *in vitro* mediante esta técnica. Como ejemplo, en el Artículo 1 del presente manuscrito se presenta la ES y conversión en planta del híbrido *Pinus strobus x Pinus wallichiana.* Además, la ES proporciona el material idóneo para la transformación genética de coníferas y lo más importante, permite la criopreservación de las líneas transgénicas obtenidas que pueden ser utilizadas posteriormente para su caracterización molecular y fisiológica.

Las coníferas son consideradas especies recalcitrantes para ser transformadas genéticamente y regeneradas mediante cultivo *in vitro*, y aunque actualmente los protocolos disponibles para la transformación y criopreservación de las líneas celulares son mucho más efectivos (Klimaszewska et al. 2004; Trontin et al. 2007), la eficiencia de transformación continúa siendo relativamente baja. Los trabajos publicados demuestran que esta tecnología es una herramienta muy útil para realizar estudios funcionales (Klimaszewska et al. 2007; Álvarez et al. 2012). En el transcurso de este trabajo (Artículo 2) hemos utilizado protocolos de ES para la generación de líneas transgénicas RNAi de *P. pinaster* para un factor de transcripción de la familia NAC (NAC1). La maduración de los embriones somáticos y la regeneración *in vitro* nos han permitido realizar análisis fenotípicos y moleculares en varias líneas silenciadas y demostrar la relación de este factor de transcripción con el desarrollo vascular de pino. Este trabajo es un claro ejemplo del potencial biotecnológico que tiene el uso de la embriogénesis somática para la transformación genética y el estudio funcional de genes en especies de interés comercial.

Aunque se ha evolucionado notablemente en el establecimiento de la ES en *P. pinaster,* es necesario avanzar en el estudio y comprensión de los mecanismos moleculares que regulan el desarrollo del embrión. Este aspecto es fundamental para conseguir optimizar todas las etapas de la ES, desde la

inducción del tejido embriogénico, hasta la maduración y germinación de los embriones somáticos (Trontin et al. 2016).

Durante la realización de esta Tesis Doctoral se ha pretendido proporcionar algunas respuestas que permitan la optimización de la ES en *P. pinaster*. Para ello se ha abordado el estudio de diversos aspectos clave en la economía del nitrógeno durante el desarrollo embrionario, como son el efecto de las diferentes fuentes de nitrógeno en la ES, la acumulación de proteínas de reserva y la regulación del metabolismo de la arginina en las primeras etapas del desarrollo.

2. Importancia de la nutrición nitrogenada durante la ES

La regeneración de plantas es el principal objetivo de la ES, pero la maduración de los embriones, su germinación y posterior conversión en plantas siguen siendo procesos limitantes para la aplicación de esta técnica a gran escala.

Una de las fases más determinantes para la obtención de plantas mediante ES es la maduración del embrión. Durante este periodo del desarrollo embrionario tiene lugar la biosíntesis y deposición de proteínas de reserva, que son proteínas ricas en aminoácidos con alto contenido en N, como la arginina. La movilización posterior de estas proteínas de reserva es esencial durante la germinación y, por tanto, para la producción de plantas viables.

Está descrito que los embriones somáticos acumulan menos proteínas de reserva en comparación con los embriones cigóticos en un mismo estado de desarrollo. Esta deficiencia en las reservas nitrogenadas se ha señalado como la causa principal de las bajas tasas de conversión en plantas y el bajo rendimiento en los estudios de campo durante las primeras etapas de germinación (Trontin et al. 2016). De forma generalizada se piensa que se podría mejorar la viabilidad de los embriones somáticos si metabólicamente se parecieran más a los embriones cigóticos. Es frecuente considerar que los embriones somáticos están preparados para iniciar la fase de germinación cuando se asemejan fenotípicamente a los embriones

cigóticos maduros, lo que tiene lugar al cabo de unas 12 semanas en *P. pinaster.* Este periodo es muy dependiente de la especie y genotipo, y está principalmente basado en la observación de la morfología que presenta el embrión. En realidad, este procedimiento sólo es indicativo de que la maduración ha tenido lugar, pero no es hasta más tarde cuando se verifica si son capaces de convertirse en plantas viables. Disponer de información acerca de la calidad de los embriones somáticos respecto a las proteínas de reserva o el contenido en agua, podría indicar cuando es el mejor momento para que inicien la fase de germinación y alcancen mayores tasas de conversión en plantas (Klimaszewska et al. 2004).

En el trabajo descrito en el Artículo 3 de esta Tesis Doctoral hemos tratado de determinar el estado óptimo de maduración de los embriones somáticos en *P. pinaster* analizando las proteínas de reserva mayoritarias en diferentes etapas del desarrollo y realizando análisis comparativos con la embriogénesis cigótica (EC). Entre las proteínas mayoritarias acumuladas durante la maduración de los embriones cigóticos se encuentran proteínas de reserva de la familia de las vicilinas y leguminas. Esas mismas proteínas se han identificado en embriones somáticos de *P. strobus* (Klimaszewska et al. 2004), *P. pinaster* (Morel et al. 2014), *P. taeda* y *P. oocarpa* (Lara-Chavez et al. 2012), y se ha visto que se acumulan desde estados precotiledonarios a cotiledonarios (Lippert et al. 2005; Lara-Chavez et al. 2012). En *P. pinaster* hemos observado que la acumulación de estas proteínas en el estado cotiledonario del embrión somático es muy inferior al del embrión cigótico inmaduro.

Otras proteínas importantes durante la embriogénesis y que han sido propuestas como posibles marcadores del estado cotiledonario durante la maduración del embrión (Morel et al. 2014) son las proteínas relacionadas con la tolerancia a la desecación, entre las que destacan las proteínas LEA (del inglés *late embryogenesis abundant protein*s), y proteínas HSP relacionadas con el estrés térmico (del inglés *heat-shock proteins*). La menor acumulación de proteínas LEA en embriones somáticos sugiere que posiblemente estos no adquieran la resistencia a la deficiencia de agua de igual forma que los embriones cigóticos. Se ha demostrado que la

desecación parcial de los embriones somáticos tiene un efecto positivo durante la germinación en zanahoria (Tetteroo et al. 1995) y en algunas piceas (Lelu-Walter et al. 1995). Por lo tanto, teniendo en cuenta estas nuevas evidencias se podrán adecuar las nuevas condiciones de maduración en *P. pinaster* para modificar el contenido proteico de los embriones somáticos, conseguir que se asemeje más al de los embriones cigóticos y promover que adquieran mayor resistencia a la desecación.

Otra de las fases limitantes de la ES es la germinación, ya que las tasas de conversión en plantas siguen siendo muy bajas. En el trabajo realizado en el Artículo 1, se ensayaron diferentes fuentes nitrogenadas en el medio de germinación, para tratar de optimizar esta etapa y aumentar la viabilidad de las plántulas somáticas.

Actualmente no existen muchos trabajos relacionados con modificaciones del medio de germinación que supongan un incremento en la tasa de conversión en planta, y la mayoría de los resultados obtenidos al respecto se encuentran bajo patente. La composición del medio de germinación varía según la especie, siendo los medios más habituales el MLV en *P. Strobus* (Klimaszewska et al. 2001), los medios DCR y MLV en *P. pinaster* (Harvengt, 2005; Lelu-Walter et al. 2006; Miguel et al. 2004; Ramarosandratana et al. 2001a), el LV en *P. sylvestris* (Lelu-Walter et al. 1999) y el DCR en *P. nigra* (Salajová y Salaj, 2005).

Para hacer frente a condiciones de baja disponibilidad de nitrógeno (N), las plantas desarrollan diferentes mecanismos para un uso más eficiente del N (Good et al. 2004, Hermans et al. 2006, Nacry et al. 2013), entre ellos un mayor desarrollo radicular. Está descrito que el desarrollo de la raíz es dependiente, entre otros factores, de la composición del medio, del agua y de los nutrientes disponibles (Hodge et al. 2009; Lima et al. 2010), y que la disponibilidad de diferentes fuentes de nitrógeno y de carbono durante el desarrollo está directamente relacionada con una distribución diferenciada de biomasa entre la parte aérea y la raíz (Bauer et al. 2001; Bown et al. 2010).

La eficiencia en el uso del N puede dividirse principalmente en dos componentes, la eficiencia de adquisición o captación del N disponible, y por otro lado la eficiencia en la utilización de N. Está descrito que la captación

activa y la arquitectura de la raíz son los mejores indicadores de una eficiente adquisición de las fuentes nitrogenadas (Glass 2003; Garnett et al. 2009; Xu et al. 2012). Un mayor desarrollo de la raíz influye directamente en la capacidad que tienen las plantas para incorporar nutrientes nitrogenados. En este estudio, las plantas que fueron germinadas sólo con fuentes de N orgánico desarrollaron raíces completamente funcionales, capaces de captar los nutrientes del medio de una forma más eficiente durante su aclimatación.

Los análisis realizados en estas plantas revelan la expresión diferenciada de numerosos genes relacionados con procesos esenciales en el crecimiento y desarrollo debido a las diferentes formas de N disponibles en el medio. Los mecanismos reguladores que están involucrados en respuestas a cambios en la disponibilidad de N son numerosos, y sugieren la existencia de una red reguladora compleja que implica también una regulación hormonal (Kiba y Krapp, 2016). En los últimos años se ha progresado mucho en el estudio de los genes y rutas de señalización implicadas en la arquitectura de la raíz y captación de N en *Arabidopsis*. En esta especie, las auxinas tienen un papel fundamental en respuestas locales a nitrato y a limitación de N (Krouk et al. 2010; Ma et al. 2014).

De este modo, el estudio realizado en el Artículo 1 nos ha permitido optimizar la composición del medio, mejorar el desarrollo de la raíz y la capacidad adaptativa de las plántulas somáticas, aumentando la tasa de supervivencia durante la fase de aclimatación.

3. Estudio del metabolismo de la arginina y su regulación en las primeras etapas del desarrollo de *P. pinaster*

Las proteínas de reserva constituyen la fuente primaria de aminoácidos y N durante el desarrollo de las plantas (Shewry et al. 1995) y debido a que la arginina constituye una parte importante de estas proteínas de reserva, su biosíntesis se puede considerar como una de las rutas metabólicas más relevantes durante la embriogénesis somática en coníferas.

El estudio del metabolismo de la arginina ha sido parcialmente abordado en plantas, pero existe aún un gran desconocimiento sobre la

función de esta vía metabólica y la regulación molecular de las numerosas enzimas implicadas en la ruta.

La determinación de la compartimentación subcelular de la ruta, que se presenta en Artículo 3, ha sido determinante para entender cómo se produce la biosíntesis de arginina y poder considerar las vías de su posterior utilización metabólica. La ornitina, precursora de la síntesis de arginina, es sintetizada en el citosol a través de una ruta lineal, y a su vez es liberada en el cloroplasto a través de una ruta cíclica. Esto implica que la síntesis de ornitina tiene lugar en ambos compartimentos celulares; en el citosol, la ornitina será principalmente canalizada hacia la síntesis de poliaminas y otros compuestos nitrogenados, mientras que en los plastos será utilizada como precursor para la biosíntesis de arginina. Esta arginina sintetizada en los plastos tendrá que ser transportada al citosol para formar parte de las proteínas de reserva que se sintetizan durante la maduración y se acumulan en vesículas proteicas en este compartimento celular. Durante la germinación estos acúmulos proteicos se movilizan, y su degradación mediante la acción de enzimas hidrolíticas libera arginina, que junto a la directamente proveniente del plasto será degradada en las mitocondrias generando urea y glutamato. Estos compuestos se utilizarán metabólicamente para la síntesis de glutamina en el citosol. Los aminoácidos glutamina y glutamato son los donadores de nitrógeno para la biosíntesis de todos los compuestos nitrogenados de la plántula en desarrollo, en particular de los necesarios para la construcción de la maquinaria fotosintética que le permitirá el crecimiento autótrofo (Cánovas et al. 1998).

El estudio comparativo de expresión de todos los genes implicados en la biosíntesis de arginina en embriones somáticos y cigóticos sugiere que el metabolismo de la arginina no está bien regulado en los embriones somáticos. En general, los embriones cigóticos presentaron una mayor expresión de todos los genes de la ruta al final de la etapa de maduración, lo que implica un aumento en la biosíntesis de arginina y una mayor acumulación de proteínas de reserva. Por el contrario, en los embriones somáticos, son los genes que codifican la arginasa (PpARG) y la ornitina-δ-aminotransferasa (PpδOAT), enzimas relacionadas con el catabolismo de la

arginina, los que presentan mayores niveles de expresión. Estudios previos ya habían mostrado que los niveles de expresión de la arginasa eran elevados en embriones somáticos de pino durante la maduración (Pérez Rodríguez et al. 2006). En conjunto, estos resultados indican que la síntesis de arginina tiene lugar durante todo el proceso de ES, y que simultáneamente se está produciendo también su degradación, lo que sería consistente con la menor acumulación de proteínas de reserva observada en los embriones somáticos maduros.

La regulación del metabolismo de la arginina es clave para la acumulación y movilización de proteínas de reserva durante la embriogénesis y posterior germinación de *P. pinaster*. En este trabajo nos hemos centrado en caracterizar las proteínas tipo PII (Artículo 4), previamente descritas en bacterias y *Arabidopsis* como sensores del balance C/N y con un papel regulador en la biosíntesis de arginina (Heinrich et al. 2004; Mizuno et al. 2007; Chellamuthu et al. 2014).

La disponibilidad de recursos transcriptómicos nos ha permitido identificar dos secuencias que codifican proteínas tipo PII (PIIa y PIIb) en varias especies de pino: *P. pinaster, P. taeda, P. contorta y P. lambertiana.* La presencia de dos variantes de PII parece que está restringida en eucariotas a los pinos, ya que no se han encontrado en otras coníferas como las piceas, y en plantas herbáceas solo se ha identificado una forma de PII (Chen et al. 2006). En bacterias, fue descrita la primera proteína PII codificada por el gen glnB (van Heeswijk et al. 1993), y posteriormente se determinó la existencia de un segundo gen glnK, que codifica una segunda proteína PII, posiblemente generado por duplicación génica (Thomas et al. 2000).

En el Artículo 4 se han estudiado las características estructurales y funcionales de ambas formas de PII en *P. pinaster*, PpPIIa y PpPIIb, y los datos obtenidos sugieren que podrían corresponden a dos isoformas generadas por *splicing* alternativo. El modelo tridimensional de ambas proteínas es prácticamente idéntico, y muy similar a los descritos en otras especies como *Synechococcus* y *Arabidopsis* (Mizuno et al 2007; Llacer et al. 2008), a diferencia del lazo Q en la región C-terminal, que es exclusivo de

plantas (excepto para la familia *Brassicaceae* a la que pertenece *Arabidopsis*). En la estructura tridimensional de ambas proteínas se ha identificado un cambio significativo de un glutamato por una lisina en la posición 18, próximo al lazo T, que podría afectar en la interacción con el lazo Q. Debido a que PII forma complejos con otras proteínas utilizando el lazo flexible T como zona de interacción, la modificación de un residuo de este lazo también podría influir en la capacidad de formación de estos complejos.

En la caracterización de la interacción de ambas proteínas con PpNAGK se ha observado un incremento notable de su actividad específica. Sin embargo, se han encontrado diferencias de afinidad en la formación del complejo, ya que en el caso de PpPIIb se necesita menor cantidad de proteína para alcanzar el nivel de saturación de la actividad de PpNAGK. Se ha descrito que el efecto inhibitorio de la arginina sobre la actividad NAGK es menor cuando forma el complejo con proteínas PII y que la presencia de glutamina favorece la estabilidad del complejo (Chellamuthu et al. 2014). Los resultados obtenidos en este trabajo son consistentes con estas premisas, existiendo diferencias en la estabilidad del complejo formado entre cada isoforma y PpNAGK. A concentraciones bajas de arginina y en presencia de glutamina el complejo PpNAGK-PpPIIa presenta mayor actividad, la cual decae de forma significativa al incrementar la concentración de arginina. Mientras que el efecto estabilizador de la glutamina es mayor para el complejo PpNAGK-PpPIIb, el cual mantiene los niveles de actividad con concentraciones más altas de arginina. Estos resultados podrían indicar una mayor afinidad de PpPIIa por la glutamina, lo que sugiere un posible papel de esta proteína como sensor del estado nitrogenado en condiciones de menor disponibilidad de N. Esto podría justificar la presencia de dos proteínas PII funcionales en pino que podrían actuar como sensores en función de la disponibilidad de nitrógeno en el medio.

Los resultados recogidos en esta Tesis Doctoral aportan nuevos datos que ayudarán a comprender los mecanismos que regulan la biosíntesis de arginina en pino, en función de las condiciones nutricionales disponibles en el medio.

4. El transporte de arginina en *P. Pinaster*

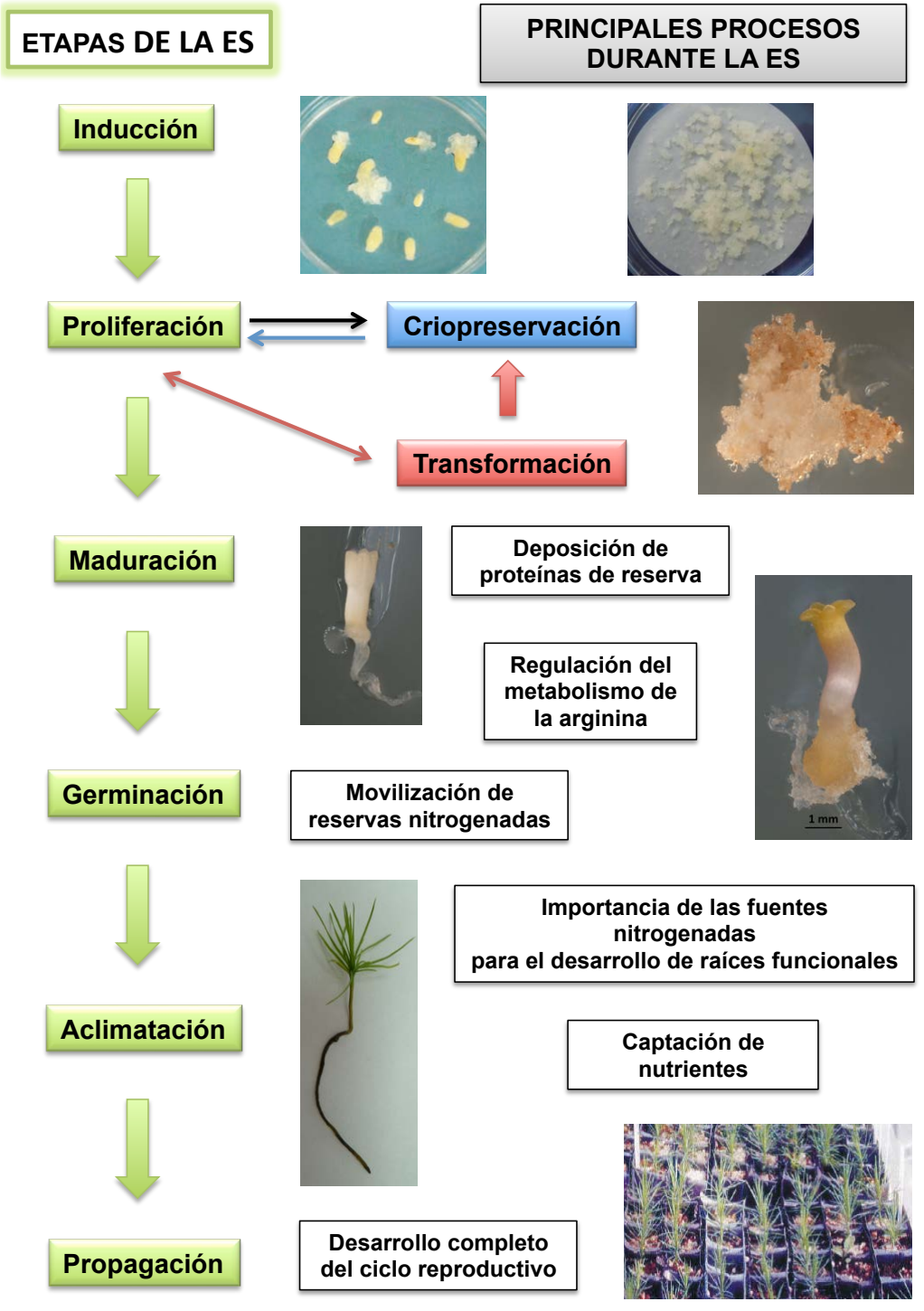
Otro punto clave que hemos tenido en cuenta para abordar el estudio del metabolismo de la arginina, ha sido tratar de esclarecer cómo se produce en pino la incorporación de los nutrientes nitrogenados durante el desarrollo. Como se ha demostrado en este trabajo, el metabolismo de la arginina en *P. pinaster* está distribuido en diferentes compartimentos celulares, lo que genera la necesidad de sistemas de transporte, tanto para la arginina, como para los intermediarios en las rutas de síntesis y degradación.

El transporte a larga distancia de arginina hacia los órganos de almacenamiento de N o semillas tiene lugar a través del sistema vascular, y es dependiente de transportadores de aminoácidos pertenecientes a la familia de permeasas de aminoácidos (AAP). En *Arabidopsis,* se ha determinado que AtAAP5 es especialmente importante en el transporte a larga distancia de arginina, ya que además de ser un transportador de alta afinidad para arginina y lisina (Svennerstam et al. 2008), desempeña un papel clave en la absorción de aminoácidos básicos en la raíz (Svennerstam et al. 2011). En coníferas, no se ha descrito por el momento ningún transportador implicado en la absorción y transporte de arginina.

En el trabajo recogido en el Artículo 5 de este manuscrito, hemos identificado diez miembros pertenecientes a la familia de transportadores AAP*,* y hemos caracterizado funcionalmente *PpAAP1*, pues es el que presenta mayor similitud de secuencia con *AtAAP5* de *Arabidopsis* y además se expresa en tejidos en desarrollo. La expresión de *PpAAP1* aumenta notablemente en el ápice de la raíz de plántulas de pino cuando se suministra arginina externa en el medio, lo que sugiere que podría desempeñar un papel como sensor de este aminoácido y, por tanto, podría estar relacionado con su movilización durante el desarrollo. Los estudios de complementación con *PpAAP1* realizados en levadura permitieron determinar que se trata de un transportador de arginina que funciona en el rango de concentración micromolar, que es en el que normalmente se encuentran los aminoácidos disponibles en el medio (Öhlund, 2004).

Líneas mutantes de AtAAP5 presentan una capacidad de absorción de arginina muy reducida respecto a plantas controles (Svennerstam et al. 2011) presentando un fenotipo claro cuando crecen con altos niveles de arginina en el medio, y produciendo más biomasa que las plantas silvestres. La complementación de estas líneas mutantes con *PpAAP1* genera plantas con una reducción drástica de su desarrollo y con un sistema radicular reducido, lo que indica la recuperación de la capacidad transportadora de arginina en estas plantas.

Será necesario realizar estudios adicionales para determinar funciones complementarias que puedan atribuirse a este transportador, así como continuar investigando las funciones y características de los restantes miembros de esta familia AAP en *P. pinaster*.



Esquema representativo de las etapas de desarrollo de la ES en coníferas y los principales procesos que tienen lugar en las primeras etapas de desarrollo

5. Bibliografía

Álvarez JM, Cortizo M, Ordás RJ (2012). Characterization of a type-A response regulator differentially expressed during adventitious caulogenesis in *Pinus pinaster. Journal of Plant Physiology* 169: 1807-1814.

Bauer GA, Berntson GM (2001). Ammonium and nitrate acquisition by plants in response to elevated CO_2 concentration: The roles of root physiology and architecture. *Tree Physiol* 21:137-144.

Bown HE, Watt MS, Clinton PW, Mason EG (2010). Influence of ammonium and nitrate supply on growth, dry matter partitioning, N uptake and photosynthetic capacity of *Pinus radiata* seedlings. *Trees* 24:1097-1107.

Cánovas FM, Cantón FR, García-Gutiérrez A, Gallardo F, Crespillo R (1998). Molecular physiology of glutamine and glutamate biosynthesis in developing seedlings of conifers. *Physiol. Plantarum* 103, 287-294 .

Chellamuthu VR, Ermilova E, Lapina T, Luddecke J, Minaeva E, Herrmann C, Hartmann MD, Forchhammer K (2014). A widespread glutamine sensing mechanism in the plant kingdom. *Cell* 159: 1188-1199.

Chen YM, Ferrar TS, Lohmeier-Vogel EM, Morrice N,cMizuno Y, Berenger B, Ng KKS, Muench DG, Moorhead GBG (2006). The PII signal transduction protein of Arabidopsis thaliana forms an arginine regulated complex with plastid N-acetyl glutamate kinase. *Journal of Biological Chemistry* 281: 5726-5733.

Garnett T, Conn V, Kaiser BN (2009). Root based approaches to improving nitrogen use efficiency in plants. *Plant, Cell & Environment,* 32: 1272-1283.

Glass (2003). Nitrogen use efficiency of crop plants:Physiological constraints upon nitrogen absorption. *Critical Reviews in Plant Sciences* 22(5):453-470.

Good AG, Shrawat AK, and Muench DG (2004) Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? *Trends Plant Sci.* 9, 597-605.

Harvengt L (2005). Somatic embryogenesis in maritime pine (*Pinus pinaster* Ait). *Forest. Sci.* 77:107-119.

Heinrich A, Maheswaran M, Ruppert U, Forchhammer K (2004). The *Synechococcus elongatus* PII signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. *Mol. Microbiol.* 52: 1303-1314.

Hermans C, Hammond JP, White PJ, Verbruggen N (2006). How do plants respond to nutrient shortage by biomass allocation? *Trends Plant Sci.* 11(12):610-7.

Hodge A, Berta G, Doussan C, Merchan F, Crespi M (2009). Plant root growth, architecture and function. *Plant Soil* 321:153-187.

Kiba T, Krapp A (2016) Plant nitrogen acquisition under low availability: regulation of uptake and root architecture. Plant Cell Physiol 57: 707-714.

Klimaszewska K, Park Y-S, Overton C, Maceacheron I, Bonga JM (2001) Optimized somatic embryogenesis in *Pinus strobus* L. In Vitro Cell Dev Biol-Plant 37:392-399.

Klimaszewska K, Morency F, Jones-Overton C, Cooke J (2004) Accumulation pattern and identification of seed storage proteins in zygotic embryos of *Pinus strobus* and in somatic embryos from different maturation treatments. Physiol Plant 121:682-690.

Klimaszewska K, Trontin JF, Becwar MR, Devillard C, Park Y-S, Lelu-Walter MA (2007). Recent progress in somatic embryogenesis of four *Pinus* spp. En: da Silva J, Shima K (eds). *Tree and Forestry Science and Biotechnology Global Science Books*, 1(1): 11-25.

Krouk G, Lacombe B, Bielach A, Perrine-Walker F, Malinska K, Mounier E, Hoyerova K, Tillard P, Leon S, Ljung K, Zazimalova E, Benkova E, Nacry P, Gojon A (2010). Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. *Developmental Cell*, 18 (6): 927-937.

Lara-Chavez A, Egertsdotter U, Flinn BS (2012) Comparison of gene expression markers during zygotic and somatic embryogenesis in pine. In Vitro Cell Dev-Pl 48:341-354.

Lelu-Walter MA, Klimaszewska K, Pflaum G, Bastien C (1995). Effect of maturation duration on desiccation tolerance in hybrid larch (Larix X leptoeuropaea Dengler) somatic embryos. *In Vitro Cellular & Developmental Biology . Plant* 31(1):15-20.

Lelu-Walter MA, Bastien C, Drugeault A, Gouez ML, Klimaszewska K (1999). Somatic embryogenesis and plantlet development in *Pinus sylvestris* and *Pinus pinaster* on medium with and without growth regulators. *Physiologia Plantarum*. 105: 719-728.

Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2006) Simplified and improved somatic embryogenesis for clonal propagation of *Pinus pinaster* (Ait.). *Plant Cell Rep* 25: 767-776.

Lima JE, Kojima S, Takahashi H, von Wirén N (2010). Ammonium triggers lateral root branching in Arabidopsis in an AMMONIUM TRANSPORTER1;3-dependent manner. *The Plant Cell* 22:3621-3633.

Lippert D, Zhuang J, Ralph S, Ellis DE, Gilbert M, Olafson R, Ritland K, Ellis B, Douglas CJ, Bohlmann J (2005). Proteome analysis of early somatic embryogenesis in *Picea glauca. Proteomics 5,* 461-473.

Llacer JL, Fita I, Rubio V (2008). Arginine and nitrogen storage. *Curr Opin Struct Biol 18*: 673-681.

Ma W, Li J, Qu B, He X, Zhao X, Li B, Fu X, Tong Y (2014). Auxin biosynthetic gene *TAR2* is involved in low nitrogen-mediated reprogramming of root architecture in Arabidopsis. The Plant Journal, 78: 70-79.

Miguel C, Gonc¸alves S, Tereso S, Marum L, Oliveira MM (2004). Somatic embryogenesis from 20 open-pollinated seed families of Portuguese plus trees of maritime pine. Plant Cell Tissue Org Cult. 76:121-130.

Mizuno Y, Moorhead GB, Ng KK (2007). Structural basis for the regulation of *N*-acetylglutamate kinase by PII in *Arabidopsis thaliana*. *J. Biol. Chem.* 282, 35733-35740.

Morel A, Trontin JF, Corbineau F, Lomenech AM, Beaufour M, Reymond I, Le Metté C, Ader K, Harvengt L, Cadene M, Label P, Teyssier C, Lelu-Walter MA (2014). Cotyledonary somatic embryos of *Pinus pinaster* Ait. most closely resemble fresh, maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses. Planta 240 (5): 1075-95.

Nacry P, Bouguyon E, Gojon A (2013). Nitrogen acquisition by roots: physiological and developmental mechanisms ensuring plant adaptation to a fluctuating resource. *Plant Soil* 370: 1-29.

Öhlund J and Näsholm T (2004). Regulation of organic and inorganic nitrogen uptake in Scots pine (*Pinus sylvestris*) seedlings. *Tree Physiol.* 24:1397-402.

Pérez-Rodríguez J, Suárez MF, Heredia R, Ávila C, Breton D, Trontin, JF, Filonova L, Bozhkov P, von Arnold S, Harvength L, Cánovas FM (2006). Expression patterns of two glutamine synthetase genes in zygotic and somatic pine embryos support specific roles in nitrogen metabolism during embryogenesis. *New Phytol* 169: 35-44.

Ramarosandratana A, Harvengt L, Bouvet A et al. (2001). Influence of the embryonal-suspensor mass (ESM) sampling on development and proliferation of maritime pine somatic embryos. *Plant Sci* 160:473-479.

Salajová T, Salaj J (1992). Somatic embryogenesis in European black pine (Pinus nigra Arn.). *Biologia Plantarum* 34: 213-218.

Shewry PR, Napier JA, Tatham AS (1995). Seed storage proteins: structures and biosynthesis. *The Plant Cell,* 7: 945-956.

Svennerstam H, Ganeteg U, Näsholm T (2008) Root uptake of cationic amino acids by Arabidopsis depends on functional expression of amino acid permease. *New Phytologist* 180: 620-630.

Svennerstam H, Jämtgård S, Ahmad I, Huss-Danell K, Näsholm T, Ganeteg U (2011) Transporters in Arabidopsis roots mediating uptake of amino acids at naturally occurring concentrations. *New Phytologist* 191: 459-467.

Trontin JF (2007). Recent progress in genetic transformation of four Pinus spp. *Transgenic Plant Journal* 1: 314 329.

Trontin JF, Klimaszewska K, Morel A, Hargreaves C, Lelu-Walter M-A (2016). Molecular aspects of conifer zygotic and somatic embryo development: a review of genome-wide approaches and recent insights. In: Germanà MA, Lambardi M (eds) *In Vitro Embryogenesis in Higher Plants*, Methods in Molecular Biology, vol. 1359, Chapter 8, Springer Science+Business Media New York, pp. 167-207.

Tetteroo FAA, Hoekstra FA, Karssen CM (1995). Induction of complete desiccation tolerance in carrot (*Daucus carota*) embryoids. *Journal of Plant Physiology* 145: 349–356.

Thomas G, Coutts G, Merrick M (2000). The *glnKamtB* operon: a conserved gene pair in prokaryotes. *Trends Genet* 16: 11-14.

Xu G, Fan X, Miller AJ (2012) Plant nitrogen assimilation and use efficiency. *Annu Rev Plant Biol* 63: 153-182.

V. CONCLUSIONES

1. La embriogénesis somática es una herramienta biotecnológica de gran utilidad para el análisis funcional de genes esenciales de coníferas mediante transformación genética.

2. El estudio de la economía nitrogenada durante las primeras etapas del desarrollo embrionario en pino es fundamental para comprender la formación del embrión. Los estudios realizados ponen de manifiesto la importancia de las distintas fuentes de nitrógeno para la obtención de plántulas somáticas viables.

3. Las proteínas de reserva acumuladas durante la embriogénesis de *P. pinaster* se consideran marcadores del estado de maduración de los embriones. La mayor acumulación de proteínas de reserva en embriones cigóticos podría determinar la mayor tasa de germinación observada respecto a los embriones somáticos.

4. La desregulación del metabolismo de la arginina detectada en los embriones somáticos podría ser la causa principal de la menor acumulación de proteínas de reserva.

5. El estudio de las proteínas PII en pino es importante para comprender la regulación del metabolismo de la arginina. La caracterización del complejo que forman las dos isoformas de PII con su diana NAGK revela diferencias de afinidad por la glutamina, y sugiere un papel complementario de ambas isoformas como sensores de N en función de la disponibilidad de este nutriente en el medio.

6. Se han identificado los miembros de la familia de permeasas de aminoácidos en pino. Los estudios funcionales de la permeasa PpAAP1 en levaduras y en plantas mutantes (*aap5*) han demostrado su capacidad para transportar arginina del medio. Por tanto, PpAAP1 es el primer transportador de arginina identificado en coníferas.

Conclusión final: Los estudios recogidos en esta Tesis Doctoral ponen de manifiesto la importancia del nitrógeno y del metabolismo de la arginina durante la embriogénesis de coníferas. Los resultados obtenidos contribuyen a una mejor comprensión de los mecanismos moleculares que regulan el desarrollo del embrión y además, proporcionan nuevas vías para el perfeccionamiento de métodos biotecnológicos de producción forestal vía embriogénesis somática.