



## **Joint Doctoral Thesis (PhD)**

Dissertation to be presented at University of Talca, to obtain the degree of :  
“Docteur en Biologie Cellulaire et Moléculaire” Bordeaux1 University

and

«Doctorado en Ciencias mención Ingeniería Genética Vegetal”, Universidad de Talca

by

**Marcelo Arnoldo Garcés Cea**

**PROTEOMIC STUDY OF WOOD FORMATION IN MARITIME PINE**

November, 2008

Examination commission:

**Phillipe Label**, Research Director, INRA, UR AGPF, Orleans, France

**Sofía Valenzuela**, Research Director, Biotechnology Center, University of Concepción, Chile

**Raúl Herrera**, Associate Profesor, University of Talca, Talca Chile (co-supervisor)

**Christophe Plomion**, Research Director, INRA, UMR Biogeco, Bordeaux, France (supervisor)

## Résumé

Les propriétés du bois de pin maritime varient aux niveaux chimique, anatomique et mécanique. Six types de bois peuvent être trouvés au sein d'un même arbre : bois précoce, bois tardif, bois de couronne, bois de base, bois de compression et bois opposé.

Au cours de cette thèse, nous avons testé l'hypothèse selon laquelle la variabilité phénotypique des propriétés de bois, serait liée à l'expression différentielle des protéines lors de la xylogénèse.

Par une approche protéomique basée sur l'électrophorèse bidimensionnelle et la spectrométrie de masse en tandem (LC ESI MS/MS), nous avons identifié 165 protéines différentiellement exprimées le long d'un gradient d'âge cambial (bois juvénile vs. bois mature) ainsi que 93 protéines différentiellement exprimées au cours de la saison de végétation (bois de printemps vs. bois d'été) chez le pin maritime. Une analyse chimique complémentaire des échantillons a été réalisée par pyrolyse analytique.

Nos résultats montrent que le xylème secondaire formé en début de saison ainsi que celui qui est initié par un cambium jeune présentent une sur-expression de protéines participant à la division cellulaire. Dans le xylème issu d'un cambium âgé ou formé à la fin de l'été nous avons mis en évidence des protéines impliquées dans la défense cellulaire (dont le rôle serait de retarder la mort cellulaire programmée), ainsi que des protéines impliqués dans la biosynthèse des éléments constitutifs de la paroi. Cette étude contribue à renforcer nos connaissances sur les acteurs moléculaires intervenant lors de la xylogénèse. Elle ouvre par ailleurs des pistes de recherche sur la détection de gènes impliqués dans le contrôle génétique des propriétés du bois dans un objectif de sélection assisté par marqueurs.

### Mots-clé:

Protéomique, Formation du bois, Pin maritime, *Pinus pinaster*, xylème, cambium

## **Abstract**

Wood properties in maritime pine are highly variable at chemical, anatomical and mechanical levels. Six types of wood can be found in a single tree, early wood, late wood, crown wood, base wood, compression wood and opposite wood.

In this thesis report, we tested the hypothesis that the observed variability at the phenotypic level, can be bound to the differential expression of proteins during the process of wood formation.

We use the tools of proteomics, Bidimensional electrophoresis and LC ESI MS/MS for the discovery of 165 proteins differentially expressed in a cambial age gradient, (from base wood to crown wood), an 93 overexpressed proteins in a seasonal gradient (from early wood collected at the beginning of the growing season, to late wood, collected at summer)

Complementary, chemical characterization of the samples was performed using analytical pyrolysis.

Our results showed that the secondary xylem formed at the beginning of the growing season, and the xylem formed by a young cambium, present a overexpression of proteins participating in the intense cell division, characteristic of those tissues, e.g. Biogenesis of cytoskeleton and hemicelluloses, RNA transcription, synthesis, folding and modification of proteins. In the xylem formed at the base of the trunk and at the end of the growing season we have found an over-expression of proteins from cell defense (they role will be to delay programmed cell death) and cell wall formation related proteins e.g. lignin biosynthesis. This study contributes to reinforce our knowledge over the molecular actors involved in the xylogenesis process. It opens , in another hand , research guides for the detection of genes involved in the genetic control of wood properties towards an objective of marker assisted selection.

## **Keywords:**

Proteomics, Wood formation, maritime pine, *Pinus pinaster*, xylem, cambium

## **Resumen**

Las propiedades de la madera en pino marítimo varían tanto a nivel químico, anatómico y mecánico. Seis tipos de madera se pueden encontrar al seno de un mismo árbol: madera temprana, madera tardía, madera de copa, madera de base, madera de compresión y madera opuesta.

En este trabajo de tesis, hemos testeado la hipótesis según la cual la variabilidad fenotípica de las propiedades de la madera, están liadas a la expresión diferencial de proteínas a través de la xilogénesis.

Por una aproximación proteómica, basada en la electroforesis bidimensional y la espectrometría de masas en tándem (LC ESI MS/MS), hemos identificado 165 proteínas diferencialmente expresadas a lo largo de un gradiente de edad cambial ( madera juvenil y madera madura) y 93 proteínas diferencialmente expresadas en el curso de una temporada de crecimiento (madera de primavera y madera de verano) en pino marítimo . Un análisis químico complementario se realizó a las muestras por pirolisis analítica.

Nuestros resultados mostraron que el xilema secundario formado al inicio de la temporada así como el formado por un cambium joven presentan una sobre expresión de proteínas que participan en la división celular. En el xilema formado a partir de un cambium maduro o al fin de la temporada hemos descubierto proteínas implicadas en la defensa celular ( su rol sería evitar la muerte celular programada), y proteínas implicadas en la biosíntesis de elementos constitutivos de la pared. Este estudio contribuye a reforzar nuestros conocimientos sobre las actividades moleculares que intervienen durante la xilogénesis. Se descubren por otro lado, pistas de investigación a seguir sobre la detección de genes implicados en el control genético de las propiedades de la madera en un objetivo de selección asistida por marcadores

## **Agradecimientos**

Esta investigación fue financiada por el proyecto GNP05013C Génoplante, de la Region de Aquitania, Francia, además agradezco el apoyo del programa ALFA II-0266-FA (GEMA, Unión Europea) por el financiamiento de mi estadía y al INRA-UMR Biogeco. Gracias a la Beca de movilidad del ministerio encargado de la investigación de Francia, que me permitió asistir a congresos. Por ultimo gracias a la Universidad de Talca por la beca de postgrado, mantenida desde el 2004.

2.2.9 Variaciones en la Formación de madera	19
2.2.9.1 Variación durante la temporada de crecimiento	
2.2.9.2 Variación ontogénica : madera de cima/juvenil vs. madera de base/madura.	
2.2.10 Avances en Genómica de Especies Forestales	22
2.2.10.1 Secuencias de Genomas	
2.2.10.2 Secuencias de EST	
2.2.10.3 Estudios de expresión	
2.2.10.4 Proteómica	
2.3 Planteamiento del Problema	28
2.4 Formulación de Hipótesis	28
2.5 Objetivos Generales y Específicos	29
<b>3. Materiales y Métodos</b>	<b>30</b>
3.1 Diseño general de los experimentos.	31
3.1.1 Gradiente de edad cambial 2003	
3.1.2 Gradiente de edad cambial 2005	
3.1.3 Gradiente de edad cambial 2006	
3.1.4 Gradiente de Temporada 2006	
3.2 Técnica de recolección de tejido formador de madera	33
3.3 Procedimientos	33
3.3.1 Extracción de proteínas	
3.3.2 Electroforesis bidimensional	
3.3.2.1 Primera Dimensión: Isoelectroenfoco con el sistema IPGphor: migración según el pI	34

## Índice

### 1. Resumen

Abstract

Resumen

Índice I

### 2. Introducción 1

2.1. Formulación del marco teórico 2

2.2 Antecedentes Bibliográficos 7

2.2.1 Importancia de la actividad forestal en Aquitania 7

2.2.2 El pino marítimo 7

2.2.2.1 Descripción biológica y hábitat

2.2.2.2 Taxonomía

2.2.2.3 Repartición geográfica

2.2.3 Programa de mejoramiento INRA 11

2.2.4 Nuevas propiedades de la madera son necesarias 13

2.2.5 Calidad de la madera 14

2.2.6 El Cambium Vascular 14

2.2.7 Diferenciación del Xilema 14

2.2.8 La formación de madera 15

3.3.2.2	Equilibración de strips (cintas)	37
3.3.2.3	Segunda Dimensión: Electroforesis denaturante en gel de poliacrilamida. Migración según el peso molecular.	38
3.3.2.4	Condiciones de migración.	39
3.3.2.5	Teñido de geles	40
3.3.3	Adquisición de imágenes.	
3.3.4	Detección de spots (manchas)	
3.3.5	Tratamientos Estadísticos	41
3.3.5.1	ANOVA	42
3.3.5.2	Clustering	43
3.3.6	Identificación de proteínas por espectrometría de masas.	43
3.3.6.1	Digestión de proteínas en el gel.	
3.3.6.2	Análisis capilar HPLC nanoespray trampa de iones MS/MS	44
3.3.7	PCR en tiempo real (qPCR)	47
3.3.8	Caracterización Química	48
3.3.8.1	Preparación de muestras	49
3.3.8.2	Análisis de pared celular	
3.3.8.2.1	FTIR Espectroscopia infrarroja con Transformada de Fourier	
3.3.8.2.2	Pirólisis Analítica	50
<b>4.</b>	<b>Resultados y Discusión.</b>	<b>51</b>
4.1	Análisis fenotípico y molecular de un gradiente base a copa	52
4.1.1	Caracterización anatómica	53
4.1.2	Comparación de datos ecofisiológicos, para las temporadas 2003, 2005 y 2006	
4.1.3	Composición química de la pared celular en un gradiente	

de base a corona.	57
4.1.4 Análisis molecular y fenotípico de un gradiente base a copa en tejido formador de madera en pino marítimo. Temporada 2003	59
4.1.5 Análisis proteómico de un gradiente base a copa temporadas 2005 y 2006.	80
4.1.5.1 Mapas proteicos bidimensionales de tejido formador de madera, gradiente desde base a copa 2005.	
4.1.5.2 Identificación de proteínas por espectrometría de masas	81
4.1.6 Mapas proteicos bidimensionales de tejido formador de madera. Gradiente base a copa 2006.	86
4.1.8 Expresión de genes y proteínas a lo largo del gradiente base a copa	
Genes y Proteínas relacionados con síntesis de pared celular.	92
La prolongación del depósito de pared celular secundaria en tejido formador de madera de base, involucra la sobreexpresión de genes relacionados con defensa celular.	95
Genes relacionados con “síntesis de proteínas” y “energía” son sobreexpresados en madera formación de copa.	97
Isozimas diferentes son recrutadas en madera de copa/juvenil o de base/madura.	99
4.1.9 Registro de datos proteómicos en base de datos PROTIC.	99
4.2 Análisis Fenotípico y Proteómico del Gradiente de temporada 2006	101
4.2.2 Variabilidad de la composición química de la pared celular a lo largo de la temporada	103
4.2.2.1 Análisis de Infrarrojo con transformada de Fourier (FTIR)	
4.2.2.2 Pirólisis analítica.	105
Análisis de componente principal.	107

4.2.3 Mapas bidimensionales de tejido formador de madera.	108
4.2.4 Identificación de proteínas en tejido formador de madera y clasificación funcional.	109
4.2.4.1 Proteínas sobrepresadas en Madera Temprana o de Primavera.	115
4.2.4.2 Proteínas sobrepresadas en Madera Tardía o de verano.	118
4.2.5 Discusión de los resultados del gradiente de temporada	121
4.2.6 Artículo “A Combined Proteome and Transcriptome analysis in a seasonal and base to crown gradient of wood forming tissue in maritime pine”.	123
4.3 Discusión Final	124
<b>5. Conclusiones</b>	<b>128</b>
<b>6. Referencias Bibliográficas</b>	<b>129</b>
<b>7. Anexos</b>	

## 2. INTRODUCCION

## **2.1 Formulación del marco teórico**

Actualmente en la región de Aquitania, en el sur oeste de Francia, la gestión de manejo forestal tiene como objetivo el mantener la sustentabilidad del bosque, tanto sobre el plano ecológico, como económico. En los últimos 40 años, la optimización de métodos silvícolas y la introducción de variedades mejoradas han permitido aumentar considerablemente la productividad y homogeneidad de la madera. Por ejemplo, en el programa de mejoramiento tradicional de Pino marítimo iniciado en los años sesenta por el INRA UMR Biogeco, se han obtenido ganancias genéticas de un 30% en rendimiento y rectitud del tronco (Raffin and Pastuszka, 2002). Las condiciones actuales de cultivo, de crecimiento acelerado han provocado una disminución en la edad de rotación (de 60 a 40 años), lo cual ha producido un aumento en la proporción de madera juvenil. Estudios relativos a este cambio han demostrado una correlación negativa entre productividad y calidad de la madera (Rozenberg and Cahalan, 1997; Pot et al., 2002) y se ha llegado a decir que uno de los problemas principales de la industria forestal a una escala mundial, es el exceso de inventario de madera juvenil en pinos (Zobel, 1984). Junto con la reducción en la calidad de la madera, el anunciado fenómeno del cambio climático propone nuevos desafíos a la industria forestal, ya que se esperan transformaciones significativas del ambiente y con ello nuevas condiciones de estrés biótico y abiótico.

Con el objetivo de asegurar una materia prima de calidad a los distintos actores de la industria forestal, maderera y papelera por medio de un recurso genético mejorado, al final de los años 90 se inició en INRA UMR Biogeco un programa multidisciplinario de investigación sobre el determinismo genético y molecular de la calidad de madera en pino marítimo. La estrategia seguida involucra el uso de diferentes fuentes de información: información funcional de especies modelo (Le Provost, 2003), patrones de diversidad de genes candidatos (Dantec et al., 2004; Pot et al., 2005), comparación de QTLs (Chagné et al., 2003), plasticidad molecular a nivel transcripcional y traduccional (Le Provost et al., 2003; Le Provost, 2003; Paiva, 2006; Paiva et al., 2008), buscando identificar genes potencialmente involucrados en la calidad de la madera (interés económico) y caracteres de adaptación al estrés hídrico (interés ecológico) (Moreau, 2005; Eveno et al., 2007), que puedan servir de diagnóstico para la mejor administración y explotación de los recursos forestales. En paralelo y siguiendo el objetivo de proveer a los seleccionadores de herramientas para evaluar la diversidad genética y su

potencial de evolución, se han iniciado estudios de asociación, para verificar los genes candidatos en poblaciones no relacionadas.

El gran tamaño del genoma del pino marítimo (51pg comparado con 0,3pg de arabisopsis y 6pg del humano) y su complejidad, ha influido en la elección de una estrategia de identificación de genes candidatos en INRA UMR Biogeco. Los genes candidatos pueden ser definidos como genes con polimorfismos moleculares asociados estadísticamente a características fenotípicas, por ejemplo, de importancia económica o también como genes asociados a elementos de expresión cuantitativa (QTLs) con una posición definida en el genoma (Pflieger et al., 2001). En la identificación de genes candidatos se ha definido en tres etapas:

- 1) Elección de genes candidatos. Los genes candidatos son elegidos en base a tres criterios: i) Analizando enzimas que participan en mecanismos moleculares o fisiológicos conocidos (por ej., genes involucrados en lignificación), ii) Analizando la colocalización de genes candidatos con QTLs (genes candidatos posicionales) o resultando de estudios de expresión, comparando condiciones contrastantes o estados de desarrollo (genes candidatos expresionales).
- 2) Screening (detección precoz) de genes candidatos: Se realiza por medio de estudios de asociación, colocalizando estadísticamente polimorfismos de genes candidatos con variaciones fenotípicas o QTLs, en un grupo de individuos genealógicamente no relacionados.
- 3) Validación de genes candidatos: Se realiza mediante estudios de genética reversa (transformación genética) o mapeo comparativo de QTLs, para confirmar que el gen candidato efectivamente esta relacionado con el carácter en estudio.

Este trabajo de tesis se inscribe dentro de la primera etapa de búsqueda de genes candidatos expresionales: la identificación de genes/proteínas expresados en la xilogénesis de pino marítimo.

La madera es un material variable tanto en el plano químico, anatómico y en sus propiedades físicas. Junto a la variación genética, el ambiente y la edad del cambium afectan igualmente las propiedades de la madera. En este contexto probaremos la hipótesis que esta variabilidad puede estar ligada a la expresión diferencial de proteínas durante la formación de madera.

En los trabajos pioneros de Le Provost (2003) y posteriormente de Paiva (2006) se abordó la caracterización del transcriptoma expresado durante la formación de madera en pino marítimo. Brevemente, se realizó una librería de expresión de cDNA de xilema secundario en diferenciación, para posteriormente abordar su secuenciación y análisis, finalizando con la anotación de 10.000 ESTs. Posteriormente, mediante análisis transcriptómicos, se identificó la expresión diferencial de ESTs analizando dos tipos de variables: i) ambientales (efecto temporada y gravitacional) y ii) de desarrollo (gradiente base a copa). Las técnicas utilizadas incluyeron cDNA-AFLP, Northern reverso, RT-PCR, RT-PCR en tiempo real, y macroarreglos. Igualmente en esta etapa del estudio se realizó un análisis global del proteoma expresado en la xilogénesis (Gion et al., 2005), concluyendo que estudios proteómicos más detallados deberían ser abordados para identificar un mayor número de proteínas expresadas y así, tratar de dilucidar los metabolismos involucrados en la formación de los distintos tipos de madera.

Un esquema de la estrategia de identificación de genes candidatos expresionales por técnicas transcriptómicas y proteómicas se presenta en la **Figura 1**, en la parte superior se ilustra el trabajo realizado en transcriptómica, donde se ha comparado la expresión de genes en un gradiente de temporada y en un gradiente de edad cambial, asimismo se realizó la identificación de genes expresados en distintos tejidos: polen, acículas, floema, raíces, brotes y específicamente en xilema. Esto llevó a la identificación de una lista de genes candidatos expresionales en pino marítimo (Paiva, 2006).

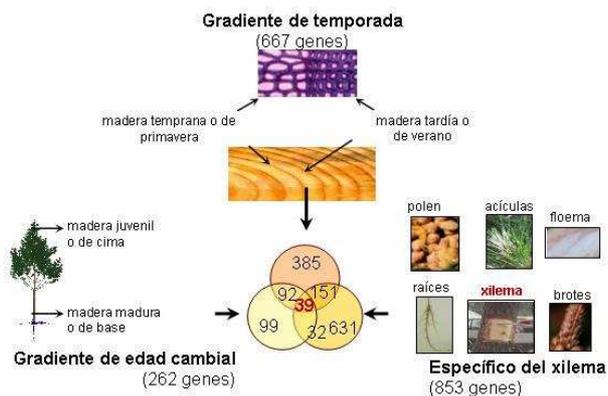
El objetivo de esta tesis, se ilustra en la región inferior de la **Figura 1** y consiste en identificar el proteoma expresado en la xilogénesis, analizando una gradiente de temporada, desde la primavera a verano y una gradiente de edad cambial, desde la base del árbol a la copa. Los estudios proteómicos fueron complementados con análisis químicos.

La organización de este reporte se compone de una introducción, en la revisión bibliográfica se describe la importancia de la madera para la región de Aquitania, las características botánicas del pino marítimo, el programa de mejoramiento del INRA, nuevas propiedades de la madera a ser incorporadas a este programa, la calidad de la madera, el proceso de xilogénesis y variaciones en la formación de madera, finalmente se describió el estado del arte en la investigación molecular de la formación de madera en coníferas.

## ELECCION DE GENES CANDIDATOS EXPRESIONALES

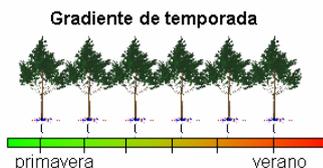
### TRANSCRIPTOMICA

Análisis transcriptómico de 3500 genes, asociados a diferentes tipos de madera y diferentes tejidos (Le provost, 2001; Paiva, 2007)

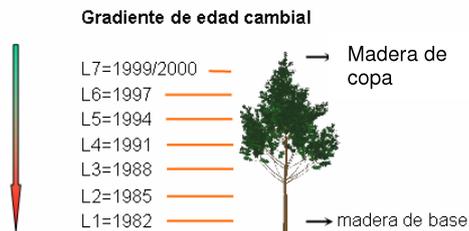


### PROTEOMICA

Dos genotipos, 8 muestras tomadas cada 10 días desde la primavera a verano.



Dos genotipos, 4 muestras tomadas en internodos desde la base a la cima del árbol



**Figura 1:** En la parte superior del esquema, se detallan los experimentos de transcriptómica que ayudaron a la identificación de genes candidatos expresionales. En la parte inferior, se detallan los experimentos de proteómica realizados en esta tesis.

En la sección materiales y métodos se encuentran los diseños experimentales usados y una explicación detallada de los procedimientos utilizados para el análisis proteómico de la madera en formación, así como del tratamiento estadístico de los datos.

En la sección resultados, se agrupan los resultados obtenidos por temporada.

1. **Análisis de un gradiente de base a copa en la temporada 2003.** Se realizó el análisis proteómico de xilema en formación en 4 niveles a través de un gradiente base a copa, estos datos fueron integrados con análisis transcriptómicos y químicos realizados por Paiva y colaboradores, los resultados fueron publicados en *New Phytologist* **178 (2), 283–301(2008)**. *Jorge A.P. Paiva, Marcelo Garcés, Ana Alves, Pauline Garnier-Géré, José Carlos Rodrigues, Céline Lalanne, Stéphane Porcon, Grégoire Le Provost, Denilson da Silva Perez, Jean Brach, Jean-Marc Frigerio, Stéphane Claverol, Aurélien Barré, Pedro Fevereiro, Christophe Plomion. Molecular and phenotypic profiling from the base to the crown in maritime pine wood-forming tissue doi:10.1111/j.1469-8137.2008.02379.x*
2. **Análisis proteómico de un gradiente de base a copa para la temporada 2005 y 2006.** En este capítulo se realizaron análisis proteómicos para un árbol en un gradiente base a copa en la temporada 2005 y análisis proteómicos para dos clones en la temporada 2006, utilizando muestras contrastantes de xilema en formación.
3. **Análisis proteómico y químico de un gradiente de temporada 2006**  
En este capítulo se analizó un gradiente de temporada, utilizando dos genotipos e integrando análisis proteómicos, químicos y Ecofisiológicos para la temporada 2006. La integración de estos datos nos permitió obtener interesantes conclusiones acerca de las categorías funcionales que participan en la formación de madera a lo largo de la temporada.

Por último, se encuentra el capítulo de conclusiones y perspectivas futuras. En la sección anexos se incluyó un apéndice técnico detallado con los protocolos y análisis estadísticos utilizados.

## **2.2 Antecedentes Bibliográficos**

### **2.2.1. Importancia de la actividad forestal en Aquitania**

La superficie total de bosques en Francia comprende a 15,2 millones de hectáreas, cubriendo un 28% del país galo, se estima que esta superficie se expande en 68.000 hectáreas anuales.

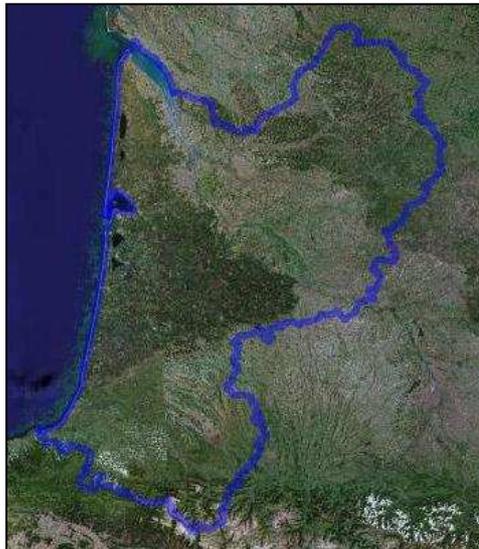
El pino marítimo, representa el 7,5 % de la superficie forestal francesa, con 1,3 millones de hectáreas, el 92,5% restante se divide entre 136 especies siendo las principales el roble y la haya.

La región de Aquitania, en el suroeste de Francia, tiene una superficie total de 1,8 millones de hectáreas de bosques, un 42,9% de su superficie total, siendo la principal zona forestal del país galo. Casi un millón de hectáreas corresponden a Pino marítimo, esta especie conífera se concentra en el Macizo forestal de Landas de Gascone, (**Figura 2**) constituyendo un bloque homogéneo de 200km de norte a sur a lo largo de la costa y 150km de este a oeste. La actividad forestal regional genera 28.000 empleos y exportaciones por 900 millones de euros. El pino marítimo es por lo tanto, un componente de importancia dentro de la silvicultura francesa, sobretudo en la región de Aquitania, a mayor escala es la conífera más utilizada para reforestación en el sud-oeste de Europa (España, Portugal).

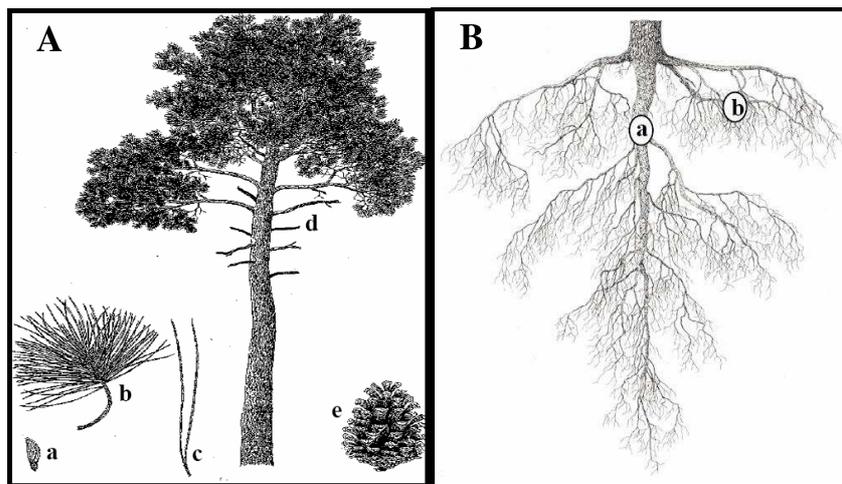
### **2.2.2 El pino marítimo**

#### **2.2.2.1 Descripción biológica y hábitat**

Pino marítimo es una especie perenne de longevidad mediana, que alcanza la madurez en 45 o 60 años y puede vivir hasta 200 años. De crecimiento inicial rápido, a la edad adulta puede alcanzar una altura de 20 a 35 metros. Su aspecto es de tronco sinuoso, especialmente en la base, la copa es piramidal en plantas jóvenes y redondeada en los adultos, ocupando solo el tercio superior del árbol, en individuos longevos es acampanada, deprimida y vacía en la parte interna del árbol. Las ramas son de color rojizo. Las hojas aciculares están agrupadas de dos en dos, de 15-25cm x 0,1-0,2cm, poseen estomas por ambas caras, son espesas, rígidas relucientes y de color verde oscuro, permanecen en el árbol por un tiempo de 2-3 años. El sistema radicular es doble y sus porciones primarias y secundarias son profundas (**Figura 3**). Es una especie monoica (cada árbol es bisexual, posee células reproductoras masculinas y femeninas), alógama y florece en abril/mayo produciendo abundante polen de dispersión anemófila. Las flores masculinas son de 1 a 2cm de largo y 4-6mm de ancho, cilíndricas o



**Figura 2** : Vista satelital de la región de Aquitania, en el sud oeste de Francia, a lo largo de su región costera, se aprecia el bosque, o masivo forestal de Landas de Gascone, compuesto exclusivamente de Pino marítimo.



**Figura 3:** El pino marítimo **A:** Parte aérea (a) semilla (b) agujas, (c) fascículo (d) tronco y ramas (e): cono. **B:** Parte subterránea (a): raíz principal, (b) raíz secundaria (Extraído de Le Provost, 2003a).

apuntadas y bastante numerosas, dispuestas en espigas de hasta 7cm de longitud y de color amarillento; los estambres, con dos sacos polínicos, se disponen helicoidalmente sobre un eje. Los conos femeninos aparecen precozmente entre 5 a 8 años y forman conos voluminosos (10-18cm), pardo rojizos y que contienen los piñones de hasta 9mm de longitud, con un ala articulada cuatro veces mas larga que la semilla. En un plano molecular, el cariotipo del pino marítimo es diploide y presenta 24 cromosomas ( $2n=2x=24$ , (Saylor, 1964)). La cantidad de ADN contenida por célula diploide es de 51pg (Chagné et al., 2003) una de las mayores cantidades del reino vegetal.

A nivel de su hábitat, pino marítimo es un árbol que se adapta muy bien a un suelo arenoso, pobre y ácido, como el existente en Landas de Gascone, por esto, es frecuente en sistemas dunares costeros, aunque no es exclusivo de este medio y se puede encontrar hasta 1500m de altitud. Bien adaptado al clima marítimo temperado y a una temperatura suave y regular, exige una ligera humedad del aire, alta luminosidad y soporta una sequía estival moderada, es parcialmente sensible a fuertes heladas, sobretodo en las procedencias mediterráneas.

#### 2.2.2.2 Taxonomía

**Tabla 1.** Taxonomía del pino marítimo

Reino	<i>Plantae</i>
Filo o Division	<i>Pinophyta</i>
Clase	<i>Pinopsida</i>
Orden	<i>Pinales</i>
Familia	<i>Pinaceae</i>
<b>Genero</b>	<b><i>Pinus</i></b>
Sub-genero	<i>Pinus</i>
Sección	<i>Pinaster</i>
Sub-sección	<i>Australes</i>
<b>Especie</b>	<b><i>pinaster</i></b>

### **2.2.2.3 Repartición geográfica**

La distribución actual de pino marítimo parece ligada directamente a la última glaciación (de -0,7 a -0,01 millones de años (Baradat and Marpeau-Bezard, 1988). Además, durante los últimos 150 años su distribución geográfica ha sido largamente modificada por la mano del hombre (Le Maitre, 1998).

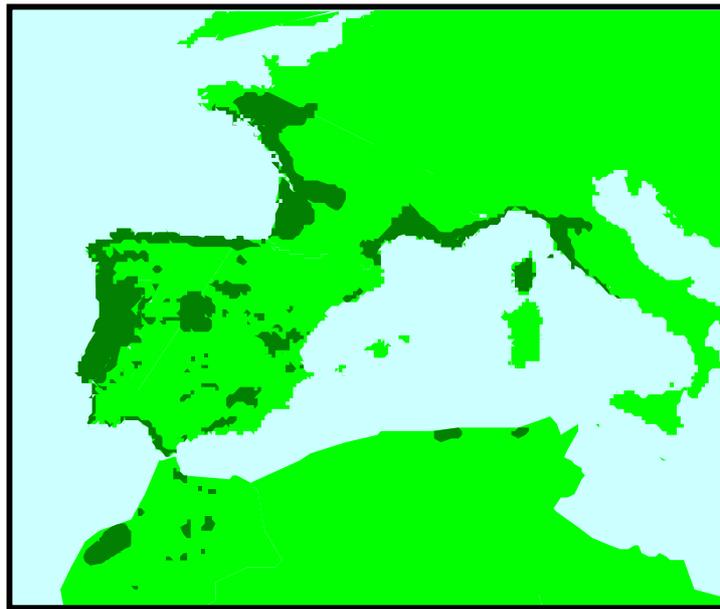
Esta es una especie autóctona de la costa mediterránea occidental y de la costa atlántica, entre las latitudes 31° a 46° norte y las longitudes 9° y 13° este. Aunque necesita de un hábitat característico, posee un área de distribución bastante extensa, con condiciones climáticas y pedológicas distintas. En efecto, el área natural del pino marítimo se extiende de sur a norte, desde Maghreb (Túnez, Algeria y Marruecos) a Vendée (Francia) y de oeste a este, desde Portugal a Italia (**Figura 4**).

Esta especie ha sido introducida a Australia, donde existen 500.000 hectáreas, África del Sur, Argentina, Nueva Zelanda y Grecia. Se ha tratado de introducir en Chile, Corea, Estados Unidos y Uruguay.

En Francia, se encuentra en estado natural a lo largo de la costa atlántica, a lo largo de la costa del mediterráneo y en Córcega. Su área de distribución fue extendida en el siglo XIX por plantación, desde los Pirineos en la región de Aquitania a la región de Bretaña por el norte. Se encuentra también en la región central, notablemente en el bosque de Orleáns.

### **2.2.3 Programa de mejoramiento INRA**

En los años 60 se inició un programa de mejoramiento genético de pino marítimo para Aquitania en el INRA, este programa se encuentra actualmente en su tercera generación de mejoramiento. Este programa comenzó con un ensayo de procedencia, correspondiente a toda el área natural de distribución de pino marítimo. Se eligió la procedencia Landesa como la mejor adaptada al hábitat local, por ser la más vigorosa y con mayor resistencia al frío. A fin de aumentar la rectitud del tronco, la procedencia Corsa fue introducida en los años 80. La estrategia seguida fue la de "selección recurrente", que permite preservar la diversidad y producir ganancia genética a través de varias generaciones. Su principio se basa sobre la selección suave y progresiva de una población de mejoramiento, compuesta de un gran número de individuos. A cada generación, un ciclo de selección se cumple en dos fases: un



**Figura 4:** Área de distribución del pino marítimo. Extraído de Le Provost 2003

inter-cruzamiento de individuos de manera de recombinar sus caracteres favorables y la selección de los mejores recombinantes entre los descendientes obtenidos. Estos individuos seleccionados van a formar la nueva generación de la población de mejoramiento. A lo largo de las generaciones, los genes favorables aportados por los diferentes progenitores son reunidos en los descendientes, que progresivamente tienen todas las propiedades buscadas. Este esquema de generación de la población, es una garantía para mantener la variabilidad genética y la diversidad a largo plazo, ya que la población esta sujeta a una débil intensidad de selección, lo que permite conservar con una fuerte probabilidad los caracteres genéticos neutros. Estos caracteres genéticos neutros, es decir que no son favorables o desfavorables con respecto a la selección, son la fuente de variabilidad genética potencialmente necesaria para responder a nuevos criterios de selección que pueden surgir en el futuro.

Con respecto a las ganancias genéticas obtenidas por el programa de mejoramiento del INRA, estas alcanzan al 30% en productividad y rectitud en tronco para las variedades comercializadas actualmente en Aquitania (Raffin and Pastuszka, 2002).

#### **2.2.4 Nuevas propiedades de la madera son necesarias**

Características importantes en la calidad de la madera como la composición química, propiedades de las fibras y densidad de la madera no han sido incluidas en el programa de mejoramiento del INRA, este tipo de características deben ser prontamente incluidas. Observaciones recientes corroboran esta afirmación, por ejemplo la adopción de nuevas prácticas silvícolas y la adopción de variedades mejoradas han reducido el tiempo de cosecha de los árboles (de 60 a 40 años en Francia), produciendo el mismo volumen de madera, pero con una mayor proporción de madera juvenil (Larson et al., 2001), lo que causa una reducción intrínseca en la calidad de la madera.

Junto con el problema de la calidad de la madera, existe una creciente preocupación con respecto a la respuesta de los bosques a la sequía y en general al estrés abiótico y biótico, por el masivo cambio de las condiciones climáticas anunciadas debido al calentamiento climático (Saxe et al., 2001). Sin embargo, este trabajo de tesis esta orientado sólo a obtener una mayor comprensión de la xilogénesis, para abordar el problema de la calidad de la madera.

### **2.2.5 Calidad de la madera**

(Pot, 2004), definió la calidad de la madera como un concepto compuesto, que puede ser estudiado a diferentes niveles ya sea químico (por ej. Composición y contenido de ligninas, celulosa, hemicelulosa), anatómico (morfología celular), físico (elasticidad) o tecnológico (densidad). Estudios tradicionales de genética cuantitativa han demostrado que las propiedades de calidad de la madera son variables y heredables (Cornelius, 1994) por lo tanto, pueden ser seleccionadas y proveer ganancias genéticas significativas (Zobel and Van Buijtenen, 1989; Nyakuengama et al., 1999; Pot et al., 2002). El problema reside en el alto costo de evaluar un número elevado de individuos, el extenso trabajo que esto significa y el excesivo tiempo que se debe esperar para evaluar árboles adultos. Es por esto que la estrategia de marcadores moleculares es vista como una herramienta molecular a desarrollar para ayudar a los mejoradores en su selección.

### **2.2.6 El Cambium Vascular**

En plantas leñosas, los haces vasculares primarios son reemplazados por el sistema vascular secundario. Este sistema comprende el floema secundario (corteza interna) el cambium vascular y el xilema secundario (madera). El xilema tiene funciones mecánicas y de transporte (agua, sales minerales, hormonas) y el floema conduce fotoasimilados sintetizados en la fotosíntesis. El cambium vascular es una zona compuesta por células meristematicas que consiste en 5 a 15 células en división (Han, 2001) que rodea el tronco, ramas y raíces del árbol, las características de este tejido han sido revisadas por Le Provost (2003); Paiva (2006).

### **2.2.7 Diferenciación del Xilema**

Por repetidas divisiones, el cambium se mantiene a si mismo y genera células radiales de floema hacia el exterior junto con células radiales de xilema hacia el interior. Elementos genéticos claves originalmente identificados en la regulación de células meristematicas de los meristemas apicales también fueron encontrados expresados en el cambium vascular durante el crecimiento de la madera (Schrader, 2003), revisado por Groover (2005). La diferenciación de células cambiales en madera involucra una sutil y sofisticada red de eventos regulados espacial y temporalmente, donde fitohormonas, auxinas en particular (Tuominen et al., 1997; Uggla et al., 1998; Hellgren, 2003) citoquininas (Aloni, 1987), giberelinas (Eriksson et al.,

2000) y etileno (Hellgren, 2003) juegan un rol crucial en la coordinación de la expresión de cientos de genes a través del proceso de formación de madera.

### **2.2.8 La formación de madera**

El xilema secundario se origina desde el cambium vascular. La formación de madera o xilogénesis, se compone de cuatro etapas principales: división, expansión celular, formación de una pared secundaria y la muerte celular (**Figura 5**). Este proceso ha sido revisado extensamente por Fukuda (1996); Lachaud et al. (1999); Roberts and McCann (2000); Mellerowicz et al. (2001); Plomion et al. (2001).

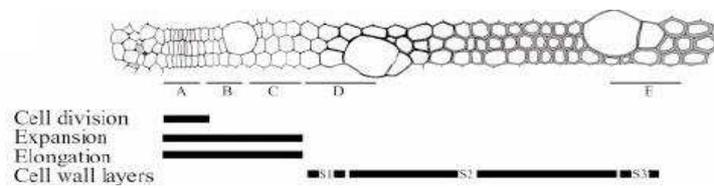
La xilogénesis es un complejo programa de desarrollo influenciado por el ambiente, el genotipo y la edad del cambium. Hay un gran número de genes expresados coordinadamente, principalmente involucrados en las vías de síntesis de polisacáridos, proteínas parietales y síntesis de ligninas.

#### **1) División de células del cambium**

El cambium comprende la zona de división celular. La velocidad de división de las células del cambium (Índice mitótico) y el número de células en la zona cambial determinan la producción de células de xilema. La división celular y el número de células en la zona del cambium presentan variaciones estacionales. En primavera, luego de la reactivación del cambium, la división celular llega rápidamente a un estado de equilibrio. Luego de un periodo principal de crecimiento del cambium, la división celular disminuye gradualmente hasta que el crecimiento cambial cesa, entrando en dormancia (Wilson, 1966; Gregory, 1971).

#### **2) Expansión radial de células hijas**

Siguiendo la zona de división, una zona de expansión puede ser distinguida donde las células se expanden radial y longitudinalmente durante la formación de la pared primaria hasta que alcanzan su tamaño final. En gimnospermas, el grado de elongación no excede el 10%. La fase de expansión continua hasta que la rigidez de la pared supera la presión de turgor de las células (Abe et al., 1997).



**Figura 5:** Representación esquemática de la xilogénesis, con sus cuatro etapas fundamentales, división celular, expansión y elongación, y depósito de las paredes celulares secundarias. Figura modificada de Hertzberg, 2001

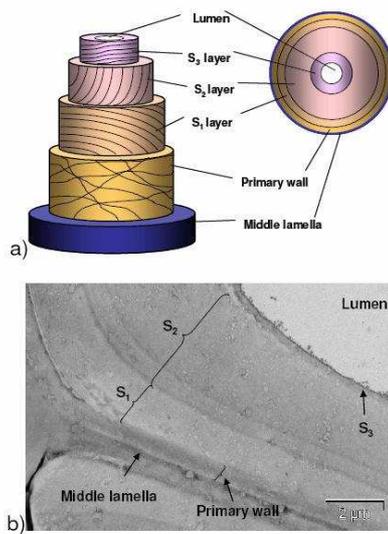
### **3) Engrosamiento de la pared celular** (depósito de celulosa, hemicelulosa, proteínas y lignina)

Una vez concluida la expansión, las células en diferenciación del xilema entran la zona de engrosamiento de pared, donde ocurre la biosíntesis de cuatro compuestos principales: polisacáridos (celulosa y hemicelulosa), ligninas, proteínas de pared celular y otros compuestos solubles (estilbenos, flavonoides, taninos y terpenoides) e insolubles en solvente neutro (pectinas y proteínas de pared celular) (Higuchi, 1996).

### **4) Muerte celular programada**

Finalmente, una vez completada la lignificación, las células entran a la fase de muerte celular programada. Durante las etapas tempranas de diferenciación de xilema, se acumulan enzimas autolíticas en las vacuolas. La vacuola colapsa (Jones, 2001), regulada por señales de Calcio y libera las hidrolasas (por ej. DNAsas y proteinasas) que degradan el contenido celular, pero no la pared secundaria. El periodo de diferenciación y maduración es usualmente corto, extendiéndose por 10-15 días en pino marítimo (Stokes, 1999). Al final de este periodo, la traqueida se encuentra metabolitamente inactiva, pero aun así comienza el depósito una gruesa pared secundaria dentro de la pared primaria. El principal componente de la pared celular secundaria es la celulosa, con sus microfibrillas ordenadas en forma paralela en cada capa. La pared secundaria (**Figura 6**) contiene en una capa externa ( $S_1$ ; 0,1  $\mu\text{m}$  a 0,35  $\mu\text{m}$  de espesor), una capa media ( $S_2$ : 1  $\mu\text{m}$  a 10  $\mu\text{m}$  de espesor) y una capa interior ( $S_3$ : 0,5  $\mu\text{m}$  a 1,10  $\mu\text{m}$  de espesor), cada una con diferente orientación de las microfibrillas de celulosa (Harada and Côté, 1985). El ángulo de las microfibrillas con respecto al eje es 60° a 80° para la capa  $S_1$ , 5° a 30° para la capa  $S_2$  y de 60° a 90° para la capa  $S_3$ . Otros componentes principales incluyen la lignina, hemicelulosas y proteínas de pared celular.

La composición promedio de la madera de pino marítimo es: Celulosa 42%, Lignina 30%, Hemicelulosas 28%, incluyendo mananos 13%, galactanos 4%, y xilanos 9%. (El alto contenido en lignina puede ser explicado por la presencia extendida de madera de compresión en el pino marítimo procedente de Aquitania). Los extractos totales son típicamente menores al 5%, variando entre 1,5 a 1,9% para diclorometano, etanol y agua. (Paiva, 2006).



**Figura 6:** Ultraestructura de las paredes celulares en células de xilema secundario en pino marítimo: A) Representación esquemática tridimensional de la lámina media, pared primaria y secundaria, con sus capas S<sub>1</sub>, S<sub>2</sub> y S<sub>3</sub>. Se indica también el lumen celular B) corte transversal mostrando la misma región Figura modificada de Paiva, 2006

Durante la formación de la pared celular, su composición puede cambiar en respuesta a factores bióticos y abióticos. Ya que la capa S<sub>2</sub> es responsable por el 75% al 85% del grosor total de la pared, su composición química y la orientación de sus microfibrillas van a influenciar las propiedades químicas, mecánicas y físicas de la célula y afectara por lo tanto la calidad de la madera y sus productos derivados.

## **2.2.9 Variaciones en la Formación de madera**

El diámetro final de la traqueida y el grosor de la pared celular son el resultado de la duración y velocidad de dos etapas: la expansión celular y la etapa de formación de pared celular secundaria (Withmore and Zanher, 1966; Dodd and Fox, 1990).

La madera en coníferas esta formada en un 90% de traqueidas, lo que haría pensar que se trata de un material simple, sin embargo es altamente variable (revisado por Zobel and Van Buijtenen, (1989). En un mismo árbol se pueden encontrar hasta seis tipos de madera diferente: temprana, vs. tardía, juvenil/de copa vs. madura/de base y de compresión vs. opuesta. La mayor fuente de variación depende de la característica considerada, por ejemplo, el largo de las traqueidas es más variable dependiendo de la edad cambial, la densidad de la madera es mas variable dependiendo de la temporada de crecimiento, con respecto a las características químicas, la mayor variabilidad es encontrada en la madera de compresión vs. la opuesta.

### **2.2.9.1 Variación durante la temporada de crecimiento**

En zonas templadas, la mayor variabilidad observada en las características de la madera es la transición de madera temprana a madera tardía, esta variación ocurre dentro de cada anillo anual de crecimiento (**Figura 7A**). La madera temprana se forma en primavera, al comienzo de la temporada de crecimiento, cuando las condiciones de temperatura y precipitación son favorables para el crecimiento activo. La madera tardía aparece al final de la temporada de crecimiento, en verano, cuando condiciones adversas disminuyen la tasa de división celular en el cambium, la etapa de expansión celular se reduce y la etapa de formación de pared secundaria extiende (Uggla et al., 2001). A nivel anatómico las diferencias son evidentes (**Figura 7B**), la madera tardía posee lumen celular de menor diámetro y traqueidas más largas, con paredes celulares engrosadas, lo que origina una madera de mayor densidad. A nivel químico, la madera temprana tiene mayor contenido de lignina y hemicelulosas y un menor contenido de celulosa y extractivos totales, comparado con la madera tardía (**Tabla 2**), Larson

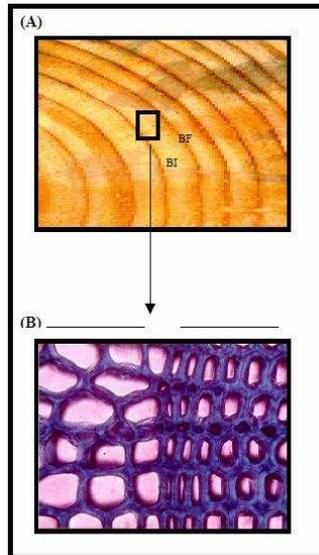
et al., (2001), explica este fenómeno por un control del engrosamiento de la pared celular dependiente de la disponibilidad de fotoasimilados, en la fase de crecimiento temprana de la temporada, los fotoasimilados producidos por acículas antiguas, son predominantemente usados en el crecimiento de nuevos tejidos meristemáticos y en elongación de las acículas, luego, avanzada la temporada, los fotoasimilados producidos por las acículas de la temporada anterior y nuevas pueden ser usados en la síntesis de pared secundaria típica de madera de verano.

**Tabla 2:** Comparación de la composición química entre Madera temprana y Madera tardía. Tabla extraída de Paiva, 2007

composición química	Madera temprana (%)	Madera Tardía (%)
Lignina	29,7	28,5
Celulosa	42,9	44,8
Hemicelulosa	27,4	26,7
Mananos	12,2	12,8
Xilanos	10,5	9,5
Galactanos	2,9	3,1
Extractivos	2,94	2,99
Diclorometano	0,95	0,82
Etanol	1,01	0,96
Agua	0,98	1,21

### **2.2.9.2 Variación ontogenica: madera juvenil/de copa vs. madera madura/de base.**

En un árbol maduro la madera difiere significativamente, desde el centro del tronco a la corteza y desde la base a la copa (Zobel and Van Buijtenen, 1989; Zobel and Sprague, 1998). Un aumento en la proporción de madera juvenil esta relacionado positivamente con la proximidad a la copa (Larson et al., 2001) y la disponibilidad de fotoasimilados. Durante los primeros años de desarrollo las ramas se extienden a lo largo de todo el árbol, en otras palabras, el árbol es solo copa (Larson et al., 2001) y la madera formada por todo el árbol es madera

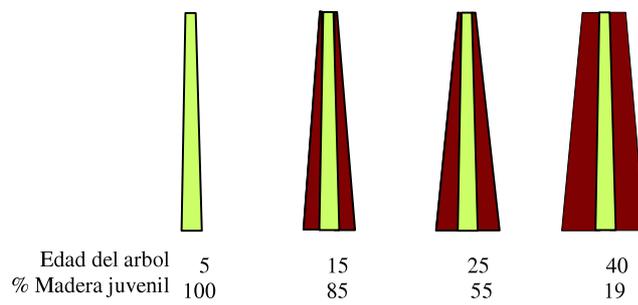


**Figura 7:** Anillos de crecimiento anuales típicos de climas templados B) madera temprana de primavera vs madera tardía o de verano. Figura modificada de Le Provost, 2001

juvenil, al correr de los años cuando las ramas solo se encuentran en la parte superior del árbol, se comienza a producir un nuevo tipo de madera en la base del árbol, la madera madura o de base. La progresiva extensión hacia la copa de la madera madura o de base y la regresión de la madera juvenil o de copa, resultan en una zona de madera juvenil extendiéndose a partir del tronco. Se considera como madera juvenil también, la madera generada durante los primeros años de vida, en el centro del tronco. En el transcurso de esta tesis se usarán las definiciones: madera juvenil o de copa y madera madura o de base, para diferenciar la madera originada a partir de tejido cambial de distinta antigüedad. La madera juvenil/de copa esta caracterizada por traqueidas mas cortas, menor gravedad específica, mayor lumen celular, y células con pared más delgada con respecto a la madera madura/de base. Estas características resultan en productos madereros débiles, papel con menor resistencia a la tensión y mayor resistencia al rasgado (Zobel and Van Buijtenen, 1989). Típicamente se encuentra un mayor porcentaje de madera de compresión en madera juvenil/de copa. La madera juvenil/de copa produce un menor rendimiento de celulosa por su alto contenido de lignina y xilanos, además produce un mayor costo de extracción y una mayor producción de desechos tóxicos. La cantidad relativa de madera juvenil se ve reducida a lo largo de la vida de un árbol, En *Pinus taeda*, se ve reducida de un 85% a los 15 años a un 55% a los 25 años y 19% a los 40 años (Zobel, 1972) (**Figura 8**). Mientras mayor sea la edad que posea el árbol, tendrá menor proporción de madera juvenil, ya que el centro de madera juvenil permanece sin cambios. La transición desde madera juvenil a madera madura, estimada desde datos de densidad, ocurre entre el anillo 10 a 12 (Zobel, 1972; Dumail et al., 1998). Debido al intenso manejo forestal, las plantaciones pueden alcanzar tamaños explotables a menor edad. Y ya que el crecimiento en diámetro es usualmente mayor durante los primeros años de crecimiento, cuando se produce madera juvenil, el centro de madera juvenil puede ser una parte muy significativa de la cosecha.

#### **2.2.10 Avances en Genómica de Especies Forestales**

La genómica puede ser definida como el desarrollo y aplicación de aproximaciones que toman en cuenta la totalidad del genoma para abordar la estructura y función de genes, lo que incluye secuenciación de ADN, mapeo de genes y perfiles de expresión génica.



**Figura 8:** Representación de el porcentaje de madera juvenil/de copa (en verde) y madura/ de base (en café) durante el crecimiento de un árbol desde 5 a 40 años. Representación de un corte tangencial de un segmento del tronco.

Las especies forestales son excelentes modelos para estudiar la formación de madera (Lev-Yadun; Taylor, 2002). Álamos y pinos son considerados los modelos de mayor importancia para estudiar la formación de madera en maderas duras (angiospermas) y maderas blandas (gimnospermas), respectivamente.

La era genómica para el ámbito forestal comenzó en los años noventa, con estudios de genética y mapeo genético de QTLs, (Cervera et al., 2001; Costa et al., 2000). Gracias al avance técnico, a fines de los noventa se inició en distintos laboratorios el análisis simultáneo de miles de transcritos y proteínas, dando una oportunidad de comprender la función de proteínas, regulación génica y eventualmente como son formados estos organismos de larga vida (Plomion et al., 2005).

#### **2.2.10.1 Secuencias de Genomas**

En 2004, se completó el genoma de *Populus trichocarpa* <http://genome.jgi-psf.org/Poptr1/Poptr1.home.html>, su impacto es analizado por Tuskan et al., (2004), muy recientemente en junio del 2007 se anunció en la última reunión de IUFRO el inicio del proyecto para la secuenciación del genoma de *Eucalyptus grandis* por la red EUCAGEN (<http://www.ieugc.up.ac.za/>), durante los próximos dos años. Sin embargo, en coníferas la estrategia de secuenciación del genoma no es técnicamente posible en la actualidad por limitaciones derivadas de su gran tamaño de genoma.

#### **2.2.10.2 Secuencias de EST**

Las librerías de ESTs para coníferas, de acceso público y de organismos privados, probablemente alcancen a un millón de entradas (FAO, 2004), siendo las principales especies estudiadas *Pinus taeda*, *Pinus pinaster* y *Picea glauca*. Boerjan (2005) presenta una revisión de las bases de datos disponibles. Datos públicos de distintos laboratorios se han agrupado en el Instituto del Cáncer Dana Farber y el Índice de Genes de Pino “DFCI Pine Gene Index” (<http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=pine>). En su versión 6.0 de Julio 2005, contiene 327.484 ESTs, con un total de 45.557 genes únicos. Los trabajos de Le Provost (2003) y Paiva (2006) permitieron enviar 4.413 secuencias únicas de xilema (con un 54.2% de redundancia) a un centro de referencia, la Plataforma para manejo Integrado de Clones (Platform for Integrated Clone Management PICME (<http://www.picme.at>)). La librería de cDNA de xilema ha demostrado ser útil para la identificación *in-silico* de SSRs (Chagné et al., 2004) y SNPs (Dantec et al., 2004), igualmente, ha sido útil para el desarrollo de

microarreglos. Actualmente se encuentra en desarrollo un microarreglo de segunda generación con 12.000 spots que integra genes de xilema, raíz y brotes de *Pinus pinaster* (Chaumeil, 2007).

### **2.2.10.3 Estudios de expresión**

Los estudios de expresión global de genes han sido utilizados para identificar los genes o las proteínas de interés que permitan una mejor comprensión de los mecanismos moleculares implicados en la formación de madera. En pino taeda se han realizado estudios usando microarreglos para analizar la formación de madera madura y madera de compresión (Whetten et al., 2001). Estos autores identificaron genes involucrados en la biosíntesis de ligninas, citoesqueleto y proteínas parietales. Algunos genes identificados fueron específicos de Pino taeda y otros también están presentes en *Arabidopsis*. Macroarreglos fueron utilizados por (Le Provost, 2003) y (Paiva, 2006) para analizar genes expresados en madera temprana vs. madera tardía, madera juvenil vs. madera madura y en madera de compresión vs. madera opuesta.

Un importante resultado de la secuenciación del genoma del álamo, es la elaboración de microarreglos de oligonucleótidos. El alto nivel de especificidad de los microarreglos de oligonucleótidos esta siendo usado para identificar sistemáticamente miembros de familias multigénicas de factores de transcripción y proteínas asociadas a pared celular, para determinar su expresión en tejidos vasculares secundarios (Plomion et al., 2005).

La técnica de SAGE (Serial análisis of gene expresión) ha sido utilizada por Lorenz and Dean, (2002) para analizar genes involucrados en la formación de madera juvenil y madera madura en pino taeda, representando un máximo de 42.000 genes. Egertsdotter et al., (2004) analizó 350 ESTs en un microarreglo identificando 71 que mostraban sobre expresión en madera temprana o madera tardía en pino taeda. Los datos de expresión fueron relacionados con análisis de grosor de la pared celular y datos climáticos. Igualmente en pino tadea, Yang et al., (2005) analizó genes implicados en la formación de la madera a través de la temporada y comparó 2 diferentes orígenes geográficos, con un microarreglo de 2171 ESTs, identificando 53 genes diferencialmente expresados en ambas poblaciones entre madera temprana y madera tardía. Recientemente, Cato et al., (2006) utilizó “Differential Display” modificado para analizar la expresión de genes entre madera juvenil/de copa y madera madura/de base, relacionando estos datos con grosor de la pared celular y crecimiento radial en pino radiata.

Estos investigadores observaron que el radio de división celular fue 3.3 veces mayor en el cambium localizado en la copa comparado con la base, y que esta división acelerada estaba asociado con un engrosamiento reducido de la pared secundaria. También se encontró que el porcentaje de células en la fase de engrosamiento de la pared celular fue significativamente menor en la copa del árbol. Genes involucrados en división y expansión celular fueron más expresados en la copa y genes involucrados en la represión del ciclo celular y engrosamiento de la pared fueron más expresados en la base.

#### **2.2.10.4 Proteómica**

A nivel transcriptómico, técnicas como los microarreglos generan grandes volúmenes de información con respecto a la expresión de transcritos. Sin embargo, se debe tomar en cuenta que un gen no es equivalente a un transcrito y un transcrito no es equivalente a una proteína (Peck, 2005). Es por esto que resultados transcriptómicos deben ser interpretados cautelosamente ya que los niveles de mRNA y de proteína no están claramente relacionados (Gygi et al., 1999). Además de esta desviación sistemática, la discrepancia entre niveles de mRNA y proteínas es causada naturalmente por la regulación post-transcripcional, que no esta distribuida uniformemente entre tejidos ni tipos celulares vegetales. Por estas razones estudios transcriptómicos no proveen información fidedigna sobre la expresión y abundancia de las proteínas.

Además de esto, por la gran heterogeneidad de las propiedades fisicoquímicas de las proteínas expresadas en un momento dado en la célula, las proteínas son mucho más difíciles de aislar, manipular e identificar que el mRNA y no existen técnicas comparables al PCR para amplificar proteínas de baja abundancia. Avances recientes en espectrometría de masas, estandarización de los métodos de separación y en la disponibilidad de bases de datos con información de expresión para especies forestales, han significado potenciar la proteómica transformándola en una herramienta eficiente en la identificación de un gran número de proteínas. Sin embargo, la Proteómica tiene aun serias limitaciones:

- La presencia de sustancias interferentes durante el proceso de extracción (lípidos, ácidos nucleicos, carbohidratos), sumado en plantas a la extracción de proteasas, polifenoles, taninos, pigmentos , lignina y ceras (Carpentier et al., 2005).

- A diferencia de los tejidos animales, las células vegetales se encuentran rodeadas de una pared celular, complicando la extracción.
- Los organelos celulares (mitocondria , cloroplastos, etc.) tienen cada uno su proteoma.
- En tejidos verdes la presencia de una gran cantidad de RuBisCO, puede distorsionar los patrones electroforéticos. La existencia de proteínas abundantes propone desafíos al rango dinámico de detección para proteínas poco abundantes. En proteínas de semillas existen problemas análogos por la abundancia de proteínas de reserva.

La identificación de proteínas vegetales es también más problemática en comparación con el campo de investigación humano o microbiológico, ya que existe un menor número de genomas secuenciados, aunque últimamente se ha avanzado en ciertos cultivos publicando bases de datos de ESTs de libre acceso.

Los últimos avances en Proteómica Vegetal han sido revisados por (Canovas et al., 2004; Rossignol et al., 2006; Jorin et al., 2007). La separación de polipéptidos por electroforesis bidimensional ha sido utilizada en especies forestales desde la década del noventa en una variedad de aplicaciones, por ejemplo para estudiar los cambios en el perfil proteico durante las etapas tempranas de desarrollo en pino taeda (Groome et al., 1991), para caracterizar la respuesta de pinos a estrés biótico (Ekramoddoullah and Tan, 1998) y abiótico (Costa et al., 1998), para diferenciar isoenzimas involucradas en el metabolismo de nitrógeno en pino silvestre (Ávila et al., 1998), para la determinación de polimorfismos genéticos y mapeo genético en pino marítimo (Bahrman and Petit, 1995; Plomion et al., 1995; Ávila Sáez et al., 2000) y para la identificación de proteínas expresadas en madera de compresión en esta misma especie (Plomion et al., 2000).

En la década presente, dentro del campo de la Proteómica Forestal, destaca el trabajo del grupo TREENOMIX de la Universidad de British Columbia, Canadá, donde se utilizaron técnicas proteómicas para definir las cuatro etapas de la embriogénesis somática en abeto (picea blanca) (Lippert et al., 2005) y para diferenciar entre perfiles proteicos en respuesta a heridas o a infección por pulgones en esta misma especie (Lippert et al., 2007). En un proyecto liderado por INRA, Francia, se elaboró un mapa proteómico de referencia en álamo y se analizó el efecto del estrés por sequía en el proteoma de hoja y raíces (Plomion et al., 2006).

También en angiospermas, recientemente en la Universidad de Sao Paulo se describió por una aproximación proteómica la formación de madera juvenil en *Eucaliptus* (Fiorani et al., 2007). En pino marítimo, Gion et al.(2005) reportó un mapa de referencia de tejidos formadores de madera, para distintos tipos de madera (temprana vs. tardía, juvenil vs. madura, de compresión vs. opuesta). Estos autores mostraron también una débil correlación positiva entre la regulación transcripcional y translacional, sugiriendo que los niveles de proteínas pueden ser escasamente estimados a partir de niveles de transcritos. La aproximación proteómica es ciertamente relevante para investigar genes involucrados en la formación de madera y en la calidad de la madera. En este contexto el propósito del presente trabajo fue estudiar la variación del proteoma en tejidos formadores de madera en pino marítimo en un gradiente de temporada y en un gradiente de edad cambial.

### **2.3 Planteamiento del problema**

La madera es un material variable tanto en el plano químico y anatómico, al igual que en sus propiedades mecánicas. Junto con la variación genética, el ambiente y la edad del cambium afectan igualmente las propiedades de la madera. Existen seis tipos diferentes de madera al seno de un mismo árbol. La comprensión de los mecanismos moleculares implicados en la formación de madera pasa por el estudio del genoma expresado durante la xilogénesis, tanto a un nivel de transcritos como de proteínas. Los mecanismos identificados pueden brindar información relevante sobre genes que pueden ser usados como marcadores moleculares en pino marítimo o servir como base para su verificación en otras especies de coníferas.

### **2.4 Formulación de Hipótesis**

En el curso de esta tesis se testeará la hipótesis que existen categorías funcionales de proteínas sobre-expresadas en distintos tipos de madera, responsables por la variabilidad presente en un árbol. Estos genes candidatos expresionales pueden ser responsables por la variación natural de la madera y sus propiedades finales y por lo tanto pueden proveer herramientas moleculares para ayudar a los mejoradores a seleccionar por calidad de madera.

## **2.5 Objetivos Generales y específicos**

### **Objetivo General**

Identificar a nivel proteómico, metabolismos celulares responsables de la diferencia observada en la madera en el seno de un mismo árbol.

### **Objetivos específicos**

- a) Analizar por técnicas proteómicas un gradiente desde la base a la copa del árbol (de madera madura muestreada en la base a madera juvenil muestreada en la copa del tronco).
  
- b) Analizar por técnicas proteómicas un gradiente temporal (de madera de primavera o temprana a madera de verano o tardía, colectada al final de la temporada de crecimiento).
  
- c) Una vez seleccionadas las proteínas de interés, mediante análisis de imágenes y análisis estadísticos, aquellas que presenten acumulación diferencial serán caracterizadas sistemáticamente por espectrometría de masas en tandem (ESI MS/MS).
  
- d) En paralelo, las mismas muestras serán caracterizadas químicamente.

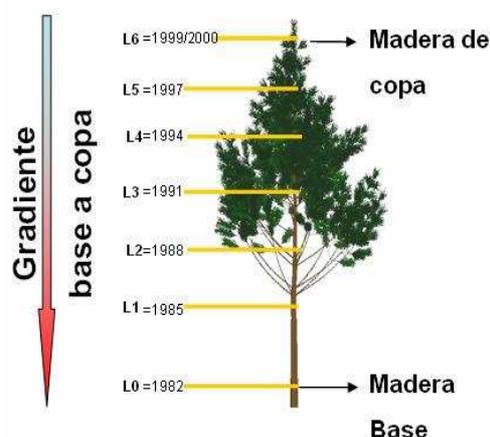
### 3. MATERIALES Y METODOS

### 3. Materiales y Métodos

#### 3.1 Diseño general de los experimentos.

##### 3.1.1 Gradiente de edad cambial 2003

Un total de 7 muestras fueron tomadas a través de un gradiente desde la base hasta la copa de un árbol recto de 30 años plantado en la Unidad Forestal Experimental de INRA – Pierroton (Cestas, Francia), (**Figura 9**) el origen de este árbol es un genotipo X procedente de Aquitania. Las muestras fueron colectadas el 12 de mayo del 2003\*, en los internodos de los años 1982, 1985, 1998, 1991, 1994, 1997 y 1999/2000, correspondiendo a la edad cambial 22, 18, 14, 12, 9, 6 y 3 años, o L0, L1, L2, L3, L4, L5 y L6 respectivamente (designadas “L” por el nivel “level”).

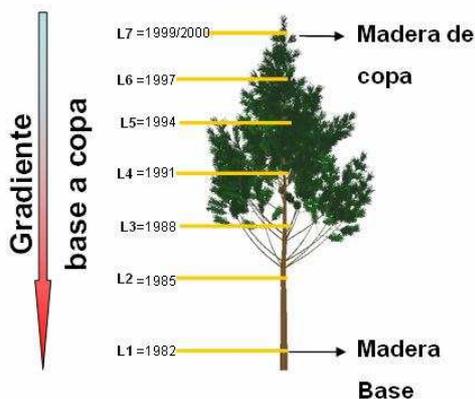


**Figura 9:** Edad cambial 2003.

Todas las muestras fueron usadas para describir la variación de la composición química de la pared celular. Las muestras de edad cambial: 22 años (L0), 14 años (L2), 9 años (L4) y 3 años (L6) fueron usadas para describir el proteoma.\* Estas muestras se encontraban ya colectadas y guardadas a -80°C al inicio de esta tesis.

##### 3.1.2 Gradiente de edad cambial 2005

De manera similar, 7 muestras fueron tomadas a través de un gradiente desde la base hasta la copa de un árbol recto de 32 años plantado en la Unidad Forestal Experimental de INRA Pierroton (Cestas, Francia), (**Figura 10**) el origen de este árbol es un genotipo X, procedente de Aquitania. Las muestras fueron colectadas el 6 de mayo del 2005\* en los internodos de los años 1982, 1985, 1998, 1991, 1994, 1997, 1999/2000 y



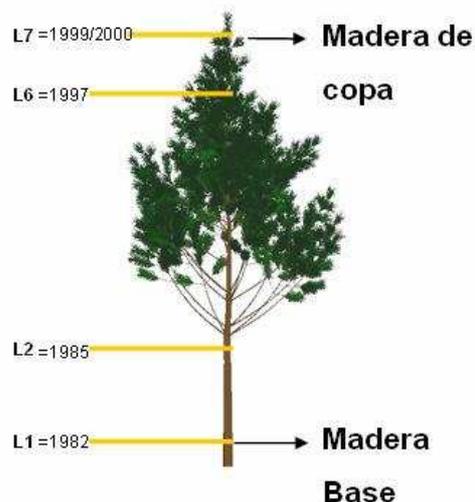
**Figura 10:** Edad cambial 2005.

correspondiendo a la edad cambial 24, 20, 16, 14, 11, 8 y 5 años, respectivamente. Todas las muestras fueron usadas para describir la variación de la composición química de la pared

celular. Las muestras de edad cambial 24 años (L1), 16 años (L3), 11 años (L5) y 5 años (L7), fueron usadas para describir el proteoma. \* Estas muestras se encontraban ya colectadas a -80°C al inicio de esta tesis.

### 3.1.3 Gradiente de edad cambial 2006

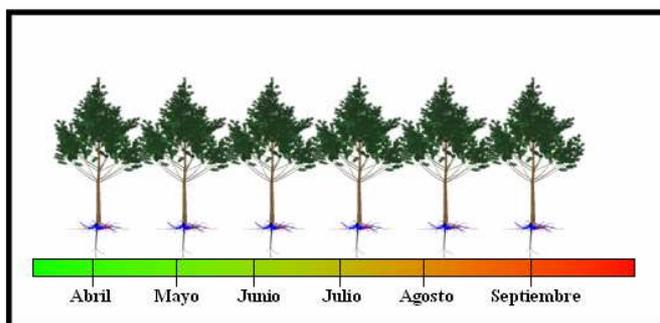
4 muestras contrastantes de madera de base y madera de copa fueron recolectadas desde dos árboles de genotipo A y B, el 19 de Junio del 2006 (**Figura 11**). 2 muestras estaban asociadas con tejido formador de madera base/madura (L1 y L2), formado por cambium de 25 y 21 años, respectivamente y 2 muestras estaban asociadas a tejido formador de madera de copa/juvenil (L6) y (L7), formado por cambium de 9 y 6 años, respectivamente. Para contrastar solo madera de base/madura vs. madera de copa/juvenil y para disminuir el número de geles, luego de la extracción y cuantificación de proteínas, las muestras similares fueron mezcladas, (L1+L2) y (L6+L7).



**Figura 11:** Edad cambial 2006.

### 3.1.4 Gradiente de Temporada 2006

Un total de 8 muestras fueron colectadas a lo largo de la temporada de crecimiento boreal, desde la mitad de la primavera a la mitad del verano 2006 en dos clones (3006 y 4015) (**Figura 12**). Las muestras fueron colectadas cada 10-11 días, el 28 de abril (S1), 12 de mayo (S2), 29 de mayo (S3), 12 de junio (S4), 26 de junio (S5), 10 de julio (S6), 24 de julio (S7) y 7 de agosto (S8), correspondiendo a los días del año 118, 132, 149, 163, 177, 191, 205 y 219, respectivamente. Todas las muestras fueron usadas para describir la variación de la composición química de la pared celular a través de la temporada, por espectrometría infrarroja con transformada de



**Figura 12:** Gradiente de temporada 2006.

Fourier y Pirólisis analítica. Solo las muestras del 12 de mayo (S2), 12 de junio (S4), 26 de junio (S5) y 7 de agosto (S8) fueron usadas para describir el proteoma a lo largo de la temporada de crecimiento. Estas fechas fueron elegidas en base a la expresión diferencial de genes candidatos verificada por PCR en tiempo real.

### 3.2 Técnica de recolección de tejido formador de madera



**Figura 13:** Recolección de tejido formador de madera.

Luego de remover la corteza, la lámina de floema con cambium adherido fue removida con un cuchillo, dejando expuesto el xilema secundario en diferenciación, este tejido se colectó desde el tronco, raspando con un cuchillo (**Figura 13**). Para los experimentos del gradiente base a copa, se muestrearon diferentes alturas a lo largo del tronco del árbol. Para el experimento de temporada de crecimiento 2006, se muestreo a la altura del pecho en las fechas indicadas. Zonas de color rojizo, con presencia de madera de compresión, no fueron colectadas. Las muestras de xilema secundario en diferenciación fueron congeladas en terreno con nitrógeno líquido, conducidas al laboratorio y guardadas a  $-80^{\circ}\text{C}$  hasta su uso posterior.

### 3.3 Procedimientos

#### 3.3.1 Extracción de proteínas

Partiendo de 2g de tejido fresco, proteínas totales de cada una de las muestras utilizadas para análisis proteómico, descritas anteriormente, fueron extraídas siguiendo el procedimiento reportado por Gion et al., (2005) (Anexo 1). Los extractos de proteínas fueron guardados a  $-80^{\circ}\text{C}$ . Tres extracciones fueron completadas para cada muestra y mezcladas para la cuantificación de proteínas. La cuantificación se realizó en seis replicados, siguiendo el

protocolo descrito por Ramagli and Rodriguez, (1985) (Anexo 2). La concentración fue calculada y 300 $\mu$ g de proteínas fueron depositados en cada “strip” o cinta IPG.

### **3.3.2 Electroforesis bidimensional**

La electroforesis bidimensional (O'Farrel, 1975) fue usada para analizar proteínas totales desde muestras de xilema siguiendo el procedimiento de Gion et al. (2005) adaptado para el sistema IPGphor (Amersham Bioscience, Uppsala, Suecia). Para el Isoelectroenfoco, se usaron cintas de 24cm con un gradiente lineal de pH de 4 a 7. Para los gradientes de edad cambial 2003 y 2005 se analizaron 4 niveles, con 4 replicados técnicos por muestra, resultando en un total de 16 geles para cada año. Para la edad cambial 2006 se analizaron 2 niveles (N1+N2) y (N6+N7), 2 clones y 3 replicados, resultando en un total de 12 geles. Para el gradiente de temporada 2006, se analizaron 4 tiempos de muestreo (S2, S4, S5, S8) en 2 clones con 3 replicados, resultando en 24 geles. En resumen, se utilizaron 44 geles para analizar el gradiente de edad cambial y 24 geles para el gradiente de temporada.

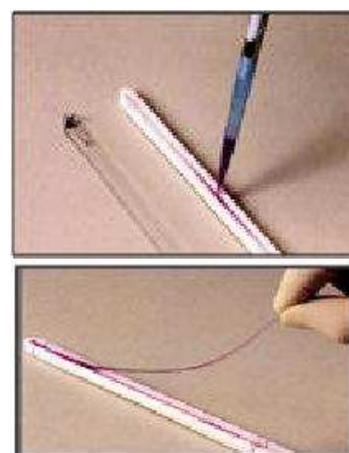
#### **3.3.2.1 Primera Dimensión: Isoelectroenfoco con el sistema IPGphor: migración según el pI**

El isoelectroenfoco (IEF) es un método electroforético que separa las proteínas de acuerdo a su punto isoeléctrico (pI). Las proteínas son moléculas anfóteras, que pueden tener carga negativa, positiva o neutra, de acuerdo a su pH circundante. La carga neta de una proteína es la suma de las cargas negativas y positivas de las cadenas laterales de sus aminoácidos y sus extremos amino y carboxilo. El punto isoeléctrico es el pH específico en el cual la carga neta de la proteína es cero. A valores de pH bajo su pI, las proteínas se encuentran cargadas positivamente y a valores de pH sobre su pI, negativamente. La presencia de un gradiente de pH es crítico para la técnica de IEF. En un gradiente de pH, bajo la influencia de un campo eléctrico, una proteína se mueve a la posición en el gradiente donde su carga neta es cero. Si una proteína que ya esta en su posición de carga neta cero, difundiera fuera de su pI, inmediatamente obtiene carga y migra de vuelta. Este es el efecto *focalizador* del IEF, que concentra a las proteínas en sus pIs y permite separar proteínas en base a muy pequeñas diferencias de carga.

El grado de resolución es determinado por el gradiente de pH y la fuerza del campo eléctrico, el IEF se realiza a altos voltajes (mas de 1000 V). Cuando las proteínas han alcanzado sus posiciones finales en el gradiente de pH, hay muy poco movimiento iónico en el sistema, resultando en una corriente final mínima (menor a 1mA). El IEF para una determinada muestra, generalmente se realiza por un número constante de Volts-hora (Volts-hora es el producto del voltaje y las horas sometidas a ese voltaje). Una denaturación y solubilización completa, alcanzada con una mezcla de Urea y detergentes, asegura que cada proteína este presente en solo una configuración y minimiza la agregación e interacciones intermoleculares, resultando en una mayor resolución y mayor consistencia de los resultados.

El método original para el IEF dependía de geles tubulares de poliacrilamida, con gradientes de pH generados por anfolitos, aunque este método fue utilizado en centenares de estudios de electroforesis bidimensional, tiene muchas limitaciones que han prevenido su amplia difusión. Los anfolitos poseen una gran variación entre lotes de fabricación, los gradientes no son estables y los geles tubulares tienen poca resistencia mecánica, haciendo la obtención de resultados altamente dependientes de la habilidad del operador. El método de gradientes de pH inmobilizados o IPG “immobilized pH gradients” fue desarrollado uniendo covalentemente grupos ácidos o básicos a geles de poliacrilamida. Cuando las cintas de IPG son usadas para el IEF, los mapas bidimensionales resultantes son superiores en términos de resolución y reproducibilidad. Existe una serie de ventajas de las cintas IPG sobre los geles de anfolitos: la separación de la primera dimensión es más reproducible porque el gradiente se encuentra covalentemente unido y no puede variar, las cintas al tener un soporte plástico son fáciles de manipular, pueden ser tomadas con pinzas o de los bordes usando guantes, el soporte plástico previene la rotura del gel y las cintas de IPG tienen una mayor capacidad de carga de proteína.

Los geles IPG “strips” o cintas son comercializadas listas para su uso, en forma deshidratada con un gradiente de pH establecido. Hay disponibles diferentes longitudes de cintas y diferentes gamas de pH. En el presente trabajo de tesis, solo se



**Fig 14:** Preparación del sarcófago.

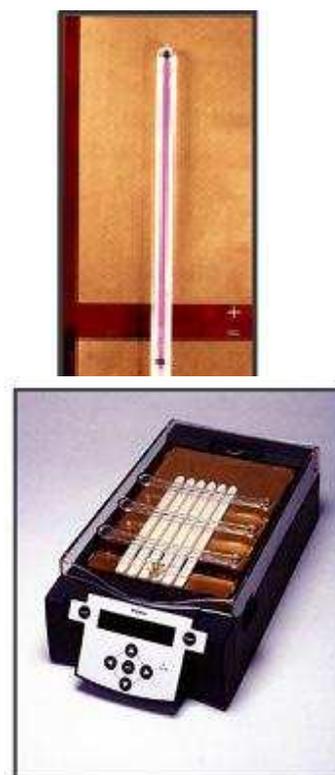
utilizaron cintas de 24cm con un gradiente de pH 4-7. Las cintas de IPG deben ser rehidratadas antes del IEF. La solución de rehidratación (Anexo 3) es aplicada junto a la muestra en el sarcófago que contiene la cinta. La importancia de los componentes de la solución de rehidratación se describe a continuación. Urea: Solubiliza y denatura proteínas, desplegando las cadenas peptídicas y exponiendo los aminoácidos ionizables. Tiourea: ayuda en la solubilización, particularmente de proteínas de membrana. Detergentes: Solubilizan proteínas hidrofóbicas y minimizan el agregamiento de proteínas, el detergente usado tiene carga neta cero. Agente reductor (DTT): Separa enlaces disulfuro, para permitir el completo desplegamiento de las proteínas. Tampon IPG: mejora la separación de las proteínas y su solubilidad. Colorante: Para verificar que el IEF conduce corriente al principio de la corrida. Nota: El colorante sale de la cinta mucho antes que la muestra termina la IEF.

Se mezcló el volumen correspondiente de muestra proteica (que contiene 300 $\mu$ g) con tampón de rehidratación. El volumen máximo fue de 470  $\mu$ l. La mezcla se adicionó al centro del sarcófago, se eliminaron eventuales burbujas (**Figura 14**).

Se quitó la película protectora de la cinta, y se instaló en el sarcófago haciendo coincidir su polaridad con este y con el gel hacia abajo, en contacto con la solución, teniendo precaución de distribuir uniformemente el líquido bajo la cinta y eliminando la presencia de burbujas. La cinta se recubrió con aceite Amersham (~2ml) para minimizar la evaporación de tampón y la cristalización de Urea.

Finalmente se cerró el sarcófago con su cubierta plástica, se limpio el exceso de aceite y se posicionó el sarcófago en el IPGphor, con el extremo en punta de este hacia el ánodo (+) (**Figura 15**).

El isoelectroenfoque se efectuó en el IPGphor siguiendo la siguiente programación: 50  $\mu$ A/cinta, 20°C, 8000V, c.s. hasta 74000 Vh.



**Figura 15:** Sistema IPGphor.

**Tabla 3:** programación del IEF

voltaje	tiempo
30V	~12 h rehidratación activa
200V	1h
500V	1h
1000V	1h *
8000V	gradiente 30min c.s. 74000 Vh

\*Al final de los 1000V se agregaron pequeños papeles filtros empapados en agua a los electrodos, para evitar el desecamiento de la cinta. Se agregó aceite de ser necesario. El sistema

IPGphor puede contener hasta 12 sarcófagos. Al lanzar la IEF a las 17h, se agregaron los papeles filtro a las 8h el día siguiente y el fin de la migración se produjo aproximadamente a las 18h.

### 3.3.2.2 Equilibración de las cintas

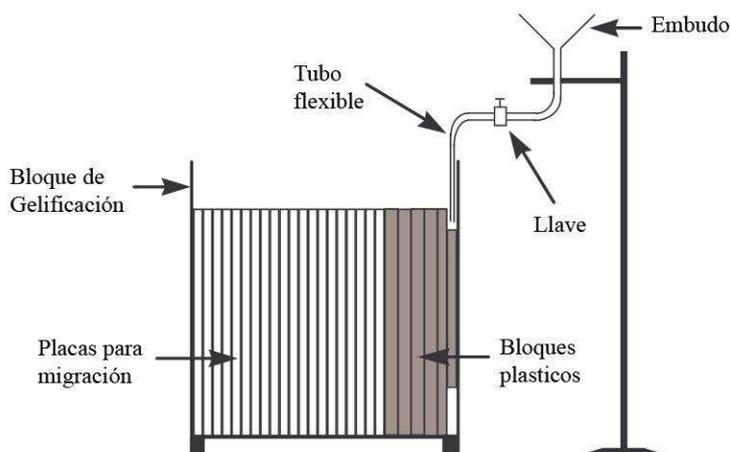
Esta etapa permitió saturar las cintas con SDS, necesario para la migración en segunda dimensión. La solución de equilibración (anexo 3) contiene Tampón, Urea, glicerol, SDS y colorante. El SDS denatura las proteínas y forma complejos proteína-SDS cargados negativamente. La cantidad de SDS unido a una proteína y su carga negativa adicional, es directamente proporcional a la masa de esta. Por esto una electroforesis de proteínas en presencia de SDS separa las separa en base a su peso molecular. El segundo paso de equilibración utiliza acetamida. La acetamida provoca una reacción de alquilación con los grupos tiol en las proteínas, previniendo su reoxidación. La equilibración de las cintas se realizo retirando las cintas del sarcófago con pinzas y estilando el exceso de aceite sobre papel absorbente. Las cintas fueron depositadas en tubos con tapas, apoyando la parte plástica de las cintas contra los tubos y dejando la cara de la cinta con gel hacia el interior. 10ml de la solución de equilibración conteniendo DTT fueron agregados y se dejo en agitación 15 min. Luego, los tubos fueron vaciados y se les agrego 10ml de solución de equilibración conteniendo iodoacetamida, dejando en agitación por 15min. Los tubos fueron vaciados, las cintas estiladas sobre papel, y finalmente guardadas en soportes plásticos rígidos, esto fue envuelto en papel aluminio y guardado a -80°C. Los sarcófagos fueron enjuagados rápidamente para retirar el aceite y dejados en remojo varias horas en una solución con 2 a 5%

de “IPGphor Strip Holder Clearing Solution” (Amersham), fueron frotados con una escobilla y enjuagados con abundante agua.

### 3.3.2.3 Segunda Dimensión: Electroforesis denaturante en gel de poliacrilamida.

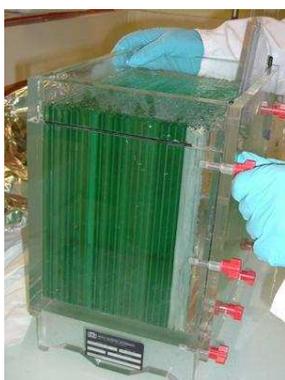
#### Migración según el peso molecular.

SDS-PAGE “SDS-poliacrilamide gel electrophoresis” es un método electroforético para separar polipéptidos de acuerdo a sus pesos moleculares. La técnica utiliza geles de poliacrilamida conteniendo sodio dodecil sulfato (SDS). La carga eléctrica intrínseca de las proteínas de la muestra no es un factor en la separación, debido a la presencia de SDS en la muestra y el gel. Las proteínas son totalmente desplegadas cuando un agente reductor como DTT es utilizado. Hay una relación aproximadamente lineal entre el logaritmo del peso molecular y la distancia relativa de migración del complejo SDS-polipéptido (Esta relación



linear es valida solo para un rango de peso molecular determinado por el porcentaje de poliacrilamida).

(La preparación del gel resolutivo y las soluciones necesarias para esta etapa se encuentran en anexo 4). En



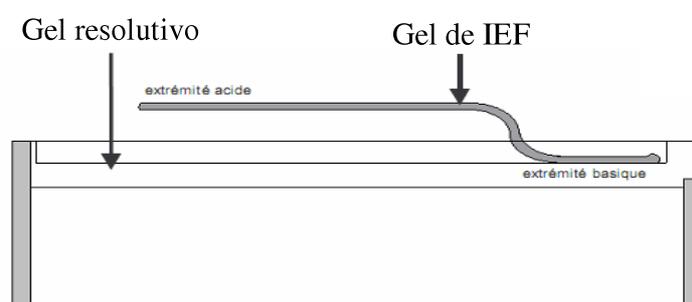
**Figura 16:** Bloque de gelificación.

una probeta de 2lt, se vació la acrilamida y luego se completó a 1lt con agua destilada. Se agrego el Tris con la sacarosa y luego agua destilada hasta completar 2lt. La solución se trasvasijó a un matraz Erlenmayer de 5lt. En una probeta de 1lt se agregaron 659,5ml de agua destilada y 40,5ml de SDS. Se mezclaron ambas soluciones y luego se filtraron al vacío los 2,7lt de esta solución con la ayuda de un embudo filtrante cilíndrico, de porosidad 40 a 100µm. Se dejo reposar media hora a fin de

disipar la espuma y se desgasifico por 2h. Debido al alto número de

placas se uso un sistema de bloque para la gelificación de los geles, en la Figura 16, se aprecia una vista lateral del sistema. Se repleta el espacio del bloque con 20 placas limpias, con espaciadores de 1mm y bloques plásticos, que rellenan el espacio sobrante. A la solución de acrilamida desgasificada se le agregaron 11,23ml de Persulfato de amonio al 10% y 1,8ml de TEMED. Finalmente se vació esta solución por el embudo. Si la temperatura de la pieza es muy elevada, la acrilamida puede polimerizar durante esta operación, por lo que la habitación y la solución deben permanecer frías. Es muy importante verificar que no se formen burbujas en el proceso de gelificación de los geles. Las placas fueron posicionadas en el bloque de gelificación con el extremo donde se deposita la muestra hacia arriba. La solución se vació hasta 1,5cm antes de la altura total de la placa, luego se recubrió la solución con butanol saturado de agua, lo que permitió obtener un frente de migración perfectamente derecho para el depósito de la cinta de la primera dimensión. Las placas gelificadas de esta manera fueron limpiadas de restos de acrilamida y preparadas para depositar la cinta proveniente del Isoelectroenfoco. Antes de depositar las cintas, se prepararon 100ml de una solución de agarosa de bajo punto de fusión y se conservo a baño maría para conservarla fundida.

Se descongelaron las cintas y se cortaron los extremos plásticos para reducir su longitud. Con el gel resolutivo en un soporte, se lleno



**Figura 17:** Preparación de gel resolutivo.

el espacio superior del gel con agarosa de bajo punto de fusión, se enjuago la cinta en el tampón de la cubeta y se depositó en la parte superior del gel ayudándose de pinzas y espátula (**Figura 17**). Se dejó polimerizar la agarosa por cinco minutos y luego se instaló el gel en la cubeta de migración (**Figura 18**).

#### 3.3.2.4 Condiciones de migración.

El amperaje se reguló al máximo y el voltaje hasta 0,3A, por la primera media hora y 0,6A por la segunda media hora. Finalmente el voltaje se aumentó hasta tener 110V máximo o 0,9A. Se dejó



**Figura 18 :** Cubeta de migración

migrando por 16 horas. La migración demora 17 horas en total y se efectuó en la noche.

### 3.3.2.5 Teñido de geles

La coloración se llevó a cabo en cajas de poliestireno transparentes con dos geles por caja en agitación (**Figura 19**)

(soluciones en Anexo 5). Se ocuparon 400ml. de solución por caja. Para el teñido de los geles se utilizó Coomassie Brilliant

Blue (Bio-Rad, Hércules, Ca, EEUU). Los geles fueron fijados por 2h en una solución ácido fosfórico 2% y etanol 50%. Luego de tres lavados de 30min., los geles fueron sumergidos en solución de incubación (metanol 34%, sulfato de amonio 17%, ácido fosfórico 2%) por 1h y posteriormente se dejaron en solución de teñido (metanol 34%, sulfato de amonio 17%, ácido fosfórico 2%, Azul de Coomassie 0,05%) por 5 días en agitación suave. Finalmente, los geles fueron digitalizados y guardados a 4°C en ácido acético 5% dentro de láminas plásticas selladas.

### 3.3.3 Adquisición de imágenes.

Los geles teñidos fueron digitalizados usando el Scanner de Imágenes (Amersham Biociences) y el software Labscan 5.0 (Amersham Biociences). Primero fue necesaria una calibración con escala de grises para transformar niveles de gris de cada píxel en la foto digitalizada de los geles, a valores de densidad óptica. El método de calibración usado fue "Colloidal Blue method" descrito en el manual de LabScan. Todas las fotos de los geles fueron guardadas como archivos ".tiff".

### 3.3.4 Detección de spots (manchas)

Para el gradiente base a copa 2003 (16 geles), el análisis de imágenes se realizó con el software Image Master 2D-Elite (IM2D, Amersham Biosciences). Para el gradiente de base a copa 2005 y 2006 (16 y 12 geles, respectivamente), el análisis de imágenes se realizó con el software Image Master 2D-Platinum 5.0 (IM2D, Amersham Biosciences). Ambas versiones del software son idénticas en el procedimiento de detección y apareamiento, solo su interfase es diferente. Se

utilizó el método de detección "wizard detection" propuesto por el software para detectar los spots. Luego, los spots detectados automáticamente fueron comprobados manualmente y algunos de ellos editados, agregados o removidos. Finalmente todos los geles fueron



**Figura 19 :** Teñido de geles.

comparados, para parear spots comunes, desde distintas imágenes. Para esto se usaron opciones automáticas de pareamiento de spots de IM2D. Luego de la revisión visual de los pareamientos de spots a través de los distintos geles, el software IM2D fue utilizado para construir un gel maestro. Para cada set de geles, cuando una mancha fue detectada en los cuatro replicados, esta mancha fue automáticamente agregada al gel maestro. Volúmenes normalizados fueron obtenidos finalmente usando el procedimiento de normalización de “total volume spot normalization” de IM2D. En este método, el volumen de cada spot es dividido por el volumen total de todos los spots de la imagen.

Para el gradiente de temporada 2006 (24 geles) el análisis de imágenes se realizó con el software Progenesis PG240 Samespots Versión 2006 (Nonlinear Dynamics, Newcastle, UK). Los geles fueron alineados antes de la detección para corregir por variaciones de corrida. Los spots fueron alineados y la intensidad de fondo fue sustraída usando el método de “Lowest on boundary”, que considera la intensidad circundante al spot y elige el píxel de menor intensidad como nivel de fondo para el spot. Los geles fueron alineados y una plantilla de detección idéntica fue aplicada para todos los geles. Volúmenes normalizados fueron finalmente obtenidos usando “total spot volume normalization”. El principio del software Progenesis es nuevo y basado en gráficos utilizados en videojuegos de última generación, con el software Image Master las imágenes de los geles permanecen intactos, luego de la detección de spots, el pareamiento se debe verificar manualmente y acomodar a leves diferencias de corrida en algunas zonas de los geles. Esta verificación del apareamiento consume una gran cantidad de tiempo, ya que se debe verificar spot por spot y gel por gel. El análisis de 16 geles con Image Master involucra aproximadamente 2 meses a tiempo completo. El software Progenesis deforma los geles, para corregir pequeñas diferencias de corrida y luego analiza los spots en todos los geles con una plantilla común. Un análisis usando esta aproximación para 16 geles reduce el tiempo de análisis a solo 2 a 3 semanas.

### **3.3.5 Tratamientos Estadísticos**

En la era post-genómica y gracias a los últimos avances técnicos en el área de la biología molecular, la generación de un gran volumen de datos se va a haciendo cada vez más común, actualmente es fundamental la interpretación de la información, para el análisis de un gran volumen de datos hace necesario el uso de modelos estadísticos, y la automatización de los

análisis, En el anexo 6 se explica en términos simples el análisis ANOVA, el concepto de interacción y algunas definiciones para la comprensión de clustering, se incluyen las rutinas de programación para R escritas para el análisis de los datos de esta tesis y bibliografía para estudio es mas detalle.

### 3.3.5.1 ANOVA

Para el gradiente base a copa 2003. Un análisis de varianza (ANOVA) fue usado para verificar el efecto de la edad cambial (L0, L2, L4 y L6) en la acumulación de proteínas, usando el siguiente modelo:  $Y_{ijk} = \mu + L_j + \epsilon_{ijk}$ , donde  $Y_{ijk}$  es el valor de la intensidad normalizada de la mancha  $i$  ( $i=1-1372$ ) en el nivel  $j$  ( $j = 0,2,4,6$ ) en el replicado  $k$  ( $k = 1-4$ ),  $\mu$  es la intensidad media de la mancha  $i$  en todos los geles,  $L_j$  es el efecto del nivel  $j$ , y  $\epsilon_{ijk}$  es el termino residual.

Para el gradiente base a copa 2005. Un análisis de varianza (ANOVA) fue usado para verificar el efecto de la edad cambial (L1, L3, L5 y L7) en la acumulación de proteínas, usando el siguiente modelo:  $Y_{ijk} = \mu + L_j + \epsilon_{ijk}$ , donde  $Y_{ijk}$  es el valor de la intensidad normalizada de la mancha  $i$  ( $i=1-743$ ) en el nivel  $j$  ( $j = 1,3,5,7$ ) en el replicado  $k$  ( $k = 1 - 4$ ),  $\mu$  es la intensidad media de la mancha  $i$  en todos los geles,  $L_j$  es el efecto del nivel  $j$ , y  $\epsilon_{ijk}$  es el termino residual.

Para el gradiente base a copa 2006, un modelo ANOVA de dos factores fue utilizado para fijar el efecto de los efectos principales (edad cambial y efecto clonal) y el efecto interacción utilizando el siguiente modelo:  $Y_{ijkl} = \mu + L_j + C_k + I_{lk} + \epsilon_{ijkl}$ , donde  $Y_{ijkl}$  es el valor normalizado del spot  $i$  ( $i=1-944$ ) en el nivel  $j$  ( $j = \text{base, copa}$ ) en clon  $k$  ( $k=1,2$ ) para el replicado  $l$  ( $l = 1 - 3$ ),  $\mu$  es la intensidad media para el spot  $i$  en todos los geles,  $L$  es el efecto nivel,  $C$  es el efecto clonal,  $I$  es el efecto interacción entre el efecto nivel y el efecto clonal y  $\epsilon$  es el termino residual.

Para el gradiente de temporada 2006, un modelo similar, ANOVA de dos factores fue utilizado para fijar el efecto de los efectos principales (temporada y efecto clonal) y el efecto interacción utilizando el siguiente modelo:  $Y_{ijkl} = \mu + T_j + C_k + I_{lk} + \epsilon_{ijkl}$ , donde  $Y_{ijkl}$  es el valor normalizado del spot  $i$  ( $i=1-1052$ ) en el tiempo  $j$  ( $j = 1-4$ ) en clon  $k$  ( $k=1,2$ ) para el replicado  $l$  ( $l = 1 - 3$ ),  $\mu$  es la intensidad media para el spot  $i$  en todos los geles,  $T$  es el efecto tiempo,  $C$  es

el efecto clonal, I es el efecto interacción entre el efecto tiempo y el efecto clonal y  $\epsilon$  es el termino residual.

ANOVAs fueron realizados con el software R (Team, 2004). Spots con valores p (p-value) menores a  $10^{-3}$  se consideraron, como significativos para el efecto considerado.

### 3.3.5.2 Clustering

Luego de identificar las proteínas expresadas diferencialmente entre los distintos niveles. Datos centrados y normalizados, fueron analizados por clustering utilizando distancia Euclidiana, y el algoritmo “k-means”.

Para el gradiente base a copa 2003 se utilizó el software Expander (Sharan et al., 2003). Para el gradiente base a copa 2005 se utilizó el software GEPAS disponible en el sitio web (<http://gepas.bioinfo.cnio.es/>). Para el gradiente base a copa y temporada 2006 se utilizó EPClust disponible en el sitio web del “European Bioinformatics Institute” (<http://ep.ebi.ac.uk/EP/EPCLUST/>). Los diferentes softwares fueron utilizados con los mismos algoritmos y su diferencia radica en los gráficos que representan los resultados.

### 3.3.6 Identificación de proteínas por espectrometría de masas.

#### 3.3.6.1 Digestión de proteínas en el gel.

Con la ayuda de una punta de pipeta azul recortada, se corta desde el gel el área que contiene el spot. Las proteínas tenidas con Coomassie brilliant blue son lavadas con  $H_2O$ /Acetonitrilo (ACN)(50:50) hasta que la mancha pierda su coloración. La mezcla de solventes es removida y reemplazada por ACN. Luego de removido el ACN los trozos de gel fueron secados en una centrifuga al vacío. Los trozos de gel fueron rehidratados en

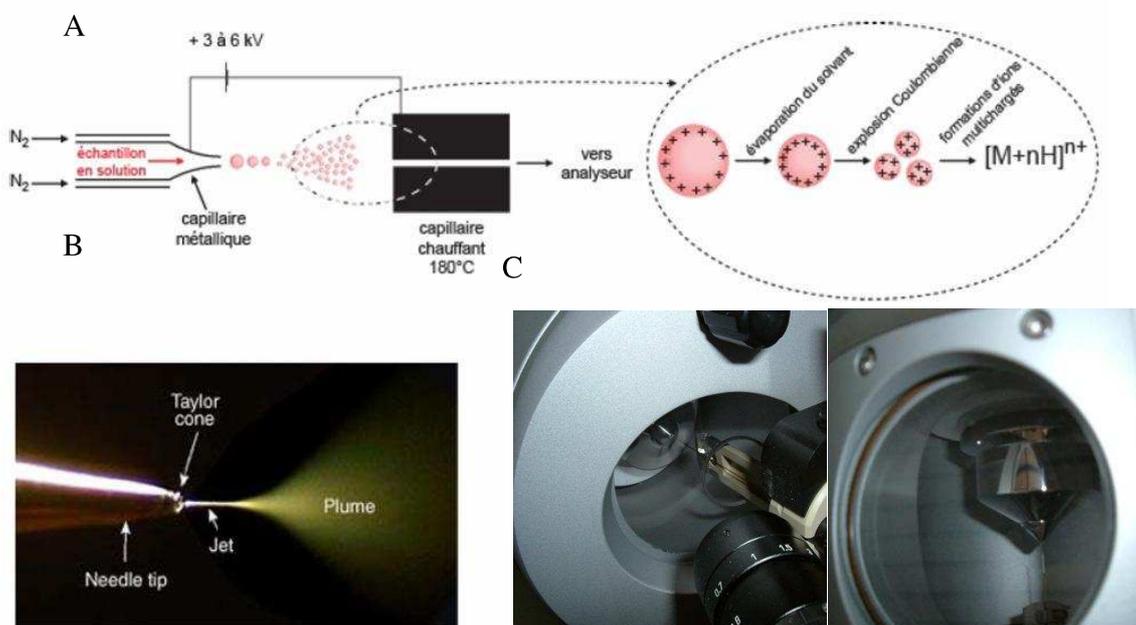


**Figura 20 :** De izquierda a derecha, inyector de muestra automático, bombas de HPLC y equipo de espectrometría de masas. El capilar de entrada al MS/MS se alinea con ayuda de una cámara (extremo superior derecho).

tripsina 10 ng/ml (Sigma, Aldrich, St. Louis, Mo EEUU) en  $\text{NH}_4\text{HCO}_3$  50mM e incubados toda la noche a 37°C. El sobrenadante fue removido y guardado a -20°C y los trozos de gel fueron incubados 15min en  $\text{NH}_4\text{HCO}_3$  50 mM a temperatura ambiente bajo agitación rotativa. Este segundo sobrenadante es mezclado con el anterior y una solución  $\text{H}_2\text{O}/\text{ACN}/\text{HCOOH}$  (47,5:47,5:5) fue agregada a los trozos de gel por 15 minutos. Este paso fue repetido dos veces. Los sobrenadantes fueron mezclados y concentrados en una centrifuga al vacío, hasta llegar a un volumen de 25µl. Los péptidos digeridos fueron finalmente acidificados con 1,2 µl de ácido acético y guardados a -20°C. La acidificación se realiza para facilitar la ionización en la posterior inyección nanoespray al equipo MS/MS.

### 3.3.6.2 Análisis capilar HPLC nanoespray trampa de iones MS/MS

Las mezclas de péptidos fueron analizadas por nano HPLC en línea (LC Packings, Amsterdam, Holanda) acoplado a un espectrómetro de masas electroespray LCQ Deca XP



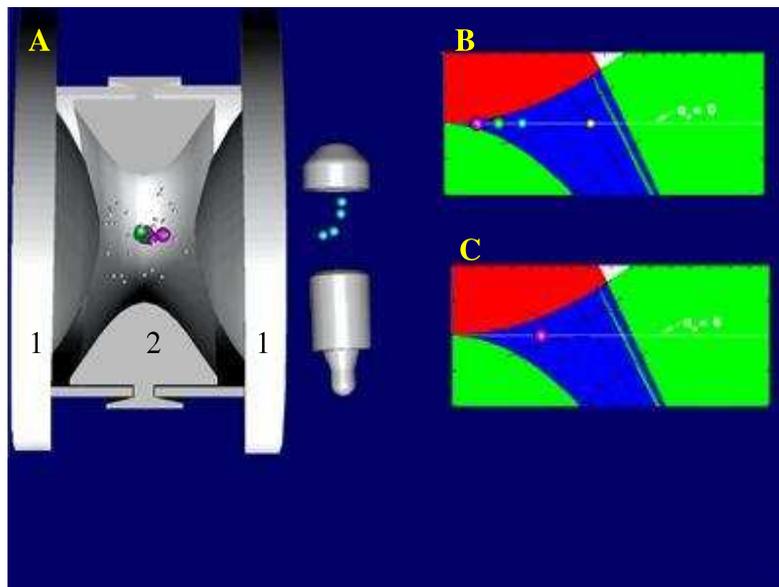
**Figura 21 :** El electroespray ;**A:** Esquema general, fuerzas electrostáticas provocan la pulverización de la muestra y el desplazamiento de la microgota ayuda a la evaporación del solvente. **B:** La muestra ingresa al equipo MS/MS como microgotas ionizadas, lo que aumenta la repulsión entre las microgotas de péptidos ionizados, **C:** Alineamiento del capilar.

con analizador de trampa de iones (**Figura 20**). Los péptidos fueron separados en una columna 75 $\mu$ m id x 15cm C18 (PepMap<sup>TM</sup> LC Packings, Ámsterdam, Holanda). Se utilizó un flujo de 200nl/min. Los péptidos fueron eluidos usando un gradiente lineal 5-50% de solvente B en 30min (Solvente A fue 0,1% ácido fórmico en 5% ACN, Solvente B fue 0,1% ácido fórmico en 80% ACN).

La espectrometría de masas electroespray (ESI-MS) es una técnica que utiliza alto voltaje para generar iones que ingresan al espectrómetro de masas desde un capilar de HPLC en línea (**Figura 21**). El líquido es cargado con alto voltaje en el capilar metálico, hasta alcanzar un punto crítico, donde no puede albergar mas carga y es despedido del inyector buscando una superficie de carga opuesta, en una nube de microgotas cargadas. Estas microgotas de 10  $\mu$ m de diámetro, evaporan el solvente en su desplazamiento, concentrando las cargas eléctricas que pueden volver a alcanzar un punto crítico y dividir la microgota nuevamente. El proceso de electroespray ha afectado profundamente el campo de la espectrometría de masas, ya que permite análisis estructurales de grandes macromoléculas y es compatible con métodos de cromatografía líquida (Fenn et al., 1990). El espectrómetro de masas fue operado en modo ion positivo a voltaje de aguja 1,8kV y voltaje capilar de 43V.

La trampa de iones es un analizador que comprende tres electrodos (**Figura 22A**) Los electrodos ubicados en sus extremos son los electrodos de tapa (**Figura 22A**) (1) y el electrodo del centro es el electrodo de anillo (**Figura 22A**) (2). Los tres electrodos forman un campo eléctrico que atrapa los iones al centro del analizador. El analizador de trampa de iones, es capaz de retener los iones en su interior y luego liberarlos hacia el detector (MS). En la **Figura 22B** y **C**, se comparan 2 gráficos de estabilidad de iones dentro de la trampa iónica. Se grafica  $a_z$  v/s  $q_z$  (parámetros de estabilidad de los iones en la trampa). La zona azul representa las condiciones donde los iones permanecen retenidos en la trampa. La trampa de iones es capaz de aislar un ion en particular (**Figura 22C**), para luego proceder a su fragmentación (MS/MS).

La adquisición de datos se realizó en un modo datos-dependiente, consistente en alternancia de una lectura exploratoria MS sobre el rango  $m/z$  300-1700, identificando todos los iones presentes (primer MS) y luego tres lecturas completas en un modo de exclusión dinámica, donde se elige el ion mas abundante, se aísla y fragmenta, este proceso se repite con el segundo



**Figura 22:** Trampa de iones. **A:** Diagrama de un corte transversal. **B** y **C** gráficos de la estabilidad iónica dentro de la trampa, esta puede ser regulada para aislar un ion en particular y fragmentarlo para conocer sus componentes (MS/MS).

ion mas abundante y con el tercer ion. Datos MS/MS fueron adquiridos usando una ventana de aislamiento de iones de  $2m/z$ , una energía relativa de colisión de 35% y una duración de exclusión dinámica de 0,5min.

El límite de detección del LC MS/MS, fue determinado analizando spots en un rango de intensidad de volumen 0,03-0,3. Una intensidad mínima de 0,05 fue confirmada como límite inferior de detección, para obtener al menos 2 péptidos, luego de la interrogación de la base de datos. La selección de los spots analizados se realizó luego de un análisis clustering, eligiendo spots de mayor intensidad en condiciones contrastantes y con una sobre expresión sostenida a lo largo de los gradientes.

Los péptidos fueron identificados con el programa Sequest utilizando la interfase Bioworks 3.2 (Thermo-Finigan, Torrence, CA, EEUU), usando 45.934 TCs (contigs tentativos) de el índice "TIGR Pinus Gene Index". La identificación de TCs corresponde a la versión de marzo 2006.

### 3.3.7 PCR en tiempo real (qPCR)

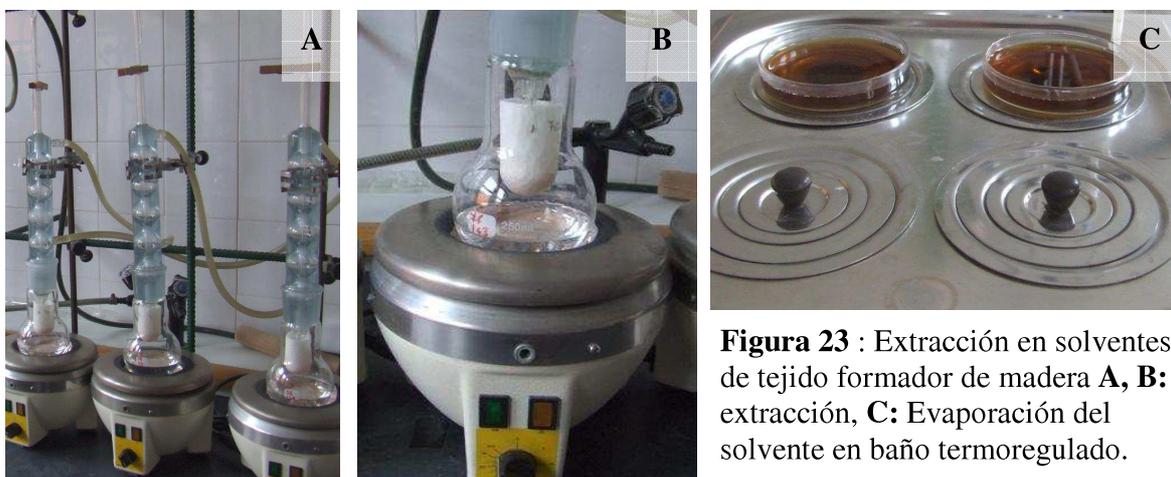
Los partidores fueron diseñados usando el software Primer 3 (Rozen and Skaletsky, 2000), para 14 genes, así como para dos controles correspondientes, una proteína ribosomal (BX252550) y un gen de profilina I (BX249454). Los dos controles fueron seleccionados en base a una prueba previa de 10 genes en diferentes tipos de tejido formador de madera colectados a lo largo del tronco (gradiente base a copa) y a través de la temporada (gradiente de madera temprana a tardía). Los partidores fueron diseñados para tener un tamaño óptimo de 22pb (18-24pb), contenido de GC de 40-60% y Tm de 58-62°C. Otros criterios como auto-alineamiento del partidador fueron tomados en cuenta. El rango de los fragmentos esperados fue de entre 103 a 228pb. Los oligonucleótidos fueron sintetizados por Eurogentec (Liege, Bélgica).

**Tabla 4:** Partidores diseñados para qPCR

<b>Base a copa</b>			
<b>Gene function</b>	<b>Primer F</b>	<b>Primer R</b>	<b>EST</b>
Putative 111	ATGTCTAATTTTCATGTGCCTTTTG	TTTTTGACACGGCTCAATTC	BX253111
Putative 778	CTCCAGATCCCTCTCTGCAC	GATCAGCCGATGTTGATTCC	BX251778
Glutamate-ammonia ligase	AAGTGGCTGGTGTGTCTC	TCTCCATAGGCAGCAATGTG	BX253698
Tubulin beta chain	CAACTTCGTTTCGAGCCTTC	ACTCCCGGGATAAGTTGGTC	BX249177
Putative 775	ATCTTCGTCGCCTTTGTTTC	ATCCCGCATGAGTTCAAAG	BX249775
Metallothionein-like protein	AACTGTACCTGCGACCCTTG	ACCTCTGACGCAGAGAGAGC	BX249412
GRP2	GTTGGGTGGGGTGAAGATAA	CCGTTCCGAGCAGAGC	BX000600
Low HSP	TTCTGCCGTAGCCAATACTC	TGCTCGCCTTATCTCTTCC	BX000656
Arabinogalactan/ proline-rich protein	ATTCCCGTTTGAATCAGCAG	TCGGTCGTAGTTTCGATCTC	BX249981
Fructose bisphosphate aldolase	CTCTTCACTGCCCTAATGC	TATTGCTGGCATCTCTGTGC	BX249425
ATH1	CAACGGGCTCTCCAACAA	CCTCCACATACATCTTCTACCA	BX249908
Peptidyl isomerase	AGGATGGGGGCATTTTAAAG	AGCTATGGCAACAGCTGGAC	BX251374
<b>Temprana a madura</b>			
<b>Gene function</b>	<b>Primer F</b>	<b>Primer R</b>	<b>EST</b>
CLAVATA1	GCAGATACCAGAGCAAAAATG	TCCTCGTTCGTTGAGTTG	BX249897
Cellulose synthase (UDP forming)	AAATGGGTGGTGTTCAGAG	GGAGCTGATCCTTGAATGC	BX249248
17.6kDA class I HSP	AAAATTTCTGCGTTTGTGG	CTTGATCGGCTACCGAAGAG	BX251102
A gamma-theonin precursor defensin	GGAAGTTCAAGTGGCAGAGG	TGCAGTAGTAGCGTTGCAG	BX254131
Putative 476	GAGGAAGTTTATCCCATCG	TGGCATATGTGGAGTTGTC	BX250476
GASA52	TACCTGGATGGAGCAAAAGC	ATAAGTCCCGGAGGAACAC	BX250252
Putative 990	CGAAATTGCGCTTAAGGTTTC	TGCAGATGCGTAGTTTCTG	BX252990

Un microgramo de RNA total fue transcrito reverso, utilizando el sistema de transcripción reversa ImProm-II (Promega, Madison, WI, EEUU) de acuerdo a las instrucciones del fabricante. Dos transcripciones reversas independientes fueron realizadas y cada una fue diluida 10X antes de cada reacción qPCR. La reacción de qPCR se realizó en el Chromo4 Multicolor Real-Time PCR Detection System (Bio-Rad Laboratorios, Inc. Hercules, CA EEUU), agregando 10 µl de IQ SYBR Green Supermix (Bio-Rad Laboratorios, Inc. Hercules, CA EEUU), 3 µl de cDNA diluido, 3 pmol de cada partidor, y agua (volumen final 20 µl). Luego de una incubación inicial de 95°C por 3min, las amplificaciones se realizaron por 40 ciclos con el siguiente programa: un paso de denaturación de 95°C por 15seg, seguido por un paso de extensión/alineamiento de 60°C por 45 seg, según lo recomendado por los fabricantes. Luego de la amplificación, una curva de “melting” fue obtenida para verificar la formación de dímeros de partidor y otros productos no específicos de la amplificación.

Para realizar la cuantificación relativa de los productos de amplificación, valores de Ct fueron obtenidos para cada muestra y los datos fueron analizados usando la macro GENEX v 1,10 (Bio-Rad Laboratorios, Inc. Hercules, CA EEUU), para Excel (Microsoft). Para cada gen cuantificado, la eficiencia promedio del PCR fue calculado automáticamente.



**Figura 23** : Extracción en solventes de tejido formador de madera **A, B**: extracción, **C**: Evaporación del solvente en baño termoregulado.

### 3.3.8 Caracterización Química

Muestras libres de extractivos fueron caracterizadas por espectroscopía Infrarroja con Transformada de Fourier (FTIR) y pirolisis analítica.

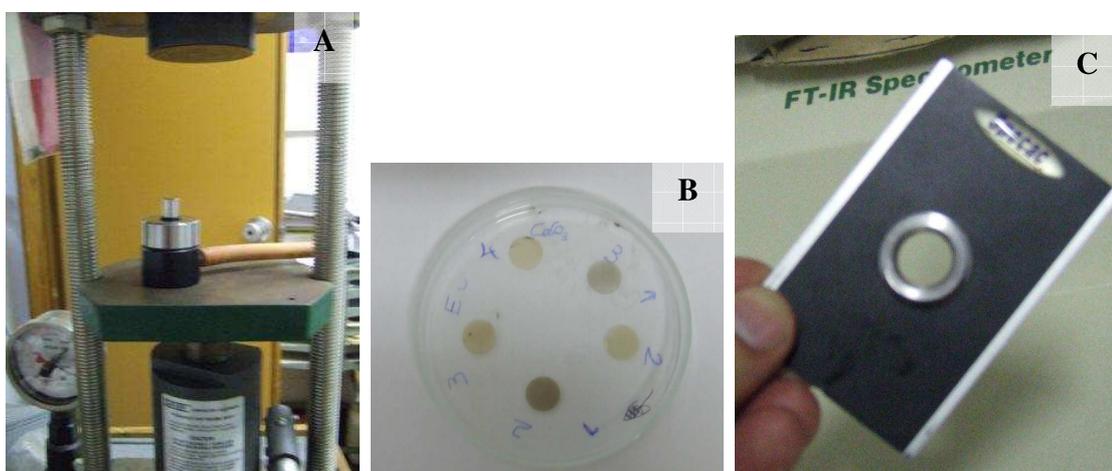
### 3.3.8.1 Preparación de muestras

Las muestras fueron congeladas y deshidratadas por una semana y pulverizadas en mortero. Una alícuota de cada muestra fue entonces extraída secuencialmente con diclorometano, metanol y agua. Un método de extracción rápida fue utilizado. Las muestras fueron extraídas en un receptáculo de extracción durante 30 minutos, en contacto directo con el solvente hirviente (extracción) y luego el receptáculo fue levantado por una hora por sobre el solvente (lavado) (**Figura 23 A, B**). El solvente fue evaporado (**Figura 23 C**) y el residuo fue cuantificado.

### 3.3.8.2 Análisis de pared celular

#### 3.3.8.2.1 FTIR Espectroscopia infrarroja por Transformada de Fourier

Aproximadamente 1,5mg de muestra finamente pulverizada fue mezclada con 200mg de KBr grado espectrópico en una pulverizadora de esferas durante 20s. Un dispositivo estándar para



**Figura 24** : Espectrometría FTIR; A: prensa para preparar pellets B: pellets C: Pellet montado en lector para ser insertado en el espectrómetro FT-IR.

preparar pellets (**Figura 24A**) fue utilizado para preparar pellets de 13mm (**Figura 24B**). El espectro fue registrado usando un espectrómetro infrarrojo Biorad FTS 165, con 64 scans por muestra, con una resolución de  $4\text{cm}^{-1}$ . Cada espectro fue ajustado en contra una línea base obtenida inmediatamente antes de medir las muestras.

Los espectros fueron obtenidos con un mínimo de  $2000\text{ cm}^{-1}$ , se utilizó un proceso de normalización "Min-Max". La normalización Min Max, escala la intensidad del espectro, teniendo como referencia mínima 0 (a  $1925\text{ cm}^{-1}$ ) y como máximo 2 (en la banda de polisacáridos a  $1059\text{ cm}^{-1}$ ). Esta normalización se realizó entre  $2000\text{ cm}^{-1}$  y  $400\text{ cm}^{-1}$ .

### 3.3.8.2.2 Pirolisis Analítica

Se efectuó Pirólisis Analítica (Py-GC/FID), utilizando una pirosonda CDS Pyroprobe 1000 con filamento, (**Figura 25B**) conectado a un cromatógrafo de gas HP 5890 series II conectada por una interfase calentada a  $270^{\circ}\text{C}$  (**Figura 25C**). Cada muestra ( $75\text{-}80\mu\text{g}$ ) fue pirolizada a  $650^{\circ}\text{C}$  por 10s con un aumento de temperatura de aproximadamente  $20^{\circ}\text{C}\text{ms}^{-1}$ . Columna capilar: DB1701 ( $60\text{m} \times 0.25\text{mm}$ ,  $0.25\mu\text{ film}$ , J&W Scientific). Las condiciones de la CG fueron: inyector  $250^{\circ}\text{C}$ , detector  $280^{\circ}\text{C}$ , programa de temperatura:  $45^{\circ}\text{C}$ , isothermal por 4min, luego  $4^{\circ}\text{C}\cdot\text{min}^{-1}$ , hasta  $270^{\circ}\text{C}$ . La identificación de los productos de pirolisis fue realizada utilizando muestras ya caracterizadas, por Py-GC/MS (CDS Pyroprobe 1000 conectado a HP 6890 con selector de masas HP 5973). Los productos fueron identificados por su espectro de masas y tiempo de retención, por comparación con la Liberia NITS y con literatura (Faix et al., 1991a) (Faix et al., 1991b) (Ralph and Hatfield, 1991). Se efectuaron por lo menos dos corridas por muestra. La cuantificación de los productos de pirolisis fue calculado con el programa Chemstation (Agilent Technologies, Palo Alto , EEUU).



**Figura 25 :** Pirólisis analítica, **A :** Pirosonda acoplada a GC **B :** Capilar con muestra instalado en pirosonda; **C:** Pirosonda encendida.

## 4. RESULTADOS Y DISCUSION

#### **4.1 Análisis fenotípico y molecular de un gradiente base a copa**

En el presente capítulo se comparan estudios fenotípicos, como análisis morfológicos en tejido formador de madera en un gradiente base a copa, análisis químicos de componentes de la pared celular y estudios genómicos, de transcriptómica y proteómica, para el gradiente base a copa, en tres temporadas 2003, 2005 y 2006. En los análisis proteómicos, se utilizaron modelos estadísticos independientes para cada año, identificando proteínas significativamente sobreexpresadas (mas de 2 “fold change” o veces de cambio) la entre los niveles de madera de base y madera de copa. Se ha encontrado solo una referencia (Cato et al., 2006) que ha tratado de comparar datos fenotípicos (grosor de pared celular, crecimiento radial y la proporción de los tipos celulares en xilema en formación), con expresión génica en tejido formador de madera de base y copa, en Pino radiata.

Se realizó la caracterización morfológica del tejido formador de madera recolectado, según la técnica de recolección utilizada en las temporadas 2003, 2005 y 2006 (Sección 4.1.1). Igualmente se compararon las condiciones ecofisiológicas de la parcela donde se recolectaron las muestras, para las temporadas 2003, 2005 y 2006 (Sección 4.1.2).

La sección 4.1.3 describe el análisis químico de los componentes de la pared celular en un gradiente base a copa en la temporada 2003 y 2006.

La sección 4.1.4 describe el análisis fenotípico y molecular de un gradiente base a copa en tejido formador de madera en pino marítimo, realizado la temporada 2003, en este estudio fueron integrados análisis proteómicos, químicos y microscópicos, con estudios transcriptómicos realizados por Paiva y colaboradores. El análisis proteómico identificó 33 proteínas sobreexpresadas en los niveles base o copa. Los resultados son resumidos en un artículo publicado en *New Phytologist* 178 (2), 283–301(2008).

La sección 4.1.5 reporta el análisis proteómico de un gradiente base a copa para la temporada 2005, 67 proteínas sobreexpresadas fueron identificadas.

La sección 4.1.6 resume los análisis proteómicos de muestras contrastantes de madera de copa y base para la temporada 2006. En esta temporada fueron identificadas 65 proteínas que explican su variación solo por el factor nivel.

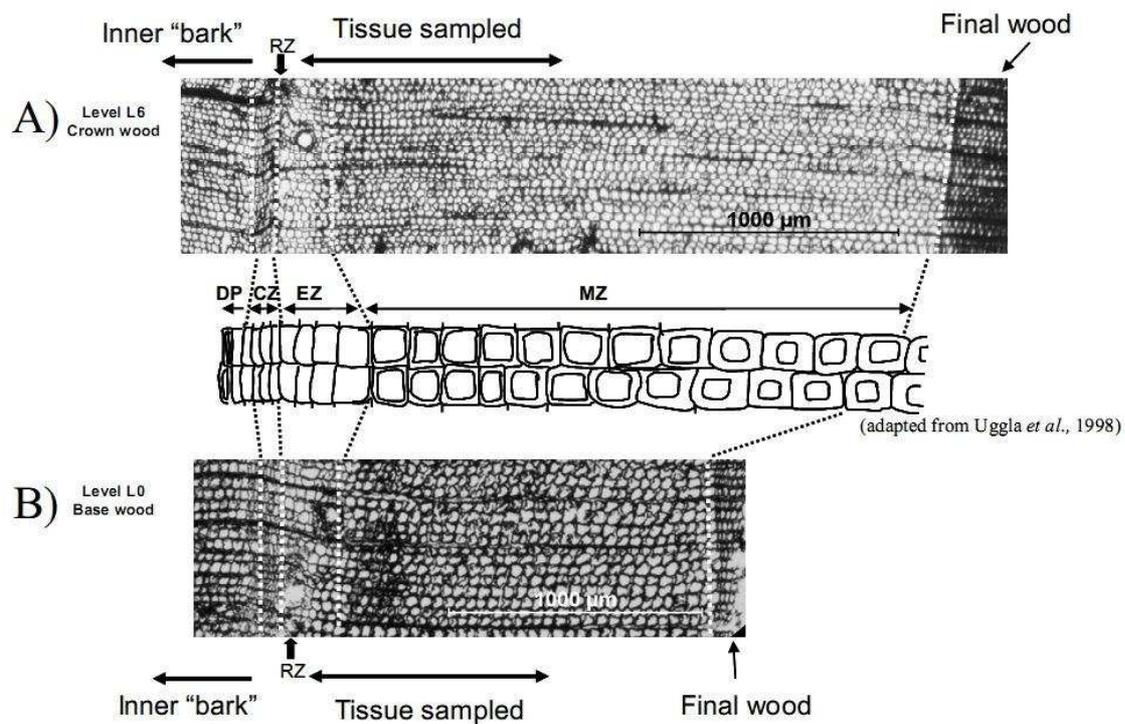
La sección 4.1.7 analiza la asignación de categorías funcionales a las proteínas identificadas y la confiabilidad de los datos.

En la sección 4.1.8 los resultados son discutidos en conjunto, comparando las características fenotípicas y las categorías funcionales de las proteínas sobreexpresadas en tejido formador de madera de base o copa.

Finalmente en la sección 4.1.9 se describe la base de datos disponible en Internet donde se encuentran los experimentos proteómicos descritos en este trabajo.

#### **4.1.1 Caracterización anatómica**

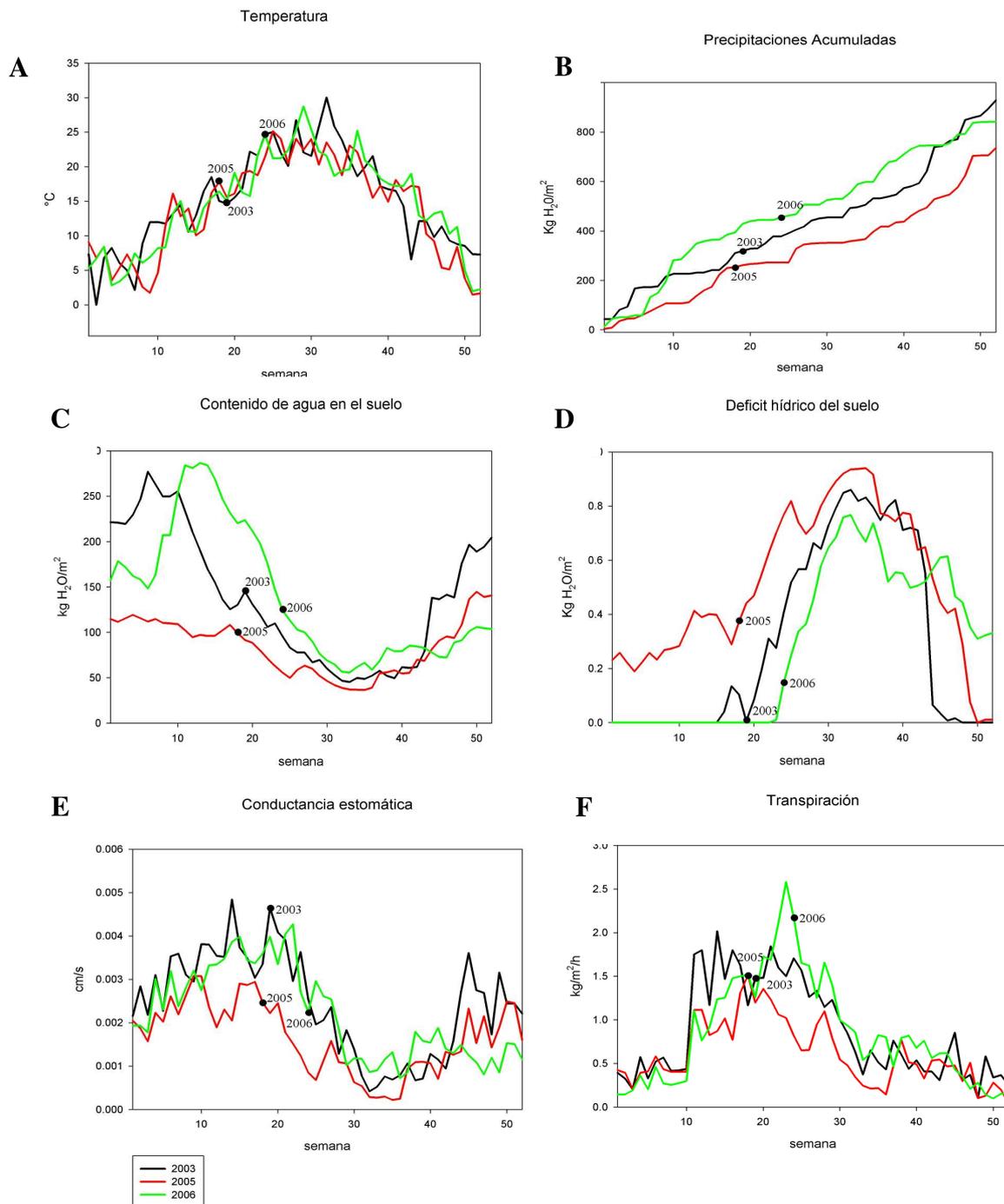
La transición de madera de copa o juvenil a madera de base o madura, se ha estimado en base a variaciones en el perfil de densidad de la madera, y ocurre alrededor del anillo número 10 y 12 (Zobel, 1972; Dumail et al., 1998). Considerando la edad de las muestras recolectadas a través del gradiente (3-22 años), el gradiente analizado debe mostrar en sus extremos las características típicas de madera de copa y madera de base. Para comprobar que las muestras extremas colectadas en la base y copa tienen las propiedades comúnmente reportadas para estos tipos de madera, se realizaron análisis morfológicos de fibra y de microscopía. En el análisis morfológico de las fibras, se obtuvieron diferencias significativas entre las muestras colectadas en la copa y en la base, en el largo promedio (410  $\mu\text{m}$  vs 1020  $\mu\text{m}$ , respectivamente) y ancho promedio (27,9  $\mu\text{m}$  vs 31,2  $\mu\text{m}$ , respectivamente). Estos datos cumplieron la expectativa de encontrar fibras más cortas y delgadas en madera de copa, en comparación con madera de base. Con respecto a las observaciones microscópicas (**Figura 26**), podemos decir que (i) el número de traqueidas producidas por el cambium en la copa es aproximadamente el doble que las traqueidas producidas en la base. (ii) el largo de cada zona correspondiente a diferentes etapas de xilogénesis secundaria es diferente entre ambas muestras y (iii) tejidos que no fueron colectados (en la base y en la copa) incluyen floema secundario con cambium adherido así como células de xilema en etapas de diferenciación temprana. El tejido expuesto colectado comprendió células en diferenciación adheridas débilmente, células en expansión celular (EZ en **Figura 26**) y traqueidas en maduración (MZ en **Figura 26**). Aunque el largo de la zona de expansión (EZ) no fue significativamente diferente entre la copa y la base del árbol, el número promedio de células sí varió significativamente entre copa y base (8,1 vs. 5,8, respectivamente;  $t$ -test  $P < 0,05$ ). El largo de la zona de maduración fue claramente diferente ( $t$ -test  $P < 0,05$ ) entre ambas muestras (2316  $\mu\text{m}$  en madera de copa y 1344  $\mu\text{m}$  en madera de base), como también el número de células (60,8 en madera de copa y 29,3 en madera de base,  $t$ -test  $P < 0,05$ ), y el radio celular promedio (11  $\mu\text{m}$  en madera de copa vs. 15  $\mu\text{m}$  en madera de base,  $t$ -test  $P < 0,05$ ).



**Figura 26:** Observaciones anatómicas de tejido formador de madera A)Madera de copa; B)Madera de base. Se muestra la porción de tejido recolectado: "Tissue sampled" y las diferentes regiones de la xilogénesis DP: Floema en división, CZ :Zona cambial, EZ: Zona de expansión, MZ : Zona de maduración. La madera de la temporada interior se etiqueta como "Final Wood".

#### 4.1.2 Comparación de datos ecofisiológicos, para las temporadas 2003, 2005 y 2006

La recolección de muestras para el gradiente base a copa se realizó el 12 de mayo, en la temporada 2003, el 6 de mayo para la temporada 2005 y el 19 de junio para la temporada 2006, correspondiendo a las semanas del año 19, 18 y 24 respectivamente. La **Figura 27** grafica las condiciones ecofisiológicas previas al momento de tomar la muestra en las tres temporadas. La fecha de recolección del xilema en formación es indicada en cada gráfico por un círculo negro. Durante las tres temporadas la temperatura promedio fue similar (**Figura 27A**), a lo largo del ciclo boreal de crecimiento. En la **figura 27B**, se muestra la precipitación acumulada hasta la fecha de recolección de muestra, el año 2005 fue el mas seco, seguido del 2003 y el 2006, con 253,5, 319,5 y 455,5 kgH<sub>2</sub>O/m<sup>2</sup>, respectivamente. El contenido de agua en el suelo, fue menor para la temporada 2005 desde el inicio de la temporada (**Figura 27C**), esto contribuyo a un aumento del déficit hídrico del suelo (**Figura 27D**), que estuvo presente durante el periodo anterior a la recolección de muestra el 2005, en contraste con las temporadas 2003 y 2006, donde las muestras fueron recolectadas apenas iniciado el déficit hídrico. La conductancia estomática y la transpiración también fueron menores para la temporada 2005 entre las semanas 10 y 18, confirmando que los árboles estuvieron sometidos a estrés hídrico antes de la colección de muestra esta temporada.



**Figura 27:** Datos ecofisiológicos anuales para temporadas 2003, 2005 y 2006. A: Temperatura, B: Precipitaciones Acumuladas, C: Contenido de agua en suelo, D: Déficit hídrico de suelo, E: Conductancia estomática, F: Transpiración.

#### 4.1.3 Composición química de la pared celular en un gradiente de base a corona.

La madera de copa/juvenil ha sido caracterizada y descrita extensamente (Maeglin, 1987; Zobel and Van Buijtenen, 1989; Zobel and Sprague, 1998; Larson et al., 2001). En comparación a la madera de base/madura, posee traqueidas más cortas, menor diámetro celular y células con pared más delgada, junto con esto, la madera de copa presenta menor gravedad específica. Estas características resultan en productos madereros débiles, papel con menor resistencia a la tensión y mayor resistencia al rasgado (Zobel and Van Buijtenen, 1989). La composición química también es diferente, la madera de copa contiene un mayor porcentaje de lignina y hemicelulosas y un menor porcentaje de celulosa que la madera de base, lo que incide en un mayor costo en la extracción de celulosa desde madera de copa y un mayor impacto ambiental (Zobel and Van Buijtenen, 1989). En pino marítimo, la composición química promedio de la madera es: celulosa 42%, lignina 30%, hemicelulosas 28%, incluyendo mananos 13%, xilanos 9% y galactanos 4% (basado en 2500 muestras de madera, Paiva 2006)

El análisis químico del gradiente copa a base analizado en la temporada 2003, mostró una mayor cantidad de proteínas (aa1) asociada a madera de copa (**Fig 1 Paiva et al.**). Igualmente en las muestras contrastantes de la temporada 2006, se calculó un porcentaje de proteínas de  $0,7\pm 0,0\%$  en la base y un  $1,0\pm 0,2\%$  en la copa.

Ligninas: la unidades de lignina H producen fenol y cresol, productos de pirólisis que también pueden derivar de tirosina (Faix et al, 1991), de hecho las unidades G siguen el mismo perfil que las proteínas, por esta razón la lignina fue estimada utilizando los productos de pirólisis originados por unidades G. Los productos de pirólisis derivados de unidades G, mostraron un perfil irregular en la temporada 2003 (**Figura 1 Paiva et al.**). Igualmente en la temporada 2006, las diferencias no fueron estadísticamente significativas, se cuantificó un  $13,4\pm 2,17\%$  en la base y un  $11,1\pm 1,48\%$  en la copa.

Las hemicelulosas son las principales contribuyentes a los productos de pirólisis “carbohidratos distintos a hexosanos y pentosanos” (c) y “Pentosanos” (cP), ambos tipos de productos, mostraron una tendencia a aumentar desde base a copa durante el 2003 (**Figura 1 Paiva et al.**), igualmente en la temporada 2006 se cuantificaron los Pentosanos (cP)  $3,2\pm 0,1\%$  en la base y un  $4,1\pm 0,3\%$  en la copa. En la temporada 2006, Carbohidratos distintos a hexosanos y pentosanos (c) mostraron una variación en torno al 40%, sin diferencias significativas. Juntos, los carbohidratos “distintos a hexosanos y

pentosanos” (c) y los Pentosanos (cP) indican que el contenido de hemicelulosas aumentan levemente desde base a copa.

El mayor contenido de celulosa (cH) se encontró en madera de base y disminuyó hasta la madera de copa (**Figura 1 Paiva et al.**).

El agrupamiento de los análisis químicos de la temporada 2003 (**Figura 3 Paiva et al.**), demostró que los niveles asociados a madera de copa (L5 y L6), agruparon junto a un mayor contenido de hemicelulosas (c y cP), aminoácidos (aa1) y un menor contenido de celulosas (aa1). Los niveles asociados a la base del árbol (L1-L4) mostraron una mayor cantidad de celulosa. Tomando en cuenta los análisis químicos efectuados el 2003 y 2006 se puede concluir que los tejidos analizados responden a las características típicas de madera de copa y madera de base, sin embargo hay un fuerte impacto del factor clonal y ambiental.

#### **4.1.4 Análisis molecular y fenotípico de un gradiente base a copa en tejido formador de madera en pino marítimo. Temporada 2003**

Los resultados de este gradiente, analizado el año 2003, se incorporaron al artículo: **“Molecular and phenotypic profiling from the base to the crown in maritime pine wood-forming tissue.** Jorge Paiva, **Marcelo Garcés**, Ana Alves, Pauline Garnier-Géré, José Carlos Rodrigues, Céline Lalanne, Stéphane Porcon, Grégoire Le Provost, Denilson da Silva Perez, Jean Brach J, Jean-Marc Frigerio, Stéphane Claverol, Aurélien Barré, Pedro Fevereiro, Christophe Plomion”.

**New Phytologist 178 (2), 283–301(2008)**

doi: 10.1111/j.1469-8137.2008.02379.x.

ISI Journal Citation Reports® Ranking: 2006: 8/147 (Plant Sciences)

Impact Factor: 4.245

# Molecular and phenotypic profiling from the base to the crown in maritime pine wood-forming tissue

Jorge A.P. Paiva<sup>1,2,3</sup>, Marcelo Garcés<sup>1,4</sup>, Ana Alves<sup>3,5</sup>, Pauline Garnier-Géré<sup>1</sup>, José Carlos Rodrigues<sup>3,5</sup>, Céline Lalanne<sup>1</sup>, Stéphane Porcon<sup>1</sup>, Grégoire Le Provost<sup>1</sup>, Denilson da Silva Perez<sup>6</sup>, Jean Brach<sup>1</sup>, Jean-Marc Frigerio<sup>1</sup>, Stéphane Claverol<sup>7</sup>, Aurélien Barré<sup>8</sup>, Pedro Feveireiro<sup>2,9</sup> and Christophe Plomion<sup>1</sup>

<sup>1</sup>INRA, UMR 1202, Biodiversity Genes and Communities, 69 route d'Arcachon, F-33610 Cestas, France; <sup>2</sup>Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. da República-EAN, 2780-157 Oeiras, Portugal; <sup>3</sup>Tropical Research Institute of Portugal (IICT), Forest and Forest Products Centre, Tapada da Ajuda, 1349-017 Lisboa, Portugal; <sup>4</sup>Instituto de Biología Vegetal y Biotecnología, Universidad de Talca, Chile; <sup>5</sup>Centro de Estudos Florestais, Departamento de Engenharia Florestal, Instituto Superior de Agronomia, ISA-DEF, Tapada Ajuda, 1349-017 Lisboa, Portugal; <sup>6</sup>FCBA InTechFibres, Laboratoire Bois Process, Domaine Universitaire, BP 251, 38044 Grenoble, Cedex 9, France; <sup>7</sup>Proteomic Facility, Université Bordeaux 2, 33076 Bordeaux, France; <sup>8</sup>Centre de Bioinformatique Bordeaux, Université Victor Segalen Bordeaux 2, rue Léo Saignat, 33076 Bordeaux Cedex, France; <sup>9</sup>Departamento de Biología Vegetal, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1700 Lisboa, Portugal

## Summary

Author for correspondence:  
Christophe Plomion  
Tel: +0 33 5 57 12 28 38  
Fax: +0 33 5 57 97 90 88  
Email: [plomion@pierroton.inra.fr](mailto:plomion@pierroton.inra.fr)

Received: 29 August 2007  
Accepted: 15 December 2007

- Environmental, developmental and genetic factors affect variation in wood properties at the chemical, anatomical and physical levels. Here, the phenotypic variation observed along the tree stem was explored and the hypothesis tested that this variation could be the result of the differential expression of genes/proteins during wood formation.
- Differentiating xylem samples of maritime pine (*Pinus pinaster*) were collected from the top (crown wood, CW) to the bottom (base wood, BW) of adult trees. These samples were characterized by Fourier transform infrared spectroscopy (FTIR) and analytical pyrolysis. Two main groups of samples, corresponding to CW and BW, could be distinguished from cell wall chemical composition.
- A genomic approach, combining large-scale production of expressed sequence tags (ESTs), gene expression profiling and quantitative proteomics analysis, allowed identification of 262 unigenes (out of 3512) and 231 proteins (out of 1372 spots) that were differentially expressed along the stem.
- A good relationship was found between functional categories from transcriptomic and proteomic data. A good fit between the molecular mechanisms involved in CW–BW formation and these two types of wood phenotypic differences was also observed. This work provides a list of candidate genes for wood properties that will be tested in forward genetics.

**Key words:** analytical pyrolysis, Fourier transform infrared spectroscopy, *Pinus pinaster*, proteome, transcriptome, wood formation.

*New Phytologist* (2008) **178**: 283–301

© INRA (2008). Journal compilation © *New Phytologist* (2008)  
doi: 10.1111/j.1469-8137.2008.02379.x

## Introduction

Trees are long-living organisms with the characteristic feature of forming wood. Wood confers mechanical strength and a long-distance pathway for water, minerals and hormones. Wood properties vary spatially within a single tree. The transition

between juvenile/crown wood (J/CW) to mature/base wood (M/BW) constitutes an important source of variation in wood traits (Zobel & van Buijtenen, 1989; Zobel & Sprague, 1998; Larson *et al.*, 2001). This transition has been defined for the variations occurring both from the pith to the bark and from the apex to the base of the tree stem.

Mature/base wood differs from J/CW in that it has thicker cell walls, narrower cell lumens, larger cellulose microfibril angles, larger spiral-grain angles and higher specific density. Juvenile/crown wood can occasionally present disproportionate amounts of compression wood, distorted grain patterns and pith deposits (Larson *et al.*, 2001). In terms of chemical composition, M/BW shows higher cellulose and lower lignin contents (reviewed by Zobel & Sprague, 1998). The anatomical, structural and chemical characteristics of J/CW adversely affect solid wood product performances (e.g. strength, stiffness and warping on drying) as well as pulp and paper manufacture (e.g. yield, tearing strength and bleaching) (Zobel & van Buijtenen, 1989; Zobel & Sprague, 1998). Owing to economic factors, there are strong trends towards the decrease of harvest age in industrial forest plantations. Together with faster growth rates and changing silvicultural practices, these factors will lead to a higher proportion of J/CW in harvested trees. Thus, it is of great importance to understand the molecular machinery involved in J/CW and M/BW formation, and to provide genetic markers for improving and manipulating wood quality.

Wood cell division and differentiation occur at the vascular cambium, the meristematic tissue found between the xylem and phloem. During wood formation, new xylem cells are produced by division of cambial initials (fusiform and ray initials). Cambial maturation (cell expansion, cell wall thickening and programmed cell death) results in changes of cell dimension and structure (e.g. orientation of microfibrils in the  $S_2$  layer of the cell wall), and chemical composition (e.g. cellulose and lignin content), which in turn affect the final anatomical, structural and chemical composition of fully developed wood.

The maintenance of a strict balance between the number of meristematic cells (cambium initials) and the programmed differentiation of their progeny (rays and tracheids) is critical for meristematic function (Nakajima & Benfey, 2002). Strong evidence has been found that key genes originally characterized as those regulating the meristematic cells of the shoot apical meristem are also expressed in the vascular cambium during woody growth (reviewed by Groover, 2005). At the top of a tree or at the end of branches, new secondary cambium is produced each year from the apical meristem. Simultaneously, at the tree stem base, the secondary cambium is maintained over the years by a delicate balance between cell proliferation and differentiation. These different cambial origins and locations result in a cambial age gradient along the tree stem; that is, at the apices of the stems and branches, secondary cambium tissues are younger than those located at the base of the stem.

The different ages of secondary cambium tissues along an adult tree stem, combined with the effects of genetic and environmental factors, for example, proximity to the crown (Larson *et al.*, 2001), are likely to affect the rate and duration of cell division and cambial maturation, and thus result in differences in anatomical structure and xylem composition.

At the beginning of the growing season, the reactivation of cambial growth in mature pine trees occurs simultaneously throughout the stem (Savidge & Wareing, 1984; Sundberg *et al.*, 1991). However, there is a high degree of variation in the rate of cell division and cambial maturation along the stem. Uggla *et al.* (1998) reported that cambial growth, measured by the radial number of tracheary derivatives produced following the start of the growing season, was greater within the crown, and decreased down the stem. They also showed that this pattern was stable over several years. Recently, Cato *et al.* (2006) found that the rate of cell division was 3.3 times higher in cambium located in the crown than at the tree base, and that this faster growth was associated with reduced cell wall thickness. They also found that the proportion of cells undergoing secondary thickening in the crown was significantly lower than at the base. Cato *et al.* (2006) also showed that genes involved in cell division and expansion tended to be more highly expressed in the tree crown, and two putative cell-cycle repressors (with sequence homology with a putative PREG-like negative regulator, and a putative CDC48 repressor protein) were expressed twofold higher in the base. They found that transcripts involved in secondary cell wall thickening were more abundant at the base.

Pines are considered to be an excellent model species to study wood formation (Lev-Yadun & Sederoff, 2000). Despite their relatively simple xylem structure (mainly tracheids) within a single tree genotype, six contrasting types of wood can be identified: compression wood vs opposite wood, early wood vs late wood and J/CW vs M/BW. The plasticity of molecular mechanisms associated with the formation of these different types of wood has been studied in pines mainly at the transcriptome level (Whetten *et al.*, 2001; Lorenz & Dean, 2002; Le Provost *et al.*, 2003; Egersdotter *et al.*, 2004; Yang & Loosptra, 2005; Cato *et al.*, 2006). RT-PCR has been used to confirm the results of over 350 genes that were differentially expressed between wood-forming tissues associated with early vs late wood, compression vs opposite wood, and juvenile (JW) vs mature wood (MW). These transcripts included genes involved in cell division, expansion and cell wall thickening. However, it is now necessary to carry out research on the proteome of wood-forming tissue (Plomion *et al.*, 2000; Gion *et al.*, 2005).

In contrast to fully developed wood, wood-forming tissues associated with different types of wood have generally not been characterized at the phenotypic level. Moreover, studies aimed at characterizing molecular modifications have often been disconnected from phenotypic characterization of corresponding wood-forming cells. In only a handful of studies has phenotypic characterization of wood-forming tissues been combined with gene expression analyses. For example, in *Populus*, Andersson-Gunneras *et al.* (2006) studied the metabolomic changes together with gene expression profile during the formation of tension wood, and in *Pinus radiata*, Cato *et al.* (2006) measured tracheid cell wall thickness,

radial growth and gene expression of wood-forming tissues associated with both BW and CW along the tree stem.

The objectives of this study were twofold: (i) to describe the variation of cell wall chemical composition in wood-forming tissues along the stem; and (ii) to study the variation in the transcriptome and proteome of differentiating xylem collected on the same samples.

## Materials and Methods

### Tissue sampling

A total of seven samples were collected along a cambial age gradient on one straight 30-yr-old tree. Samples were collected on the 12 May 2003 at internodes of years 1982 (level 0, L0), 1985 (level 1, L1), 1988 (level 2, L2), 1991 (level 3, L3), 1994 (level 4, L4), 1997 (level 5, L5), and 1999/2000 (level 6, L6). Levels L1 to L6 were used to describe the variation of cell wall chemical composition, and for transcriptome analyses. Levels L0, L2, L4, and L6 were used to describe the proteome along the tree.

Two additional straight trees (A and B) were sampled in 2006, at the same date and location as that of the tree sampled in 2003. This additional sampling was conducted in order to study the extent to which the results obtained in 2003 were related, or not, to other genotypes and climatic conditions. Samples L0 (internode 1982) and L6 (internode 1999/2000) were characterized in terms of cell wall chemical composition. Samples L1 (1985), L2 (1988), L5 (1997), and L6 (1999/2000) were used to quantify gene expression (using qPCR) for 14 genes showing differential expression along the stem of the tree sampled in 2003.

These three mature trees were sampled from a natural stand within the Forest Research Unit of INRA-Pierroton (Cestas, France). Each tree was 50 m from its nearest neighbour. Typically, natural stands from this region display high degrees of heterozygosity among individuals (Mariette *et al.*, 2001; Derory *et al.*, 2002; Ribeiro *et al.*, 2002), and parentage analysis (González-Martínez *et al.*, 2002) has shown that close mature trees within stands are unrelated most of the time.

The outer bark was removed from the trunk (Supplementary material, Fig. S1A), and the inner bark was cut and removed using a knife (Fig. S1B). This sampled layer comprised the secondary phloem cells with attached cambium and immature secondary xylem cells in the early stages of differentiation (division and beginning of cell expansion). The exposed tissue on the trunk side was then scraped down to the hard lignified layer beneath (Fig. S1C). Scraped tissues were soft and comprised loosely bound differentiating xylem cells undergoing cell expansion or initiating the cell wall thickening. Tissues were collected from around the trunk and then bulked. It took 30 min to collect all the samples from each tree. Differentiating xylem samples were immediately frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$  before use.

### Phenotypic characterisation of wood-forming tissues

**Microscopic observations** Blocks of  $0.5\text{ cm} \times 0.5\text{ cm} \times 0.5\text{ cm}$  (from the 2006 samples), consisting of developing phloem, xylem and fully developed wood, were dehydrated in ethanol and embedded in paraffin. Samples were sectioned with a microdissector at a thickness of  $20\text{ }\mu\text{m}$  and examined using bright field illumination on a Nikon Eclipse 80i microscope (Nikon Instruments Europe BV, Badhoevedorp, the Netherlands) equipped with a digital camera. NIS-Elements D 2.30 (Nikon) was used for image analysis.

**Automatic fibre analysis** Cells of the differentiating wood samples collected in 2003 were dispersed by a nondestructive method. A known mass of tissue was put in a 1 : 1 (v/v) solution of acetic acid and hydrogen peroxide for 14–16 h, at  $75^{\circ}\text{C}$ . A known mass of sample was then placed in the MorFi<sup>®</sup> analyser (Techpap, Grenoble, France), diluted to make a cell (fibre) suspension, and then homogenized. The measurement chamber consists of a transparent vein with a certain geometry that allows the flux of cell suspension and images to be captured by an optical device. Images were analysed using MorFi<sup>®</sup> V7.07.16.B software.

**Chemical characterization** Extractive-free samples were characterized by Fourier transform infrared (FTIR) spectroscopy and analytical pyrolysis.

The samples were freeze-dried for 1 wk and ground into a fine powder using a pestle and mortar. An aliquot of each sample was then sequentially extracted by dichloromethane, methanol and water using a fast extraction procedure. This was conducted for 30 min in an extraction thimble in direct contact with the boiling solvent (extraction). Finally, the thimble was raised and kept above the boiling solvent for 1 h (washing step).

**FTIR spectroscopy** Approximately 1.5 mg of finely ground tissue was mixed with 200 mg of KBr spectroscopic grade using a ball mill for 20 s. A standard pellet device was used to prepare 13-mm-diameter pellets. The spectra were recorded using a Bio-Rad FTS 165 infrared spectrometer with 64 scans per sample at a resolution of  $4\text{ cm}^{-1}$ . At least two spectra per sample were obtained. Each spectrum was rationed against a background of pure KBr.

The spectra were offset at the minimum close to  $2000\text{ cm}^{-1}$ , and two different normalization procedures were used: the min-max normalization and the vector normalization. The min-max normalization method scales spectrum intensities to the effect that the minimum absorbance unit is 0 (at  $1925\text{ cm}^{-1}$ ) and the maximum is 2 (at the polysaccharide band with maxima at  $1059\text{ cm}^{-1}$ ). This normalization was performed between  $2000\text{ cm}^{-1}$  and  $400\text{ cm}^{-1}$ .

The vector normalization method calculates the average  $y$ -value of the spectrum. The average value is then subtracted

from the spectrum, decreasing the mid-spectrum to  $y = 0$ . The sum of squares of all  $y$ -values is calculated and the spectrum is divided by the square root of this sum.

For a semiquantitative evaluation of the relative proportions of polysaccharides/proteins between samples, these two normalization procedures were used. The results were essentially the same, thus only results obtained with min-max normalization are shown.

**Analytical pyrolysis** Analytical pyrolysis (pyrolysis gas chromatography with flame ionization detector, Py-GC/FID) was performed with a CDS Pyroprobe 1000 with a coil filament connected to a HP 5890 series II by a heated interface (270°C) (Rodrigues *et al.*, 2001). Each sample (75–80 µg) was pyrolysed at 650°C for 10 s with a temperature rise time of approx. 20°C ms<sup>-1</sup>. The DB1701 (60 m × 0.25 mm, 0.25 µm film, J&W Scientific) capillary column was used. Gas chromatography conditions were as follows: injector, 250°C; detector, 280°C; temperature programme, 45°C, 4 min isothermal, then heating rate 4°C min<sup>-1</sup> to 270°C. The identification of pyrolysis products was performed using previously well characterized samples, by Py-GC/MS (CDS Pyroprobe 1000 connected to a HP 6890 with a HP 5973 mass selective detector). Products were identified by their mass spectra and retention time through comparison with the NIST library and with the literature (Faix *et al.*, 1990a,b; 1991a,b; Ralph & Hatfield, 1991). At least two runs per sample were performed. Quantification of pyrolysis products was calculated with Chemstation Software (Agilent Technologies, Palo Alto, CA, USA).

### Transcriptome analysis

**RNA extraction** Total RNA was isolated from 3 g of xylem following Chang *et al.* (1993). Genomic DNA was removed by treating RNA with DNase (0.1 U µl<sup>-1</sup> of total RNA; Promega, Madison, WI, USA) for 30 min at 37°C. RNase (0.4 U µl<sup>-1</sup> of total RNA; Promega) was added to protect RNA from RNase degradation. Treated RNA was then purified using an RNeasy<sup>®</sup> Mini Kit (Qiagen GmbH, Hilden, Germany), quantified by spectrometry and quality checked using 2% agarose gels.

**cDNA library construction, sequencing and bioinformatic analysis** A composite cDNA library was constructed using the xylem samples described in Table S1. Total RNA from each sample was extracted following the method of Chang *et al.* (1993). Equal amounts of total RNA from each sample were mixed and poly A(+) RNA was isolated from this bulked sample. A cDNA library was prepared using the λ-ZAP-cDNA synthesis kit (Stratagene, La Jolla, CA, USA). Approximately 10 000 clones were excised to generate plasmid clones. The colonies were randomly picked and clones were arranged individually in 96-well microtitre plates for storage and processing. Individual plasmid clones were grown overnight

at 37°C in wells containing 250 µl of LB-ampicillin and stored at -80°C with glycerol. The 10 000 plasmid clones were sequenced using the Templifi kit (Amersham Bioscience, Piscataway, NJ, USA), by single pass from the 5'-end to generate the EST collection.

A bioinformatic pipeline that was developed to analyse the maritime pine ESTs sequences was based on our own developments and third party software. This system, available at <http://cbi.labri.fr/outils/SAM2/>, allowed the following operations to be performed automatically on all ESTs: (i) base calling: phred (sequence and quality files); (ii) vector masking: cross\_match; (iii) vector removing (and splitting); and (iv) detection of polyA (and splitting). Only sequences longer than 60 nucleotides were kept for further analysis. Since the data from the sequencing machine vary in quality for a number of reasons (Ewing *et al.*, 1999), a visual inspection was conducted for sequences with quality score < 20, that is with a mean phred base-call accuracy < 99%, and for those containing a poly A, T, C or G. Secondly, to provide a nonredundant set of ESTs to be spotted onto nylon membranes, we used the StackPack (<http://www.sanbi.ac.za>) software for clustering and assembly of the clones. A functional annotation (gene function and metabolic classification) was then assigned for each consensus. This process is based on a search for homology with public protein and nucleic acid sequence databases, using the BLAST software (Altschul *et al.*, 1997). Homologues were sequentially searched in SWISSPROT (BLASTX), TrEMBL (BLASTX), EMBL (BLASTN) and, lastly, dbEST (BLASTN). At each step, the process was stopped if a gene with a similar sequence was found (defined by an expected value lower than 10<sup>-5</sup> for BLASTX, and 10<sup>-10</sup> for BLASTN searches). Annotation decision was facilitated by complementary analysis such as sequence comparisons with dedicated databases such as PRODOM (Corpet *et al.*, 1998), PROSITE (Falquet *et al.*, 2002) and KEGG (Kanehisa & Goto, 2000). ESTs were deposited at the EMBL database (<http://www.ebi.ac.uk/embl/index.html>) and annotated sequences made available in a dedicated server (<http://cbi.labri.fr/outils/SAM/COMPLETE/index.php>).

### Production of maritime pine xylem macroarrays, hybridization, signal quantification and data analysis

The 3512 genes of the *Pinus pinaster* Ait. xylem cDNA library (Paiva, 2006), representing 89.5% of unique sequences (based on Pinus Genes Index Tentative Contigs database (PGI), <http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=pine>) and controls, including luciferase, desmin, nebulin, RAS polylinker of insertional vector of the cDNA library and water, were amplified by nested PCR as follows. The first PCR (PCR1) was prepared in 10 µl, with 1 µl of bacteria suspension added to a PCR mix containing 5 mM dNTPs, 2 mM MgCl<sub>2</sub>, 0.2 µM of universal primers M13 (-20) and R13, and 0.2 U µl<sup>-1</sup> of Taq polymerase. The second PCR (PCR2) was prepared

in 100 µl, with 2 µl of PCR1 amplification product added to a PCR mix containing 5 mM dNTPs, 2 mM MgCl<sub>2</sub>, 0.2 µM of T7 and T3 primers, and 0.2 U µl<sup>-1</sup> of Taq polymerase. All PCR reactions were performed using a Primus Thermoblock (MWG, Reinach, Switzerland). After an initial denaturation step at 95°C for 5 min, 15 cycles for the PCR1 and 40 cycles for the PCR2 were carried out as follows: 30 s denaturation at 95°C, 45 s annealing at 55°C, and 90 s elongation at 72°C. A final elongation step was conducted for 10 min at 72°C. All PCR2 were checked for their quality and quantified on an agarose gel before spotting.

Amplified genes and controls were robotically spotted (Eurogentec, Liege, Belgium) in duplicates onto a set of two nylon membranes, using a two-level template, as indicated in Fig. S2.

Antisense mRNA (aRNA) was obtained from 2 µg of total RNA using the MessageAmp kit (Ambion, Inc, Austin, TX, USA). Antisense mRNA was quantified using the RiboGreen<sup>®</sup> RNA quantitation assay (Molecular Probes Inc., Eugene, OR, USA). All aRNA samples were diluted in RNase-free water to a final concentration of 50 ng µl<sup>-1</sup>. Desmin was used as an external control of probe labelling. A desmin cDNA fragment cloned in pBS-SK<sup>+</sup> was used for *in vitro* transcription of RNA from T3 promoter, according to instructions provided by the supplier of T3-polymerase (Riboprobe combination systems T3/T7 RNA Polymerase; Promega). Desmin aRNA was also quantified by the RiboGreen<sup>®</sup> RNA assay.

Probes were synthesized using 300 ng aRNA supplemented with 2% desmin aRNA using the Strip-EZ<sup>™</sup> RT kit from Ambion (Ambion Inc.) in the presence of random-sequence decamer primers provided with the kit, according to the manufacturer's instructions.

Filters were prehybridized for 8 h in 15 ml of hybridization buffer (5xSSC, 5X Denhardt, 0.5% SDS, 10 µg ml<sup>-1</sup> denatured salmon sperm DNA) at 65°C. Hybridizations were carried out in high stringency conditions at 65°C overnight using 15 ml of fresh hybridization buffer supplemented with a minimum of 10<sup>6</sup> cpm purified and denatured probe per ml of buffer. Membranes were washed twice (5 min each) at room temperature with 2xSSC 0.5% SDS buffer. They were then washed twice at 65°C for 15 min in 2xSSC 0.1% SDS buffer, two washes of 15 min in 1xSSC 0.1% SDS followed by two washes (15 min each) in 0.1xSSC 0.1% SDS buffer. Finally, they were wrapped in a plastic film, exposed to the General Purpose PhosphorImager screen (Amersham Bioscience) for a period of 4 d, and finally scanned using Storm System (Amersham Bioscience), set to a resolution of 50 µm.

Hybridization signals were quantified using ArrayVision (Imaging Research, Ontario, Canada). Median-based trimmed mean density (MTM density) was used for spot quantification. The background of each spot within each level 1 grid was calculated using the median of the four blanks located in the corresponding level 1 grid ('selected spot background' option of ArrayVision; see Fig. S2). The background values were

removed from individual spot density values, and the adjusted values were used for statistical analysis.

Two batches of three hybridization replicates (including labelling, overnight hybridizations, washing and screen exposition) were performed, resulting in a total of six data points for each spotted gene (three replicated hybridizations × two spots per gene). Batch A corresponded to samples L1, L3 and L5, and batch B corresponded to samples L2, L4 and L6. Overall, 18 hybridizations (three replicates × six samples) were performed.

After background correction, the average density for each membrane set and hybridization was calculated. The density of all data points from a membrane was then divided by the corresponding average mean to account for technical effects. The following ANOVA model was then applied to each gene (based on corrected density values). Statistical analyses were performed using SAS v. 6.12 (Statistical Analysis System version 6.12, SAS Institute, Cary, NC, USA). Analysis of variance (ANOVA) was carried out using the following model:

$$Y_{jkl} = \mu + B_j + L_{k/l} + \varepsilon_{jkl}, \quad \text{Eqn 1}$$

where  $Y_{jkl}$  is the gene corrected density of batch  $j$  in sample  $k$  for replicate  $l$  ( $l = 1-6$ ),  $B_j$  is the batch effect ( $j = A$  and  $B$ ), and  $L_{k/l}$  is the cambial age effect ( $k = 1, 2, 3, 4, 5, 6$ ) nested in batch effect  $B_j$ , and  $\varepsilon_{jkl}$  is the residual of the model.

The criteria used to select differentially expressed genes were based on the results from these ANOVA models. A transcript was classified as differentially expressed between the different samples if its  $P$ -value for the cambial age effect within batch was less than 10<sup>-4</sup>, if the cambial age effect also explained more than 50% of the total variation in the sum of squares (in order to guarantee that most variation was the result of cambial age), and if it displayed a normal residual distribution. Using such stringent criteria, it was our hope to decrease the rate of false positives, as the threshold used was an approximation of the Bonferroni threshold used for multiple comparison tests, which guarantees that the identification of false positives is less than one. By considering the normality of residues, we guaranteed one of the conditions of application of the parametric tests implemented in the model. However, the risk associated with these stringent statistical criteria was that true positives could be discarded, that is, genes showing a true differential expression between the cambial age samples.

Centred-reduced data of the differentially expressed genes, were analysed using the 'Click' algorithm (Sharan & Shamir, 2000) of the Expander software (Shamir *et al.*, 2005), in order to cluster genes by their expression profiles.

**Quantitative real-time PCR (qPCR) assay** PCR primers were designed using PRIMER3 software (Rozen & Skaletsky, 2000) for 14 genes, as well as for two controls corresponding to a ribosomal protein (BX252550) and a Profilin 1 gene (BX249454). The two controls were selected based on a

selection of false positives was minimized. Centred-reduced data for the differentially expressed proteins were analysed using the k-means algorithm of the Expander software.

**Protein identification by mass spectrometry** CBB-stained protein spots were washed in H<sub>2</sub>O/ACN (50 : 50) until destaining. The solvent mixture was removed and replaced by ACN. After shrinking of the gel pieces and ACN removal, gel pieces were dried in a vacuum centrifuge. Gel pieces were rehydrated in 10 ng µl<sup>-1</sup> trypsin (Sigma-Aldrich, St Louis, MO, USA) in 50 mM NH<sub>4</sub>HCO<sub>3</sub> and incubated overnight at 37°C. The supernatant was removed and stored at -20°C, and the gel pieces were incubated for 15 min in 50 mM NH<sub>4</sub>HCO<sub>3</sub> at room temperature under rotary shaking. This second supernatant was pooled with the previous one, and a H<sub>2</sub>O/ACN/HCOOH (47.5 : 47.5 : 5) solution was added to the gel pieces for 15 min. This step was repeated twice. Supernatants were pooled and concentrated in a vacuum centrifuge to reach a final volume of 25 µl. Digests were finally acidified by addition of 1.2 µl acetic acid and stored at -20°C.

Peptide mixtures were analysed by online nano HPLC (LC Packings, Amsterdam, the Netherlands) coupled to a nanospray LCQ Deca XP Plus ion trap mass spectrometer (Thermo-Finnigan, San Jose, CA, USA). Peptides were separated on a 75 µm id × 15 cm C18 PepMap™ column (LC Packings, Amsterdam, the Netherlands).

The flow rate was set at 200 nl min<sup>-1</sup>. Peptides were eluted using a 5–50% linear gradient of solvent B in 30 min (solvent A was 0.1% formic acid in 5% ACN, and solvent B was 0.1% formic acid in 80% ACN). The mass spectrometer was operated in positive ion mode at a 1.8 kV needle voltage and a 43 V capillary voltage. Data acquisition was performed in a data-dependent mode alternating in a single run, a MS survey scan over the range *m/z* 300–1700 and three full-scan MS/MS in an exclusion dynamic mode. MS/MS data were acquired using a two-*m/z*-units ion isolation window, a 35% relative collision energy, and a 0.5 min dynamic exclusion duration.

Peptides were identified with SEQUEST through the Bioworks 3.2 interface (Thermo-Finnigan, Torrance, CA, USA) using the 327,484 indexed entries of the Pinus Gene Index (<http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=pine>). The validation filters used for the database query were as follows: Peptide DeltaCN ≥ 0.1, Peptide Xcorr vs Charge State ≥ 1.5 (+1), 2.0 (+2), 2.5 (+3), 3.0 (≥ +4), Peptide Probability ≤ 0.05, Peptide # Matches = 1, Protein Number of Different Peptides ≥ 2.

## Results

### Phenotypic analysis of base to crown wood-forming tissues

**Anatomical and morphological characterization** The transition between J/CW to M/BW wood in maritime pine, estimated

from variations in wood density, occurs at around the 10th to the 12th growth ring (Zobel *et al.*, 1972; Radi, 1992; Dumail *et al.*, 1998; Fonseca & Lousada, 2000). Considering the age range of the 2003 samples, it is likely that the differentiating xylem collected along the stem included a range of wood, from BW to CW. To ensure that the extreme samples collected at the base and the crown had the properties commonly reported for these types of wood, differentiating xylem cell morphology was characterized using the MorFi<sup>®</sup> automatic analyser. Significant differences were obtained between the samples collected at the crown and the base, with regard to the mean length (410 vs 1020 µm, respectively) and the mean width of cells (27.9 vs 31.2 µm, respectively). These data fit the expectation of shorter and thinner fibres in CW as compared with BW.

Upon microscopic observations (carried out for the 2006 samples; see Fig. S3) it can be seen that: (i) the number of tracheids produced by the cambium at the crown is about twice that produced at the base; (ii) the length of each zone corresponding to the different stages of secondary xylogenesis also differs between both samples; and (iii) tissues left out from the sampling (at either the base or the crown) comprise secondary phloem with attached cambium as well as xylem cells in the early stages of differentiation. The exposed tissue that was collected comprised loosely bound differentiating xylem cells undertaking cell expansion (EZ zone in Fig. S3) and maturing tracheids (cells undertaking cell wall thickening and programmed cell death, MZ zone in Fig. S3). Although the length of the expansion zone was found to be not significantly different between crown and base samples, the mean number of cells differed significantly (8.1 vs 5.8, respectively; *t*-test *P* < 0.05). The length of the maturing zone was clearly different (*t*-test *P* < 0.05) between both samples (2316 µm in CW vs 1344 µm in BW), as well as the number of cells (60.8 cells in CW vs 29.3 cells in BW, *t*-test *P* < 0.05), and the mean cell radius (11 µm in CW vs 15 µm in BW, *t*-test *P* < 0.05).

**Cell wall chemical composition** Cell wall chemical composition of differentiating xylem samples collected in 2003 was assessed by FTIR and analytical pyrolysis on extractive-free tissues.

Fourier transform infrared spectroscopy spectra were obtained for samples L1 to L6 collected along the tree stem. Besides the usual features found in the spectra of fully developed wood, that is, the characteristic bands of lignin and polysaccharides, we also found bands of amide I, II, and III (Fig. S4) that can be attributed to the presence of amide vibrations from the peptide group (proteins). A higher protein content was found in the crown (samples L5 and L6), whereas a lower protein content was found in the base samples (L1 and L2). Intermediate amounts were observed for samples L3 and L4 (Fig. 1).

Pyrograms (Fig. S5) contained typical pyrolysis products of polysaccharides and lignin origin (see also details in Table S3).

The variation in the main groups of pyrolysis products along the stem is presented in Fig. 2:

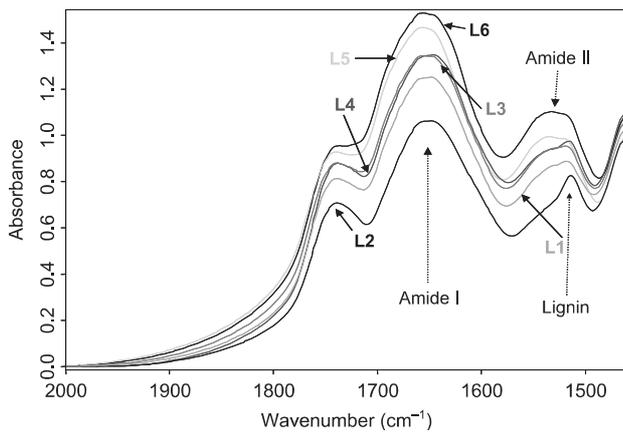


Fig. 1 Min-max normalized spectra of differentiating xylem from L1 (base wood) to L6 (crown wood) showing the broad amide I and amide II bands.

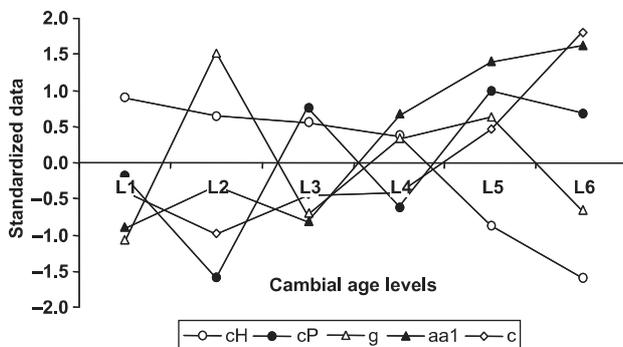


Fig. 2 Variation of main classes of pyrolysis products along the stem, from the base (L1) to the crown (L6) of the tree sampled in 2003: aa1, toluene (amino acid content); cP, pentosans (hemicellulose origin); cH, hexoxanes (mainly from cellulose); c, other carbohydrates (mainly from hemicellulose origin); G, guaiacyl lignin (G) units. Standardized data, data for each pyrolysis product was normalized to a mean of 0 and standard deviation of 1.

- Toluene aa1, a characteristic pyrolysis product of phenylalanine (Moldoveanu, 1998), was found to be at a low and stable level from L1 to L3 (BW), and then rapidly increased from L3 to L6 (CW).
- Lignin: part of H-units result from protein contribution, as it is known that tyrosine produces phenol and cresol, two main H pyrolysis products of lignin (Faix *et al.*, 1991a; Moldoveanu, 1998). Indeed, H-units followed the same pattern as phenylalanine (aa1). For this reason we preferred not to use the H-units as a reliable measure of lignin composition, but rely more on lignin determined by G-units (g). G-units lignin showed a very irregular profile along the cambial age gradient. After an abrupt increase from L1 to L2, it dropped down from L2 to L3, then increased again from L4 to L5, and finally decreased from L5 to L6.
- Carbohydrate pyrolysis products other than pentosans and hexoxanes (c) increased from BW to CW with a rapid increase

from L4 to L6. It should be noted that hemicelluloses are the main contributors to c. pentosans (cP), and that the content was quite irregular but showed a tendency to increase from the bottom (L1) to the top (L6) of the tree. Together, c and cP indicate that hemicellulose content increased from BW to CW. The higher content of cellulose (represented by hexoxanes cH and, in particular, cH7-levoglucosan) was found to slightly decrease from L1 to L4, and then decreased abruptly from L4 to L6. It should be noted that the pattern of c (mainly of hemicellulose origin) and cH (mainly of cellulose origin) presented symmetric profiles.

Expander software was used to cluster the differentiating xylem samples according to their metabolic profile. This analysis revealed two main sub-trees (Fig. 3):

- Group A included the differentiating xylem samples derived from the youngest cambium collected at the top of the tree (L5 and L6). These samples had lower cellulose (cH, cH7), higher hemicelluloses (c and cP), and higher amino acid (aa1) content.
- Group B consisted of the differentiating xylem samples derived from the oldest cambium collected at the base of the tree (L1–L4). They were characterized by higher cellulose content. In group B we observed that the metabolic profile of samples L2 and L4 was different from that of L1 and L3, L2 and L4 containing more G-lignin.

#### Environmental and genotypic effects on cell wall chemical composition

To study the environmental and genotypic effects on cell wall chemical composition, pyrolysis was carried out on wood-forming tissue collected at the base and the crown of three unrelated individuals located in the same stands and sampled mid-May (therefore forming early wood). The first set of samples were collected from the 30-yr-old tree sampled in 2003, and the second set were collected from the two 33-yr-old trees sampled in 2006. A hierarchical clustering of the data (Fig. 4) first showed that these six samples clustered into two distinct sub-trees corresponding to each sampling year (2003 vs 2006), suggesting that climatic conditions had the strongest impact on cell wall chemical composition. Despite the sampling year, it was also found that the contrast between CW and BW samples was similar and was half as much as the distance between years. Finally, the distance between the two 2006 genotypes was much lower and comparable to that found between two consecutive cambial ages (L5 vs L6 or L0 vs L1).

#### Molecular analysis of base and crown wood-forming tissues

**Variability of transcript accumulation along the stem** The same six samples from 2003 that were previously characterized at the chemical level were also analysed at the transcriptome level, using high-density filters (HDFs). The HDFs consisted of 3512 unique clones derived from a nonnormalized xylem cDNA library (Paiva, 2006). A total of 262 (7.5%) transcripts

pre-screen of 10 genes on different types of wood-forming tissues collected along the trunk (from BW to CW) and through the duration of the growing season (from early to late wood). Primers were designed to have an optimal size of 22 bp (18–24 bp), GC content of 40–60%, and TM of 58–62°C. Other criteria, such as primer self-annealing, were also taken into account. Predicted fragment size ranged between 103 and 228 bp. Oligonucleotides were synthesized by Eurogentec. Primers pairs are listed in Table S2.

One microgram of total RNA was reverse transcribed, using the ImProm-II™ Reverse Transcription System (Promega), in accordance with manufacture instructions. Two independent reverse-transcription reactions (RT) were performed using a pool of total RNA and each one was diluted 10-fold before each qPCR reaction. The qPCR reactions were performed on the Chromo4™ Multicolor Real-Time PCR Detection System (Bio-Rad Laboratories, Inc. Hercules, CA, USA), by adding 10 µl of IQ™ SYBR® Green SuperMix (Bio-Rad), 3 µl of diluted cDNA, 3 pmol of each primer, and water (to the final volume of 20 µl), to the reaction mix. After one initial incubation at 95°C for 3 min, amplifications were performed for 40 cycles with the following cycle profile: a denaturing step at 95°C for 15 s followed by an annealing/extension step at 60°C for 45 s, as recommended by the manufacturers. A fluorescent signal was obtained at each cycle at the 60°C annealing/extension step. After amplification, a melting curve was obtained allowing the detection of primer-dimers, and other nonspecific amplification products.

For relative quantification of transcript accumulation, Ct values were obtained for each of the two gene-primer-cDNA dilutions, and data were analysed using the Excel (Microsoft) macro GENEX v1.10 (Gene expression Analysis for iCycle iQ® Real-time PCR Detection System, v1.10, 2004, Bio-Rad Laboratories), using the methods derived from the algorithms of Vandesompele *et al.* (2002). For each tested gene, average PCR efficiency was calculated for each individual sample well (automatically calculated by Chromo4™ Multicolor Real-Time PCR Detection System (Bio-Rad Laboratories) software and based on fluorescence signal) and used in expression data analysis.

#### Proteome analysis using two-dimensional gel-electrophoresis (2D-E), coupled with protein identification by tandem mass spectrometry (MS/MS)

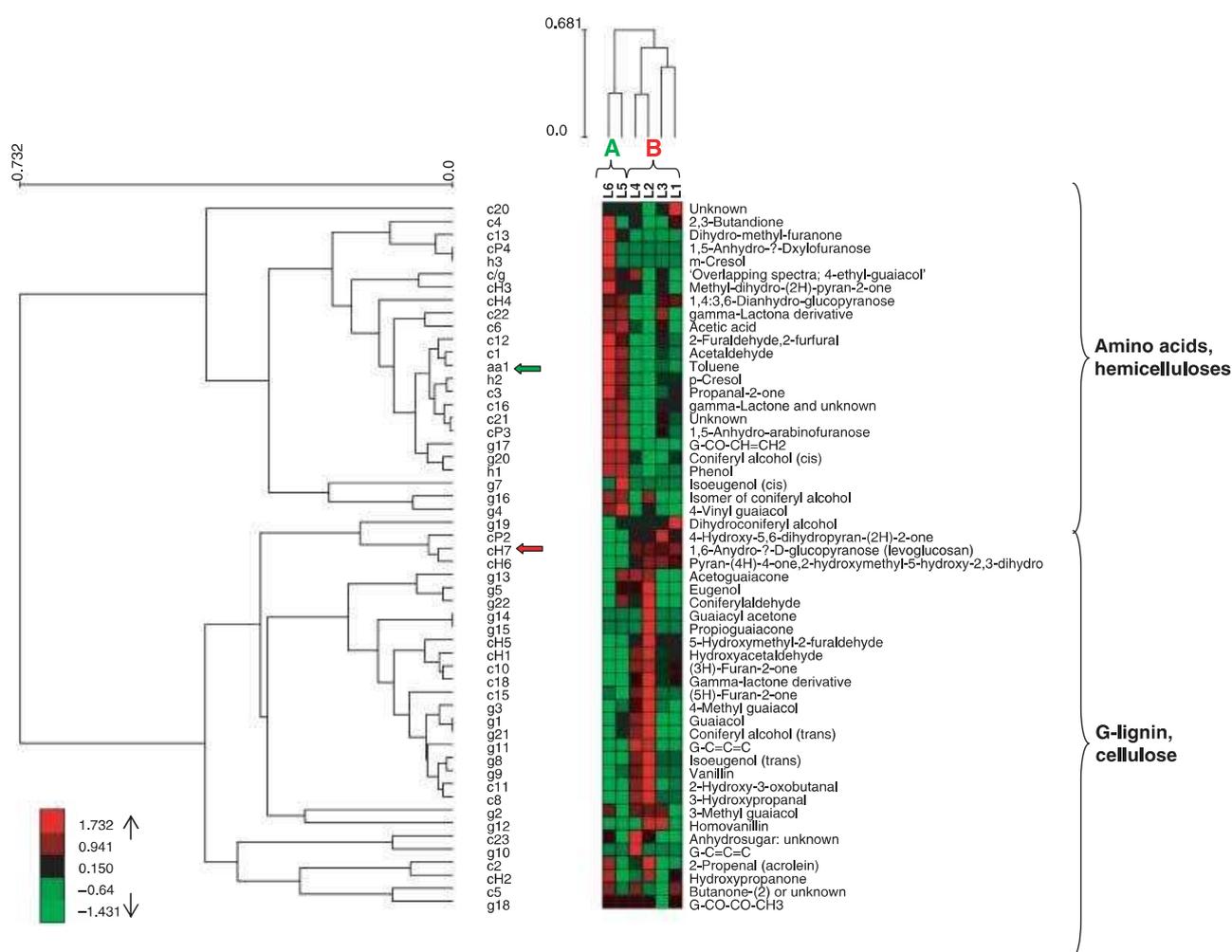
**Protein extraction and quantification** Starting with 2 g of fresh tissue, the total protein of each of the four samples described earlier (2003 tree) was extracted following the procedure described by Gion *et al.* (2005). Proteins were stored at –80°C. Three extractions were completed for each sample and pooled for protein quantification. The resulting mix was quantified over six replicated assays, using the protocol described by Ramagli & Rodriguez (1985). The mean concentration was then calculated and used to load 300 µg of proteins on each IPG strip.

**Two-dimensional gel-electrophoresis** This technique (O'Farrel, 1975) was used to analyse total protein from the xylem samples following the procedure of Gion *et al.* (2005) adapted for the IPGphor system (Amersham Biosciences, Uppsala, Sweden). For the IEF, 24 cm strips were used with a linear pH gradient ranging from 4 to 7. To ensure gel reproducibility, four replicates were performed for each sample, resulting in a total of 16 gels.

Coomassie brilliant blue G250 (CBB G-250; Bio-Rad) was used for gel staining. Gels were fixed for 2 h in a solution containing 2% phosphoric acid and 50% ethanol. After three washes with water (each of 30 min), the gels were placed in an incubation solution (methanol 34%, ammonium sulphate 17%, phosphoric acid 2%) for 1 h, and then immersed in a staining solution (methanol 34%, ammonium sulphate 17%, phosphoric acid 2%, Coomassie blue 0.05%) for 5 d. Finally, the gels were stored for several days in a 5% acid acetic solution before scanning and manual spot picking.

**Image acquisition and spot detection** Stained gels were digitalized using the Image Scanner (Amersham Biosciences) and the LabScan software (Amersham Biosciences). First, a calibration with a grey scale was necessary to transform grey levels into OD values for each pixel of the gel picture. The calibration method used was the colloidal blue method described in the LabScan manual. All the gel pictures were saved as tiff files. Image analysis was performed using the Image Master 2D-Elite software (IM2D; Amersham Biosciences). The 16 gel images were placed in one folder. The wizard detection method proposed by the software was used to detect the spots. The spots that were automatically detected by the software were checked manually. Following the detection procedure, the volume of each spot corresponded to a gross value. In order to eliminate the background from this gross value, the nonspot mode of IM2D was used. Finally, all the gels were matched to attribute a common spot identity for the same spots derived from different images. For this, we used the automatic matching options of IM2D. After a visual check of the matching, the IM2D software was used to construct a master gel. For each sample, when a protein was detected in all of the four replicates, this protein was automatically added to the master gel. Normalized volumes were finally obtained using the total spot volume normalization procedure of IM2D.

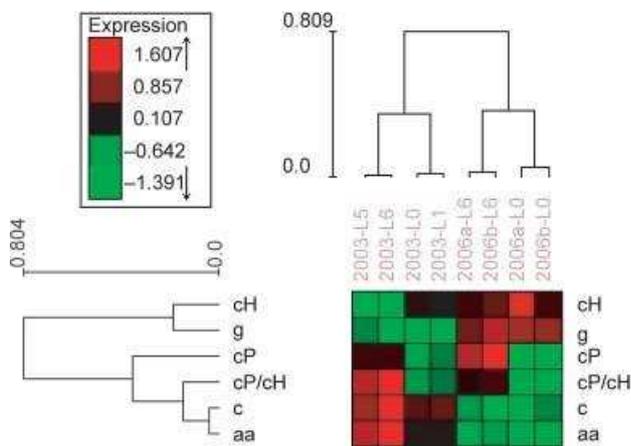
**Statistical analyses** Analysis of variance (ANOVA) was used to test the position effect (L0, L2, L4, L6) on protein accumulation, using the following model:  $Y_{ijk} = \mu + L_j + e_{ijk}$ , where  $Y_{ijk}$  is the normalized intensity value of spot  $i$  ( $i = 1-1372$ ) in level  $j$  ( $j = 0, 2, 4, 6$ ) for replicate  $k$  ( $k = 1-4$ ),  $\mu$  is the mean intensity of spot  $i$  in all gels,  $L_j$  is the effect of level  $j$ , and  $e_{ijk}$  is the residual term. ANOVAs were performed using R (R Development Core Team, 2004). A protein was classified as differentially expressed between the different samples if its  $P$ -value for cambial age effect was less than  $10^{-3}$ , so that the



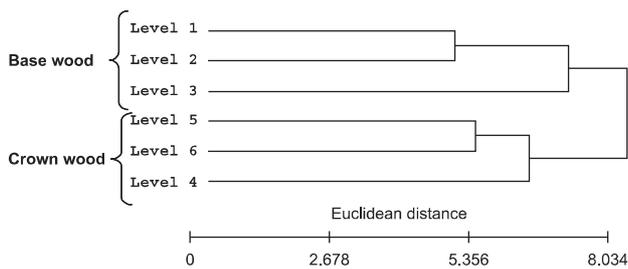
**Fig. 3** Clustering of differentiating xylem samples (levels L1 to L6) according to their metabolic profiles and clustering of pyrolysis products. ch7, marker for cellulose content; aa1, marker for amino acid content (toluene). The scale bar adjacent to each dendrogram represents the distance measurement used in the Expander software algorithm ( $(1 - \text{Pearson correlation})/2$ ). The colour scale bar represents the standardized content of pyrolysis products. For each pyrolysis product, data were standardized to give a mean of 0 and standard deviation of 1.

were found to be differentially expressed along the stem. About 55.7% of the differentially expressed transcripts could be attributed to known functional proteins (BLASTX  $< 10^{-5}$  against the SwissProt database), 17.2% matched to proteins of unknown function (BLASTX e-value  $< 10^{-5}$  for sequence homology with *Arabidopsis* unknown proteins), and 27.1% did not match any sequences in public databases and were therefore classified as putative proteins (BLASTX e-value  $> 10^{-5}$  or 'no hit'). Such a high proportion can in part be attributed to a lower sequence read length for the 'putative protein' category (436 bp on average vs 572 bp for known function proteins), but for some of the ESTs (e.g. 'unknown function' category characterized with a mean length of 576 bp), this is also likely to be the result of a general lack of information regarding the putative molecular mechanisms in which these genes are involved.

For differentially expressed genes, the fold-change ratio between maximum and minimum values varied between 1.3 and 3.9 (data available in Table S4). About 45% of these genes had a fold-change ratio below 1.6; 35.9% had a ratio varying between 1.6 and 2, whereas 19.1% of the genes had a ratio greater than 2. Among the 19 genes showing the largest changes in gene expression level (fold-change ratio higher than 2.5), we found a geranylgeranyl pyrophosphate synthetase (plastidial) (BX254083, 3.4-fold) and an  $\alpha$ -tubulin (BX249184, 2.8-fold) both up-regulated in BW, an orthologue to a *P. taeda* water deficit inducible protein LP3/abscisic stress ripening protein (BX253505, 2.7-fold), and a GASA5-like protein (BX250252, 2.6-fold) both up-regulated in CW. Nine genes classified as putative proteins or those of unknown function were found among differentially expressed genes displaying the highest variation between the samples.



**Fig. 4** Clustering of differentiating xylem according to their metabolic profiles. Samples were collected at the base (B) and the crown (C) of the three adult trees sampled in 2003 and 2006 (trees a and b) at the same period (mid-May) and in the same experimental field. The scale bar adjacent to each dendrogram represents the distance measurement used in the Expander software algorithm ( $(1 - \text{Pearson correlation})/2$ ). The colour scale bars represent the relative standardized content of main classes of pyrolysis products. For each pyrolysis product class, data were standardized to give a mean of 0 and a standard deviation of 1.



**Fig. 5** Clustering of differentiating xylem samples collected from the crown (L6) to the base (L1) of an adult tree sampled in 2003, according to their transcriptomic distance (Euclidian distance of  $\log_2$  expression valued and UPGM algorithm).

Epclus (<http://www.bioinf.ebc.ee/EP/EP/EPCLUST/>) was used to cluster the differentiating xylem samples according to their transcriptomic Euclidian distance. This was computed using  $\log_2$  scaling data based on the 262 differentially expressed transcripts. Samples were found to cluster into two distinct sub-trees (Fig. 5), that is, one sub-tree corresponding to BW samples (L1, L2 and L3) and the other corresponding to CW samples (L4, L5 and L6).

In order to summarize the complexity of the expressional data, differentially expressed genes were also clustered according to their expression profiles using the Click algorithm of Expander. Of the 262 genes, 223 were clustered into four groups (Fig. 6), with an average homogeneity of 0.885, and an average separation score of  $-0.295$ , which means that the expression profiles in each cluster were very homogeneous. Most of the genes were split into two clusters: clusters 1 and 2, comprising 35.5 and

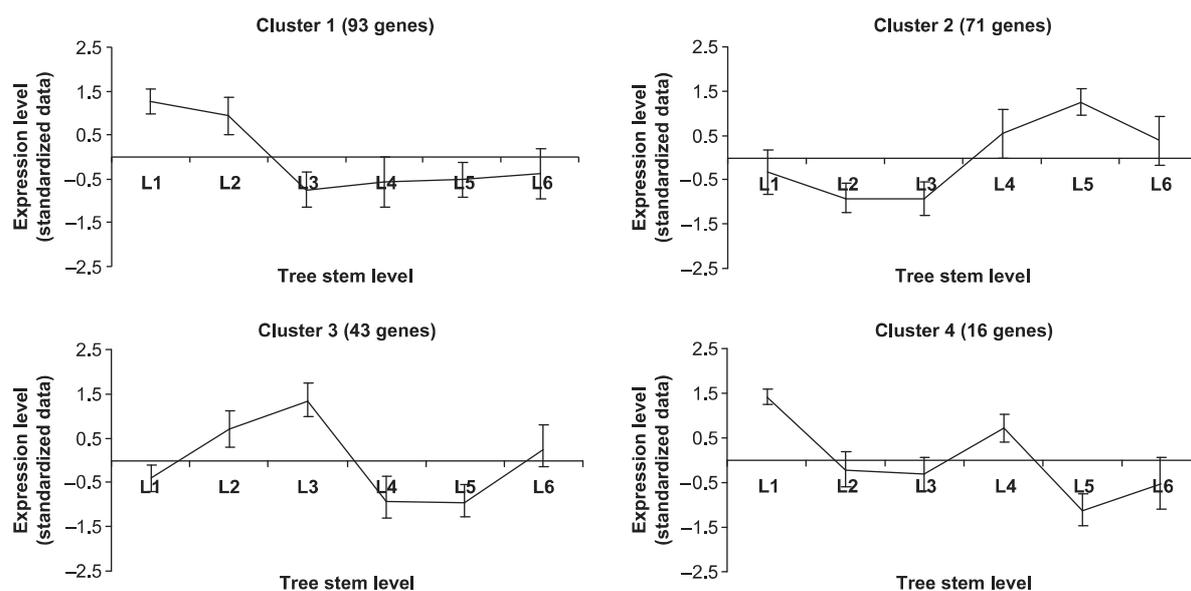
27.1% of the differentially expressed genes, respectively. Cluster 3 (16.4%) and cluster 4 (6.1%) comprised 22.5% of the differentially expressed genes, whereas 39 genes remained as singletons.

Expressed sequence tags within each cluster could be split into 14 functional categories (Table 1). EMBL accessions, PGI tentative contigs (TC) and annotations (sorted by clusters and/or functional categories) are provided as supplementary data (Table S4).

Clusters 1 and 2 were particularly interesting because they displayed very distinct profiles. Transcripts of cluster 1 were over-expressed in samples L1 and L2, and were thus referred to as BW-related genes. Their accumulations strongly decreased at L3, whilst L3 to L6 maintained a relatively constant expression level. Amongst the known function genes with a fold-change ratio greater than 2, we found a subtilisin-like protease (BX255758, twofold), a cytochrome c1 precursor (BX253984, twofold), a 40S ribosomal protein S7 (BX255511, 2.1-fold), a 17.6 kDa class I HSP (BX251102, 2.6-fold), a gamma-thionin homolog precursor (BX254131, 2.2-fold), an ATP-binding cassette transporter (BX249573, 2.6-fold) and a SEC13 related protein (BX249772, 2.6-fold). Two genes with the highest ratios were for an  $\alpha$ -tubulin (BX249184, 2.8-fold) and a geranylgeranylpyrophosphate synthetase (BX250083, 3.4-fold). Interestingly, two transcripts involved in the meristematic regulation (Homeobox protein SHOOTMERISTEMLESS (BX254965, 1.5-fold), and the receptor kinase CLAVATA 1 (BX252322, 1.5-fold)) were also included in this cluster.

Cluster 2 included transcripts which were over-expressed in crown samples (L4, L5, L6), and were thus further referred to as CW-related genes. The CW cluster contained 71 transcripts, 37% corresponding to putative or unknown function proteins. Amongst the genes with known function, 38% were classified in the 'protein synthesis' category (corresponding mainly to ribosomal proteins). The 'metabolism' (13%), 'energy' (9%) and 'cellular organization' (9%) categories were also well represented. Amongst the genes of known function (with a fold-change ratio greater than two) we found four ribosomal proteins (BX251945, BX251256, BX250069, BX255250, 2.1- to 2.3-fold), a GAS5 like protein (BX250252, 2.6-fold) and a LP3 (ASR) protein (BX253505, 2.7-fold). Three genes of unknown function, which were among the most significant (BX255469, 2.9-fold; BX252990, 2.6-fold; and BX248868, 3.0-fold) were also found in this cluster.

Cluster 3 comprised 43 genes showing a rapid increase from L1 to L3, then a rapid decrease from L3 to L4. From L4 to L5 they maintained their expression, and finally slightly increased. Almost three-quarters of these genes were classified as putative proteins or as proteins of unknown function. 'Transcription' was the most represented category, with three genes, including an Avr9/Cf-9 rapidly elicited protein (BX253198, 1.7-fold), a glycine-rich RNA-binding protein 7 (BX250887, 1.8-fold), and a zinc finger protein (BX252274, 2.3-fold).



**Fig. 6** Clustered mean expression profiles of differentially expressed genes along the trunk (tree sampled in 2003), from the bottom (level L1) to the crown (level L6). Clusters were obtained using the 'Click' function of the Expander software on standardized data (mean 0 and standard deviation 1). Error bars represent the standardized expression level variation within each level.

**Table 1** Cluster distribution of the different functional categories of differentially expressed genes along the stem

Functional category	% EST spotted onto the array	Number of EST (cluster as in Fig. 6)				
		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
Communication/signal transduction	2.7	2	3	1	–	6
Cell division and growing	2.3	3	1	–	1	5
Protein fate	3.0	5	–	2	2	9
Energy	2.6	2	4	1	–	7
Metabolism	10.2	10	6	2	6	24
Cellular organization	2.1	3	4	–	–	7
Stress response	2.8	5	1	–	3	9
Protein synthesis	4.5	8	17	2	–	27
Intracellular traffic	1.5	2	2	2	–	6
Transcription	4.0	4	1	3	–	8
Transport	2.3	5	2	–	–	7
<i>Not classified</i>						
Putative protein <sup>a</sup>	36.9	25	15	19	2	61
Unknown protein <sup>b</sup>	21.6	16	11	9	2	38
Others	3.4	3	4	2	0	9
<b>Grand total</b>	<b>100</b>	<b>93</b>	<b>71</b>	<b>43</b>	<b>16</b>	<b>223</b>

<sup>a</sup>Putative proteins (BLASTX e-value > 10<sup>-5</sup> or 'no hit')

<sup>b</sup>Unknown function proteins (BLASTX e-value < 10<sup>-5</sup> for sequence showing homology with *Arabidopsis* unknown proteins).

Cluster 4 consisted of 16 genes, presenting two peaks of expression at L1 and L4, and low to very low levels of expression at L2, L3, L5 and L6. This cluster could also be considered as characteristic of BW-related genes. 'Metabolism' (37.5% of the genes) was the main functional category of this cluster, including, for example, a cellulose synthase (BX249248, 2.8-fold) and two chitinases (BX254139 and BX254191, 2.5- and 2.1-fold, respectively).

**Variability of protein accumulation along the stem** Two-dimensional gel-electrophoresis was used to analyse the proteome variation of wood-forming tissues along the trunk of the tree sampled in 2003. A reference 2D-E map was established using the proteins extracted from four differentiating xylem samples associated with the transition from BW formation (level L0) to CW formation (level L6). A total of 1372 spots were placed on the reference map that was further used to compare the

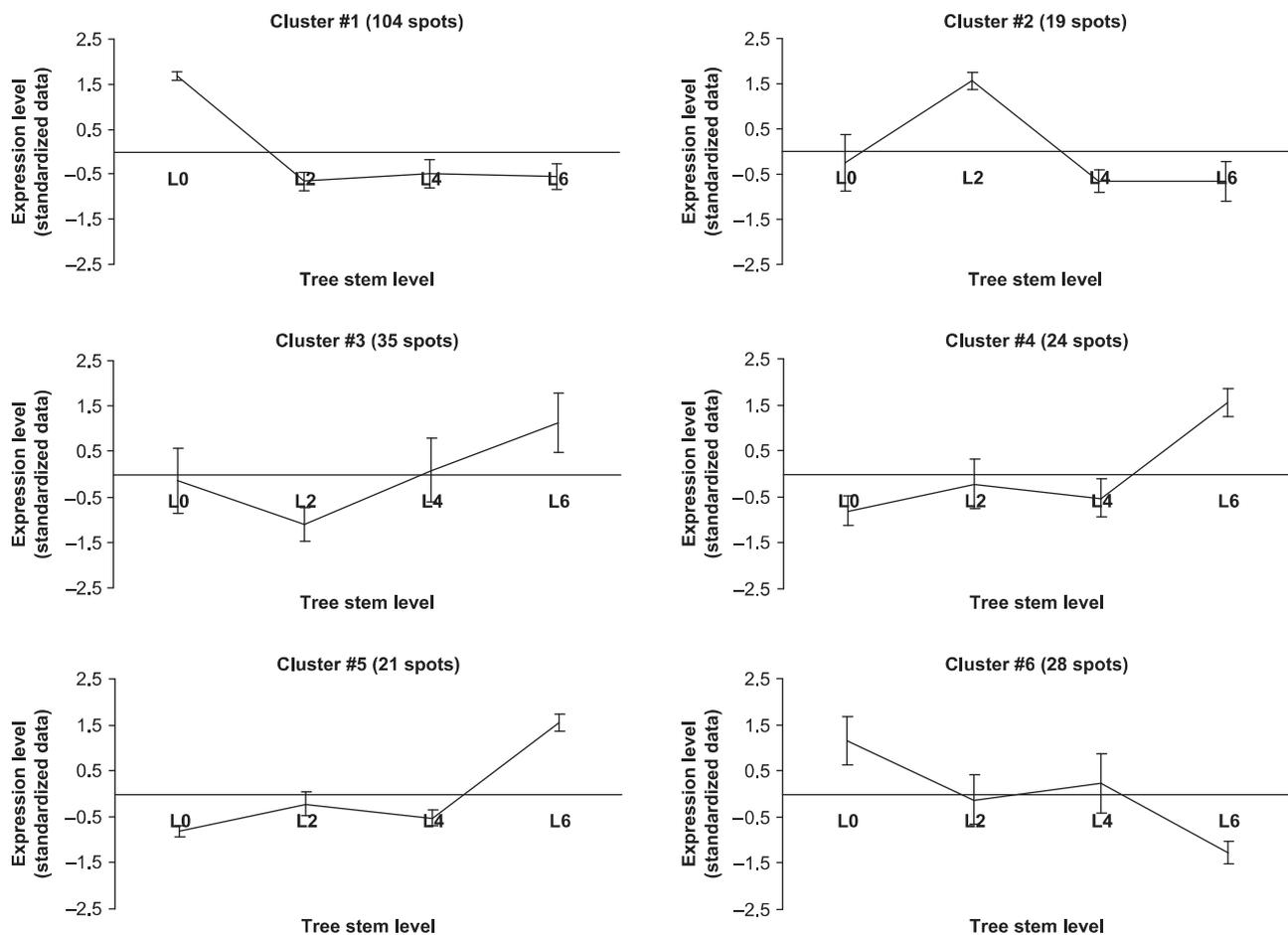


Fig. 7 Clustered mean expression profiles of differentially expressed proteins along the trunk (tree sampled in 2003, levels L0, L2, L4 and L6). Clusters were obtained using the 'K-means' function of the Expander software on standardized data (mean 0 and standard deviation 1). Error bars represent the standardized expression levels variation within each level.

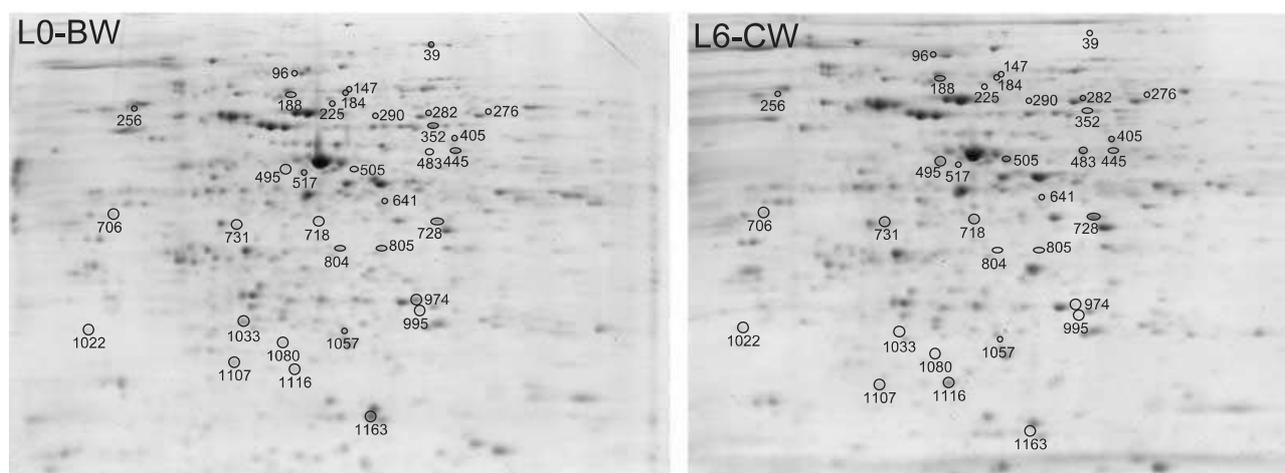
abundance of each spot/protein (spot volume) between the four levels. Statistical analysis using ANOVA allowed the detection of 231 spots, each showing significant differences between the four levels.

The 231 spots were clustered into six groups using the K-means algorithm of Expander software (Fig. 7). Clusters 1 and 6 comprised 45.0 and 12.1% of the spots showing maximum abundance at L0, that is, associated with the formation of BW. Conversely, clusters 3, 4 and 5 comprised 15.1, 10.4 and 9.1% of the spots presenting a maximum abundance at L6, that is, associated with the formation of CW. Cluster 2 presented a peak at L2.

A total of 33 spots were characterized by LC ESI MS/MS, representing spots over-expressed either in BW (22 spots) or CW (11 spots) samples. These spots (Fig. 8) were manually checked to ensure the selection of spots with high resolution and reproducibility between replicates. Table S5 summarizes the mass spectrometry results (also available at <http://www.cbib1.cbib.u-bordeaux2.fr/Protic/public/PublicMAP.php>).

For half of the spots, more than one protein was identified, which could result from the comigration of proteins with similar electrophoretic properties, thus hindering their enumeration and identification (Westbrook *et al.*, 2001).

The BW clusters comprised mainly proteins of the defence and amino acid metabolism functional categories. Defence-related proteins included peroxidases (#1033, #1107, #445; 2.6-, 5.1- and 2.4-fold variation), heat-shock proteins (#974, #1163, #995; 8.5-, 18- and 10-fold) and a chaperonin (#184, 2.7-fold). The amino acid-related proteins, in particular, included those involved in methionine metabolism, that is, methionine synthase (#1057, 4.4-fold; #39, 7.1-fold), adenosylhomocysteinases (#352, 6.4-fold; #282, 1.5-fold), adenosylmethionine synthase (#405, 5.4-fold); and glutamine synthetase (#804, 3.0-fold). In addition, structural enzymes of cellwall polysaccharides and lignin biosynthetic pathways were identified in the BW clusters. These included UDP-glucose dehydrogenase (#276, 7.5-fold), UTP-glucose-1-phosphate uridylyltransferase (#225, 5-fold) and cinnamoyl-CoA reductase (#641, 2.5-fold).



**Fig. 8** Two dimensional gel-electrophoresis (2D-E) maps of maritime pine wood-forming tissue. Proteins that were identified are marked with circles and numbered as in Table S5 (L0, base wood; L6, crown wood).

Proteins involved in vesicular traffic (GTP-binding protein, Rab family (#1060, 6-fold) and Nucleosome assembly protein 1-like 1 (NAP-1 related protein) (#256, 4.3-fold) were also observed.

The CW clusters contained proteins of the defence category, such as glyoxalase I (#731, 6.2-fold) and heat-shock proteins (#1116, 4.6-fold; #96, 2.1-fold), two proteins of the energy category (RuBisco subunit binding protein (#188, 1.6-fold) and malate dehydrogenase (#728, 2.3-fold)), two cytoskeletal proteins (actin (#495, 6.2-fold) and  $\beta$ -2-tubulin (#718, 2.6-fold)), as well as two adenosylmethionine synthases (#483, 14.5-fold; and #505, threefold). For proteins preferentially expressed in CW, the highest fold-ratio was observed for a nascent polypeptide-associated complex (#1022, 24-fold).

**Genotypic and climatic effect on gene expression** Quantitative real-time PCR (qPCR) was used to check whether or not differentially expressed genes between the crown and the base of the 2003 samples showed similar trends in 2006 (for both fold-change and rank). This investigation, using two different biological replicates (collected in 2006) and a different transcriptomic technology (qPCR), was carried out for 14 genes identified as differentially expressed in the 2003 samples: five genes more expressed in the crown (GASA5 (BX250252), LP3 protein (BX253505), and three unknown function genes (BX252990; BX255469, BX248868)) and nine genes more expressed in the base (17.6 kD HSP (BX251102), gamma-thionin (BX254131), ATP-binding cassette (BX249573), homeobox STM (BX254965), betaine-aldehyde dehydrogenase (BX253899), phytoceramidase (BX249766), geranylgeranyl-diphosphate synthase (BX254083), SEC13-putative protein (BX249772) and cellulose synthase (BX249248)). qPCR results are presented as supplementary data (Fig. S6). While some genes (whose expression could be linked to the ontogenic effect) clearly displayed the same trend of gene expression along the

stem between the 2003 and 2006 samples, such as GASA5, gamma-thionin, and smwHSP (genes of known function), others did not match the 2003 results at all. It is likely that these genes were either responding to environmental conditions or corresponded to different members of the same multigene family cross-hybridizing in reverse northern, for example, cellulose synthases. These interpretations will be discussed further in the following section.

## Discussion

We would like to highlight the importance of this study because it involves an extensive application of genomic tools (metabolome, transcriptome and proteome characterization) used to link phenotypic data with molecular mechanisms along the stem of wood-forming tissues. We only found a single reference (Cato *et al.*, 2006) that attempted to link phenotypic data (cell wall thickness and radial growth and the proportion of developing xylem cell types) with gene expression in wood-forming tissues associated with both BW and CW in *P. radiata*.

In this section we will illustrate how the phenotypic plasticity observed along a base to crown gradient of a single tree could be interpreted in terms of molecular variation as revealed by gene and protein expression profiles obtained along the stem. We will discuss to what extent these results can be generalized with other genotypes and climatic conditions.

### Anatomical and cell wall chemical composition versus gene expression along a base to crown gradient

**Carbohydrate and lignification-related genes are overexpressed in BW-forming tissue** As expected, the variation of cellulose content along the stem was similar to that found in fully developed wood; that is, cellulose content was higher at the

base of the stem where mature wood is formed. The increase in cellulose content results from the extension of secondary cell wall thickening. Cato *et al.* (2006) reported that over 30% of differentiating tracheids in BW-forming tissue produced secondary cell walls, compared with only 3% produced by the J/CW. Since secondary cell walls consist mainly of cellulose, the higher cellulose content in differentiating xylem at the base thus results in a thicker cell wall. Conversely, upper levels (L5 and L6) had lower cellulose content but higher content of hemicelluloses (Figs 2 and 3).

Two cellulose synthases (UDP-forming), *PpinCesa3* (<http://cellwall.stanford.edu/>) (BX250234, 1.8-fold) in cluster 1 and *PpinCesa1*<sup>1</sup> (BX249248, 2.8-fold) in cluster 4, were found to be preferentially expressed at the base of the tree, which is consistent with the higher cellulose content found in this tissue. *PpinCESA3* is 98% identical to *PrCESA3*, and *PpinCESA1* is 97% identical to *PrCESA2*. *PrCESA3* is an orthologue of *AtCESA7* (IRX3) and *PrCESA1* is an orthologue of *AtCESA8*, two genes known to be involved in secondary xylem formation in *Arabidopsis thaliana*. The sequence homology of these two genes suggests that both *PpinCESA1* and *PpinCESA3* could be involved in secondary cell wall thickening. Both transcripts were found to be differentially expressed in a panel of eight tissues of maritime pine (Paiva, 2006), but while *PpinCESA1* was more expressed in differentiating xylem, *PpinCESA3* was found to be preferentially expressed in other tissues. These results suggest different roles of these two cellulose synthases during the formation of the secondary cell wall. Proteins related to carbohydrate metabolism were also identified in the BW clusters, namely one UDP-glucose dehydrogenase (spot #276, 75-fold) and one UTP-glucose-1-phosphate uridylyltransferase (spot #225, fivefold). Both enzymes have important roles in cell wall formation in higher plants. UTP-glucose-1-phosphate uridylyltransferase produces UDP-glucose, which can be used in the biosynthesis of cellulose (reviewed by Kleczkowski *et al.*, 2004). UDP-glucose can then be used by the UDP-glucose dehydrogenase to form UDP-glucuronate, a key precursor in hemicellulose and pectin formation (Seitz *et al.*, 2000).

At the proteomic level we found that enzymes implicated in methionine metabolism were highly up-regulated in BW-forming tissues: namely methionine synthase (44-fold), adenosylhomocysteinase (6.4- and 1.5-fold), adenosylmethionine synthetase (SAM-S, 5.4-fold), and homocysteine methyltransferase (7.1-fold). Methionine is involved in methyltransferase reactions as S-adenosylmethionine (SAM). The preferential abundance of these proteins in BW-forming tissues may reflect the higher demand for methyl transfer reactions required for the biosynthesis of monolignols. The importance of methylation associated with lignin biosynthesis is also corroborated by the higher expression of one of the lignification enzymes, CCoAOMT (2.5-fold), in BW-forming tissue. In terms of absolute value, more lignin is expected in BW-forming tissues because of the thicker cell walls of BW tracheids.

**Prolongation of cell wall thickening in BW-forming tissue involved the up-regulation of defence-related genes** One of the characteristics of tracheids from stem base is the presence of a thicker secondary cell wall (Zobel & Sprague, 1998). This thicker cell wall is more related to the duration of cell wall thickening than to the rate of cell wall biosynthesis (Larson *et al.*, 2001; Ugglia *et al.*, 2001; Cato *et al.*, 2006).

We found that five genes of the 'stress response' category accumulated preferentially in BW-forming tissues (clusters 1 and 4 of Fig. 6), including a 17.6 kDa class I heat-shock protein (BX251102, 2.6-fold), a gamma-theonin precursor/defensin (BX254131, 2.2-fold), a late embryogenesis protein (LEA; BX255424, 1.9-fold), a DNA damage regulation gene DDR48-stress protein (BX248943, 1.5-fold), and a disease resistance protein-like (BX255778, 1.5-fold).

In plants, low-molecular-weight heat-shock proteins (smwHSP) accumulate in response to various stresses (Vierling, 1991; Waters *et al.*, 1996; Costa *et al.*, 1998), and also seem to have specific roles in developmental processes, including seed maturation, somatic embryogenesis and wood formation (Puigderrajols *et al.*, 1996, 2002; Pla *et al.*, 1998; Le Provost *et al.*, 2003). In animals, expression of smwHSP during the transition of cell division to differentiation has been related to a preventive role in differentiating cells from undergoing apoptosis (Arrigo, 2005). Gion *et al.* (2005) also observed that smwHSP proteins accumulated in BW-forming tissues and suggested that they played a role in prolonging the cell wall thickening phase of xylogenesis, by delaying entry into programmed cell death (PCD). Late embryogenesis proteins are a major group of proteins which are extremely stable and hydrophilic. They typically accumulate during the late stages of embryogenesis or in response to dehydration, low temperature, salinity or exogenous abscisic acid (ABA) treatments, thus indicating their responsiveness to cellular dehydration (Ramanjulu & Bartels, 2002). It has been proposed that LEA stabilize membranes and prevent crystallization of cellular components (Dure *et al.*, 1989; Garay-Arroyo *et al.*, 2000). Thionins are small, basic, cysteine-rich proteins, which may function as defence molecules against an array of plant pathogens (Florack & Stiekema, 1994; Broekaert *et al.*, 1995). We report here for the first time on the up-regulation of these two genes in BW-forming tissues.

In agreement with the transcriptomic data, we also found that defence-related proteins were preferentially expressed in BW-forming tissues such as heat-shock proteins (spot #974, spot #1163, and spot #995, 85- 18- and 10-fold variation, respectively), peroxidases (spot #1033, spot #1107, and spot #445, 26-, 5.1- and 2.4-fold variation, respectively), universal stress protein (spot #805, 13-fold) and one chaperonin (spot #184, 2.7-fold).

The preferential expression of a higher number of stress-responsive proteins and other molecular mechanisms involved in PCD in BW-forming tissues fully agree with the results previously obtained at the proteome level by Gion *et al.*

(2005). Furthermore, this is consistent with the hypothesis that these genes could be involved in the delay of PCD, and thus in the prolongation of cell wall deposition, therefore resulting in the higher cell wall thickness and wood density characteristic of BW.

**'Protein synthesis' and 'energy' related genes are more expressed in CW-forming tissue** The greater protein content observed at the top of the tree may be the result of a higher rate of cell division, with more cells in the division and expansion stages, as shown by the microscopic observations. This result is consistent with that found in *P. radiata*, where cell division was 3.3 times greater in cambium of the crown compared with that of the base, even during the period of late wood formation (Cato *et al.*, 2006). In Scots pines (*Pinus sylvestris*), Uggla *et al.* (2001) also reported that the rate and time of cell division and expansion were greater at the top of the tree than at the bottom.

The youngest vascular cambium is associated with CW formation and is characterized by a high rate of cell division and expansion (Uggla *et al.*, 2001; Cato *et al.*, 2006; this study). At the protein level we found two essential components of the cytoskeleton, an actin (spot #495, 6.2-fold) and a tubulin ( $\beta$ -2-tubulin; spot #718, 2.6-fold) preferentially expressed in CW-forming tissues. These cytoskeleton-related proteins might be involved in the control of cell division. In particular, actin filaments are responsible for many aspects of cell behaviour, such as cell division, intracellular movement and cell expansion.

Cell division and expansion are highly demanding in terms of protein synthesis and energy requirements, as demonstrated by the higher protein content and number of genes included in cluster 2 (Fig. 6) classified into the 'protein synthesis' category (13 genes, mainly ribosomal proteins with fold-change ratios ranging from 1.6 to 2.3, while cluster 1 comprised only three ribosomal proteins with fold-change ratios of 1.6, 1.9 and 2). Interestingly, the protein with the highest CW/BW ratio was a nascent polypeptide-associated complex (spot #1022, 24-fold), demonstrating the importance of protein synthesis in CW. Similarly, but with a fold-change ratio equal to or less than 2, we found four genes up-regulated in the CW-forming tissues, suggesting a high demand for energy and metabolic needs. These genes were cytochrome-c oxidase (BX252991, twofold), ferredoxin III (BX252874, 1.8-fold), naphthoate synthase (EC 4.1.3.36, ubiquinone biosynthesis pathway, BX250373, 1.7-fold), and cytochrome c (BX251935, 1.6-fold). At the protein level, we also found two proteins of the 'energy' category: malate dehydrogenase (spot #728, 2.3-fold) and a RuBisCo subunit binding protein (spot #188, 1.6-fold).

**GA and ABA responsive genes are overexpressed in CW-forming tissue** A gene encoding a GASA5-like protein (BX250252; TC58624, 2.6-fold) was up-regulated in CW-forming tissues. The *GASA* (for GA-stimulated) belongs to a widespread class of genes found in mono- and dicotyledonous

plants. They are all structurally related to the original gibberellins (GA) regulated *GAST1* (GA-stimulated transcript 1) gene from tomato (Aubert *et al.*, 1998). Gibberellins (GAs) constitute a class of tetracyclic diterpenoid acids involved in the regulation of major plant growth and developmental processes such as germination, cell elongation, expansion and division, flowering and fruit development (Richards *et al.*, 2001). GAs act synergistically with auxin in stimulating cambial growth (Israelsson *et al.*, 2005) and to provide positional information during wood development (Tuominen *et al.*, 1997; Uggla *et al.*, 1998, 2001). Homologues of *GAST1* have been shown to be differentially expressed in different tissues of several species, for example: (i) histochemical analysis revealed that the *GASA4* promoter is active in meristematic regions, suggesting a role in cell division (Aubert *et al.*, 1998); (ii) *GEG*, a gerbera homolog of *GAST1*, was found to inhibit petal cell expansion, but to promote radial cell expansion (Kotilainen *et al.*, 1999); (iii) the tomato *RS1* is expressed in roots and induced by auxin (Taylor & Scheuring, 1994); (iv) *GIPs* (GA-induced genes of *Petunia*) have been implicated in shoot elongation and transition to flowering (Ben-Nissan *et al.*, 2004); (v) in poplar, Israelsson *et al.* (2005) showed that the expression of a GIP-like1 gene dramatically increased (several hundred-fold) in expansion zones of wood-forming tissues, where the concentration of bioactive GA was highest.

Another gene that could be implicated in cell wall expansion, is an orthologue of LP3 (ASR) (BX253505; TC73497, 2.7-fold) that was found to be up-regulated in CW-forming tissues. LP3 is a water deficit-induced protein, which is highly homologous to ASR (ABA, stress and ripening proteins). ASR is a small gene family demonstrating similarity with ABA, stress and ripening genes (Padmanabhan *et al.*, 1997). Le Provost (2003) also reported in *P. pinaster* that a different member of the LP3 family, the *P. pinaster* LP3-1 (AJ300726, TC73079), was up-regulated 12-fold in BW compared with CW-forming tissues. The different expression patterns of these two LP3 members are likely to be the result of different functional roles. As LP3 genes are highly hydrophilic (Padmanabhan *et al.*, 1997) they might be implicated in the regulation of cell water potential during the cell expansion phase.

A BLAST search of BX255469 against the EMBL public database revealed this accession matched with a gene of unknown function in *Oryza sativa* (accession AK062043,  $e$ -value =  $9e^{-37}$ ) and *Arabidopsis thaliana* (accession NM127873,  $e$ -value =  $1e^{-33}$ ). These genes were found to be implicated in gibberellin (Matsui *et al.*, 2005) and drought stress (Yazaki *et al.*, 2004) responses in rice and *Arabidopsis* respectively. The coregulation of GASA5, LP3 and BX255469 suggests a main role of the cross-talk between GA- and ABA-responsive genes in the regulation of crown wood formation, in particular during the cell division and expansion stages.

**Genetic and climatic effects on phenotypic and molecular 'base to crown' variation** An important finding of this study

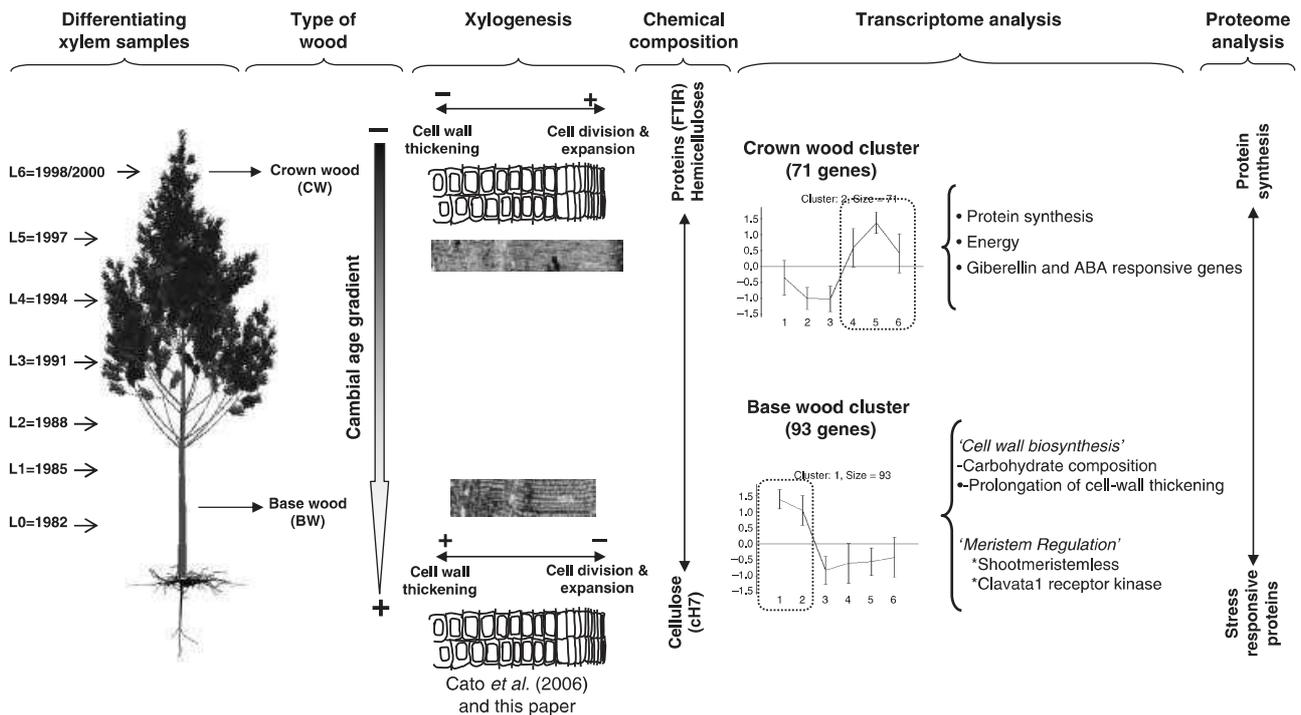


Fig. 9 Summary of the main results obtained in this study. Tree image courtesy of T. Fourcaud (AMAP CIRAD) and F. Danjon (BIOGECO INRA).

carried out in natural environment, was the lack of agreement (at both the chemical and molecular levels) between the 2003 and 2006 experiments. These differences were attributed to the effect of different climatic conditions at the date of sampling. Indeed, the year of sampling ranks first among other factors (i.e. genotype and position along the trunk) to explain the variability of chemical composition between BW and CW. Temperature and water availability are known to be among the most significant external sources of variation affecting cambial activity and the differentiation of newly divided cells, ultimately influencing wood characteristics. At the phenotypic level, it is well known that environmental factors (including edaphic and climatic variables) affect wood characteristics (Lipshitz & Waisel, 1970; Zobel & van Buijtenen, 1989; Gindl *et al.*, 2000; Rozenberg *et al.*, 2002). At the molecular level it is also clear that gene and protein expression in wood-forming tissue varies according to environmental clues, as demonstrated in pine during the growing season (Le Provost *et al.*, 2003; Egertsdotter *et al.*, 2004; Gion *et al.*, 2005; Yang & Loopstra, 2005).

In this study, 65% of the genes analysed by qPCR in the 2006 samples presented different patterns from that found in 2003 (Fig. S6). Interestingly, both trees analysed in 2006 presented similar expression profiles, again showing that the genotypic effect was probably low. However, because different gene expression technologies were used (reverse-northern vs qPCR), we cannot exclude the possibility that the differences observed between years were also the result of confounding

technical (e.g. cross-hybridizations of multiple gene family members on the macroarrays) and experimental setup (e.g. number of replicates) effects. Despite this general lack of coincidence between years, few genes, such as gamma-thionin (up-regulated in BW) or a GASA-5-like protein (up-regulated in CW), showed similar trends in 2003 and 2006, suggesting that their regulation was mainly affected by ontogenic effect.

### Concluding remarks

Samples collected from the base to the crown of an adult maritime pine tree were characterized in terms of cell wall chemical composition, transcript and protein accumulation (Fig. 9). Cell wall composition reflected the characteristics of fully developed wood, in respect to cellulose content; that is, cellulose content was higher in differentiating xylem samples at the stem base. The higher protein content found in differentiating xylem collected in the crown indicated intense protein synthesis, which might be related to numerous cell divisions and intense xylem differentiation activity. However, we found high variability between trees collected in different years, much higher than that found between different genotypes sampled in the same year, reflecting a strong environmental effect on xylogenesis.

We also provided novel insights into the patterns of protein and transcript variation of genes involved in wood formation. Out of 3512 transcripts and 1372 proteins, 262 and 231 were found to be differentially expressed along the stem, respectively.

In CW-forming tissue, more than 45% of the genes of known function coded for protein synthesis (mainly ribosomal proteins), corroborating the high protein content found in this tissue. This study also indicates a potential role of stress response-related genes in BW formation, and suggests that they could be implicated in delaying PCD, in agreement with the previous hypothesis formulated by Gion *et al.* (2005). Additionally, our study also points out the role of gibberellins/ASR-responsive genes in xylogenesis.

Finally, this study provides us with a list of potential candidate genes whose polymorphisms could be involved in the genetic control or regulation of wood quality components. The effects of nucleotide polymorphism will be further tested in association studies (Cardon & Bell, 2001). Maritime pine is indeed a potentially good candidate species for association mapping, because of the rapid drop of linkage disequilibrium along genes (Garnier-Géré *et al.*, 2005), similar to the patterns observed in *Pinus taeda* (Neale & Savolainen, 2004).

## Acknowledgements

This research was supported by grants from the European Union (GEMINI QLK-5-CT1999-00942), ANR Génoplante (GENOQB, GNP05013C), the Aquitaine Region, FCT (Fundação para a Ciência e Tecnologia, Portugal), POCTI and FEDER within research projects POCTI/AGR/47353/2002. JAPP was supported by fellowship SFRH/BD/3129/2000 and SFRH/BPD/26552/2006 from FCT/MCT Portugal. MG was supported by ALFA programme II-0266-FA (GEMA, European Union) and by postgraduate fellowship from Universidad de Talca, Chile. We thank Alexia Stokes for her review and Laura Graham for improving the English.

## References

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389–3402.
- Andersson-Gunneras S, Mellerowicz EJ, Love J, Segerman B, Ohmiya Y, Coutinho PM, Nilsson P, Henrissat B, Moritz T, Sundberg B. 2006. Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *Plant Journal* 45: 144–165.
- Arrigo AP. 2005. In search of the molecular mechanism by which small stress proteins counteract apoptosis during cellular differentiation. *Journal of Cellular Biochemistry* 94: 241–246.
- Aubert D, Chevillard M, Dorne AM, Arlaud G, Herzog M. 1998. Expression patterns of *GASA* genes in *Arabidopsis thaliana*: the *GASA4* gene is up-regulated by gibberellins in meristematic regions. *Plant Molecular Biology* 36: 871–883.
- Ben-Nissan G, Lee JY, Borohov A, Weiss D. 2004. GIP, a *Petunia hybrida* GA-induced cysteine-rich protein: a possible role in shoot elongation and transition to flowering. *Plant Journal* 37: 229–238.
- Broekaert WF, Terras FRG, Cammue BPA, Osborn RW. 1995. Plant defensins: Novel antimicrobial peptides as components of the host defense system. *Plant Physiology* 108: 1353–1358.
- Cardon LR, Bell JI. 2001. Association study designs for complex diseases. *Nature Reviews Genetics* 2: 91–99.
- Cato S, McMillan L, Donaldson L, Richardson T, Echt C, Gardner R. 2006. Wood formation from the base to the crown in *Pinus radiata*: Gradients of tracheid wall thickness, wood density, radial growth rate and gene expression. *Plant Molecular Biology* 60: 565–581.
- Chang S, Puryear J, Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11: 113–116.
- Corpet F, Gouzy J, Kahn D. (1998) The ProDom database of protein domain families. *Nucleic Acids Research* 26: 323–326.
- Costa P, Bahrman N, Frigerio JM, Kremer A, Plomion C. 1998. Water-deficit-responsive proteins in maritime pine. *Plant Molecular Biology* 38: 587–596.
- Derory J, Mariette S, González-Martínez SC, Chagné D, Madur D, Gerber S, Brach J, Persyn F, Ribeiro MM, Plomion C. 2002. What can nuclear microsatellites tell us about maritime pine genetic resources conservation and provenance certification strategies? *Annals of Forest Science* 59: 699–708.
- Dumail JF, Castéra P, Morlier P. 1998. Hardness and basic density variation in the juvenile wood of maritime pine. *Annales des Sciences Forestières* 55: 911–923.
- Dure L III, Crouch M, Harada J, Ho TH, Quatrano R, Thomas T, Sung ZR. 1989. Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology* 25: 475–486.
- Egertsdotter U, van Zyl LM, Mackay J, Peter G, Kirst M, Clark C, Whetten R, Sederoff R. 2004. Gene expression during formation of earlywood and latewood in loblolly pine: expression profiles of 350 genes. *Plant Biology* 6: 654–663.
- Ewing RM, Kahla AB, Poirat O, Lopez F, Audic S, Claverie JM. 1999. Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Research* 9: 950–959.
- Faix O, Bremer J, Schmidt O, Stevanovic T. 1991a. Monitoring of chemical changes in white-rot degraded beech wood by pyrolysis-gas chromatography and Fourier-transform infrared spectroscopy. *Journal of Analytical and Applied Pyrolysis* 21: 147–162.
- Faix O, Fortmann I, Bremer J, Meier D. 1991b. Thermal-degradation products of wood – a collection of electron-impact Ei mass-spectra of polysaccharide derived products. *Holz Als Roh-Und Werkstoff* 49: 299–304.
- Faix O, Meier D, Fortmann I. 1990a. Thermal-degradation products of wood – a collection of electron-impact Ei mass-spectra of monomeric lignin derived products. *Holz Als Roh-Und Werkstoff* 48: 351–354.
- Faix O, Meier D, Fortmann I. 1990b. Thermal-degradation products of wood – gas-chromatographic separation and mass-spectrometric characterization of monomeric lignin derived products. *Holz Als Roh-Und Werkstoff* 48: 281–285.
- Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJA, Hofmann K, Bairoch A. 2002. The PROSITE database, its status in 2002. *Nucleic Acids Research* 30: 235–238.
- Florack DEA, Stiekema WJ. 1994. Thionins: properties, possible biological roles and mechanisms of action. *Plant Molecular Biology* 26: 25–37.
- Fonseca F, Lousada J. 2000. [Wood variation in *Pinus pinaster* Ait.] Variação da Madeira de *Pinus pinaster* Ait. *Série Técnicas – Científicas, Ciências Aplicadas* 35, Vila Real, Portugal: UTAD.
- Garay-Arroyo A, Colmenero-Flores JM, Garcarrubio A, Covarrubias AA. 2000. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *Journal of Biological Chemistry* 275: 5668–5674.
- Garnier-Géré PH, Austerlitz F, Bedon F, Kremer A, Plomion C. 2005. Effets de la sélection sur la diversité et la différenciation génétiques moléculaires: résultats de simulations et application au pin maritime pour les gènes de la lignification. *Les Actes du BRG* 5: 275–291.

- Gindl W, Grabner M, Wimmer R. 2000. The influence of temperature on latewood lignin content in treeline Norway spruce compared with maximum density and ring width. *Trees – Structure and Function* 14: 409–414.
- Gion J-M, Lalanne C, Le Provost G, Ferry-Dumazet H, Paiva J, Frigerio J-M, Chaumeil P, Barré A, de Daruvar A, Brach J. *et al.* 2005. The proteome of maritime pine wood forming tissue. *Proteomics* 5: 3731–3751.
- González-Martínez S, Gerber S, Cervera M, Martínez-Zapater J, Gil L, Alía R. 2002. Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theoretical and Applied Genetics* 104: 1290–1297.
- Groover AT. 2005. What genes make a tree a tree? *Trends in Plant Science* 10: 210–214.
- Israelsson M, Sundberg B, Moritz T. 2005. Tissue-specific localization of gibberellins and expression of gibberellin-biosynthetic and signaling genes in wood-forming tissues in aspen. *Plant Journal* 44: 494–504.
- Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28: 27–30.
- Kleczkowski LA, Geisler M, Cieresko I, Johansson H. 2004. UDP-glucose pyrophosphorylase. An old protein with new tricks. *Plant Physiology* 134: 912–918.
- Kotilainen M, Helariutta Y, Mehto M, Pollanen E, Albert VA, Elomaa P, Teeri TH. 1999. GEG participates in the regulation of cell and organ shape during corolla and carpel development in *Gerbera hybrida*. *Plant Cell* 11: 1093–1104.
- Larson PR, Kretschmann DE, Clark AI, Isebrands JG. 2001. *Formation and properties of juvenile wood in southern pines: a synopsis*. General Technical Report FPL-GTR-129. Madison, WI, USA: Department of Agriculture, Forest Service, Forest Products Laboratory.
- Le Provost G, Paiva J, Pot D, Brach J, Plomion C. 2003. Seasonal variation in transcript accumulation in wood forming tissues of maritime pine *Pinus pinaster* Ait. with emphasis on a cell wall Glycine Rich Protein. *Planta* 217: 820–830.
- Lev-Yadun S, Sederoff R. 2000. Pines as model gymnosperms to study evolution, wood formation, and perennial growth. *Journal of Plant Growth Regulation* 19: 290–305.
- Lipshitz N, Waisel Y. 1970. Effects of environment on relations between extension and cambial growth of *Populus euphratica* Oliv. *New Phytologist* 4: 1059–1064.
- Lorenz WW, Dean JFD. 2002. SAGE profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine *Pinus taeda*. *Tree Physiology* 22: 301–310.
- Mariette S, Chagné D, Lézier C, Pastuszka P, Raffin A, Plomion C, Kremer A. 2001. Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity* 86: 469–479.
- Matsui K, Hiratsu K, Koyama T, Tanaka H, Ohme-Takagi M. 2005. A chimeric AtMYB23 repressor induces hairy roots, elongation of leaves and stems, and inhibition of the deposition of mucilage on seed coats in *Arabidopsis*. *Plant Cell Physiology* 46: 147–155.
- Moldoveanu SC. 1998. *Analytical pyrolysis of natural organic polymers*. Techniques and Instrumentation in Analytical Chemistry. Amsterdam, the Netherlands: Elsevier Science.
- Nakajima K, Benfey PN. 2002. Signaling in and out: control of cell division and differentiation in the shoot and root. *Plant Cell Supplement* 2002: S265–S276.
- Neale DB, Savolainen O. 2004. Association genetics of complex traits in conifers. *Trends in Plant Science* 9: 325–30.
- O'Farrell 1975. High resolution two-dimensional electrophoresis. *The Journal of Biological Chemistry* 250: 4007–4021.
- Padmanabhan V, Dias DMAL, Newton RJ. 1997. Expression analysis of a gene family in loblolly pine (*Pinus taeda* L.) induced by water deficit stress. *Plant Molecular Biology* 35: 801–807.
- Paiva JABCP. 2006. Phenotypic and molecular plasticity of wood forming tissues in Maritime Pine (*Pinus pinaster* Ait.). *Joint doctoral thesis (PhD)*. Bordeaux, France: Université de Bordeaux I. Lisbon, Portugal: Universidade Nova de Lisboa [<http://www.pierroton.inra.fr/biogeco/genetique/recherches/pin/genomique/expression/Jorge/index.html>]
- Pla M, Huguet G, Verdaguer D, Puigderrajols P, Llompарт B, Nadal A, Molinas M. 1998. Stress proteins co-expressed in suberized and lignified cells and in apical meristems. *Plant Science* 139: 49–57.
- Plomion C, Pionneau C, Brach J, Costa P, Baillères H. 2000. Compression wood-responsive proteins in developing xylem of maritime pine *Pinus pinaster* Ait. *Plant Physiology* 123: 959–969.
- Puigderrajols P, Fernández-Guijarro B, Toribio M, Molinas M. 1996. Origin and early development of secondary embryos in *Quercus suber* L. *International Journal of Plant Sciences* 157: 674–684.
- Puigderrajols P, Jofre A, Mir G, Pla M, Verdaguer D, Huguet G, Molinas M. 2002. Developmentally and stress-induced small heat shock proteins in cork oak somatic embryos. *Journal of Experimental Botany* 53373: 1445–1452.
- R Development Core Team. 2004. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Radi M. 1992. Analyse morphologique de l'arbre en vue de sa modélisation mécanique. *Doctoral thesis*. Bordeaux, France: Université Bordeaux I.
- Ralph J, Hatfield RD. 1991. Pyrolysis-Gc-MS characterization of forage materials. *Journal of Agricultural Food Chemistry* 20: 1426–1437.
- Ramagli L, Rodriguez L. 1985. Quantitation of microgram amounts of protein in two-dimensional polyacrylamide gel electrophoresis sample buffer. *Electrophoresis* 6: 559–563.
- Ramanjulu S, Bartels D. 2002. Drought- and desiccation-induced modulation of gene expression in plants. *Plant, Cell & Environment* 25: 141–151.
- Ribeiro MM, Mariette S, Vendramin GG, Szmidi AE, Plomion C, Kremer A. 2002. Comparison of genetic diversity estimates within and among populations of maritime pine using cpSSR and AFLP data. *Molecular Ecology* 11: 869–877.
- Richards DE, King KE, Ait-ali T, Harberd NP. 2001. How gibberellin regulates plant growth and development: a molecular genetic analysis of gibberellin signaling. *Annual Review of Plant Physiology and Plant Molecular Biology* 52: 67–88.
- Rodrigues J, Graca J, Pereira H. 2001. Influence of tree eccentric growth on syringyl/guaiacyl ratio in *Eucalyptus globulus* wood lignin assessed by analytical pyrolysis. *Journal of Analytical and Applied Pyrolysis* 58: 481–489.
- Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. *Bioinformatics methods and protocols: methods in molecular biology*. Totowa, NJ, USA: Humana Press, 365–386.
- Rozenberg P, Van Loo J, Hannrup B, Grabner M. 2002. Clonal variation of wood density record of cambium reaction to water deficit in *Picea abies* (L.) Karst. *Annals of Forest Science* 59: 533–540.
- Savidge RA, Wareing PF. 1984. Seasonal cambial activity and xylem development in *Pinus contorta* in relation to endogenous indol-3-yl-acetic and S-abscisic acid levels. *Canadian Journal of Forest Research* 14: 676–682.
- Seitz B, Klos C, Wurm M, Tenhaken R. 2000. Matrix polysaccharide precursors in *Arabidopsis* cell walls are synthesized by alternate pathways with organ-specific expression patterns. *Plant Journal* 21: 537–546.
- Shamir R, Maron-Katz A, Tanay A, Linhart C, Steinfeld I, Sharan R, Shiloh Y, Elkon R. 2005. EXPANDER – an integrative program suite for microarray data analysis. *BMC Bioinformatics* 6: 232.
- Sharan R, Shamir R. 2000. CLICK: A clustering algorithm with applications to gene expression analysis. In: *Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology ISMB*. Menlo Park, CA, USA: AAAI Press, 307–316.

- Sundberg B, Little CHA, Cui K, Sandberg G. 1991. Level of indole-3-acetic acid in the stem of *Pinus sylvestris* in relation to seasonal variation of cambial activity. *Plant, Cell & Environment* 14: 241–246.
- Taylor BH, Scheuring CF. 1994. A molecular marker for lateral root initiation: the rsi-1 gene of tomato (*Lycopersicon esculentum* Mill) is activated in early lateral root primordia. *Molecular & General Genetics* 243: 148–157.
- Tuominen H, Puech L, Fink S, Sundberg B. 1997. A radial concentration gradient of indole-3-acetic acid is related to secondary xylem development in hybrid aspen. *Plant Physiology* 115: 577–585.
- Uggla C, Magel E, Moritz T, Sundberg B. 2001. Function and dynamics of auxin and carbohydrates during early/latewood transition in Scots pine. *Plant Physiology* 125: 2029–2039.
- Uggla C, Mellerowicz EJ, Sundberg B. 1998. Indole-3-acetic acid controls cambial growth in *Pinus sylvestris* L. by positional signaling. *Plant Physiology* 117: 113–121.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3: RESEARCH0034. Epub June 18 2002.
- Vierling E. 1991. The roles of heat shock proteins in plants. *Annual Review of Plant Physiology* 42: 579–620.
- Waters ER, Lee JL, Vierling E. 1996. Evolution, structure and function of the small heat shock proteins in plants. *Journal of Experimental Botany* 47: 325–338.
- Westbrook J, Yan J, Wait R, Welson S, Dunn M. 2001. Zooming-in on the proteome: very narrow-range immobilised pH gradients reveal more protein species nad isoforms. *Electrophoresis* 22: 2865–2871.
- Whetten R, Sun YH, Zhang Y, Sederoff RR. 2001. Functional genomics and cell wall biosynthesis in loblolly pine. *Plant Molecular Biology* 47: 275–291.
- Yang SH, Loopstra CA. 2005. Seasonal variation in gene expression for loblolly pines *Pinus taeda* L. from different geographical regions. *Tree Physiology* 25: 1063–1073.
- Yazaki J, Shimatani Z, Hashimoto A, Nagata Y, Fujii F, Kojima K, Suzuki K, Taya T, Tonouchi M, Nelson C. *et al.* 2004. Transcriptional profiling of genes responsive to abscisic acid and gibberellin in rice: phenotyping and comparative analysis between rice and *Arabidopsis*. *Physiology Genomics* 17: 87–100.
- Zobel BJ, van Buijtenen JP. 1989. *Wood variation: its causes and control*. Berlin, Germany: Springer-Verlag.
- Zobel B, Kellison R, Matthias M, Hatcher AV. 1972. *Wood density of southern pines*. Technical Bulletin No. 203. Raleigh, NC, USA: North Carolina Agricultural Experiment Station.
- Zobel BJ, Sprague JR. 1998. *Juvenile wood in forest trees*. Berlin, Germany: Springer-Verlag.

## Supplementary Material

The following supplementary material is available for this article online:

**Fig. S1** Sampling procedure of differentiating xylem.

**Fig. S2** Template grid used for spotting and signal quantification.

**Fig. S3** Anatomical observations of wood-forming tissue.

**Fig. S4** Min-max normalized spectra of fully developed wood (red spectrum) and differentiating xylem (blue spectrum, sample L6) of maritime pine from 4000 to 400 cm<sup>-1</sup>.

**Fig. S5** Pyrograms of differentiating xylem samples (genotypes sampled in 2003) taken along the trunk (level L1 (base) to L6 (crown)) showing the main differences between both samples, in terms of protein (aa1) and cellulose content (CH7, levoglucosane).

**Fig. S6** qPCR expression profiles in the 2006 samples (genotypes A and B) for selected 'base' (a–i) and 'crown' (j–n) up-regulated genes in the 2003 samples.

**Table S1** Origin of the differentiating xylem samples used for cDNA library construction

**Table S2** List of qPCR primers pairs

**Table S3** Analytical pyrolysis data of the 2003 samples

**Table S4** Transcriptomic data of the 2003 samples

**Table S5** Proteomic data of the 2003 samples

This material is available as part of the online article from:

<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1469-8137.2008.02379.x>  
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the journal at *New Phytologist* Central Office.

Table S5: Overexpressed proteins in base and crown wood forming tissue of maritime pine

OVEREXPRESSED IN BASE WOOD				OVEREXPRESSED IN CROWN WOOD			
spot.ID	nb pro. (a)	% coverage (b)	TC	spot.ID	nb pro. (a)	% coverage (b)	TC
974	14	16.2	TC64985	1116	6	12.3	TC58057
	10	19.7	TC73231		20	30.7	TC66404
	6	12.6	TC73405		3	6.2	TC57833
1163	4	9.0	TC65842		15.8	TC58604	
	3	10.2	TC73213		13.0	TC58205	
995	10	14.8	TC64985		7.8	TC73537	
	3	14.4	TC65048		16.3	TC74096	
	4	8.7	TC73231		4.8	TC59225	
	3	10.2	TC75123		20.9	TC58167	
1033	11	32.0	TC65955		15.3	TC65958	
	2	12.3	TC76630		6.2	TC72878	
1107	4	14.5	TC59008		2.9	TC73522	
445	4	14.1	TC66142		2	4.4	TC74635
	5	5.5	TC58167		17.4	TC73559	
	28.5	4.8	TC73481		14	24.9	TC57392
184	15	25.5	TC58627		14	31.1	TC58550
	3	4.8	TC73481		16.7	TC67616	
	3	4.8	TC60591		3	17.1	
805	3	12.3	TC73526				
	10	14.4	TC57052				
352	3	7.8	TC73526				
	9	11.5	TC58167				
282	9	20.5	TC58167				
405	10	20.5	TC72880				
39	13	14.3	TC72880				
1057	13	19.4	TC64985				
	5	9.0	TC58314				
804	4	9.3	TC65955				
	2	7.9	TC65515				
517	2	5.8	TC74635				
	5	9.3	TC57104				
276	14	24.1	TC68264				
147	2	7.8	TC74687				
225	7	16.3	TC58505				
	15.2	9.5	TC57034				
290	8	9.5	TC57034				
	9.0	7.1	TC58024				
	3.4	5.07E-08	TC57316				
	9.8	2.19E-05	TC57671				
641	15.7	5.90E-03	TC59063				
	16.0	1.73E-10	TC74067				
	9.5	1.67E-17	TC67217				
	2	5.2	TC67216				
1080	13.1	1.00E-07	TC68017				
256	3	8.3	TC73058				

OVEREXPRESSED IN BASE WOOD				OVEREXPRESSED IN CROWN WOOD			
spot.ID	scans (a)	% coverage (b)	TC	spot.ID	scans (a)	% coverage (b)	TC
1116	6	12.3	TC58057	1116	6	12.3	TC58057
96	20	30.7	TC66404	96	20	30.7	TC66404
	3	6.2	TC57833		3	6.2	TC57833
731	15.8	TC58604		731	15.8	TC58604	
	8	13.0	TC58205		8	13.0	TC58205
728	4	7.8	TC73537		4	7.8	TC73537
	16.3	TC74096			16.3	TC74096	
188	9	16.3	TC74096		9	16.3	TC74096
	3	4.8	TC59225		3	4.8	TC59225
483	8	20.9	TC58167		8	20.9	TC58167
	6	15.3	TC65958		6	15.3	TC65958
	2	6.2	TC72878		2	6.2	TC72878
505	2	2.9	TC73522		2	2.9	TC73522
	2	4.4	TC74635		2	4.4	TC74635
495	14	24.9	TC73559		14	24.9	TC73559
178	14	31.1	TC57392		14	31.1	TC57392
1022	4	16.7	TC58550		4	16.7	TC58550
706	3	17.1	TC67616		3	17.1	TC67616

(a) nb pro: number of different peptides identifying the protein.  
 (b) coverage: the percentage of aminoacids covered over the total protein aminoacids.  
 (c) p value for the TC identification, it must be less than 1.00E-03.  
 (d) Presents the relative abundance of different proteins in a mix. Its calculated by the Peppan function, using the area of the peaks, the values are estimatives only.  
 (e) Functional categories were assigned taking into account the Gene Ontology provided in the Pinus Gene Index (<http://compbio.dcf.harvard.edu/cgi-bin/ontology/gimain.pl?tab=pine>). The TC identifications correspond to the March 2006 assembly.  
 (f) For proteins over-expressed in base or crown wood, fold-change were calculated dividing the expression value of each sample level by expression level of L6 or L0, respectively.  
 (g) Mw in (Da) and pI were estimated with LMW calibration kit for SDS Electrophoresis (Amersham Biosciences)  
 (h) Mw in (Da) and pI were calculated with a tool available at URL <http://us.expasy.org/tools/peptide-mass.html>

#### **4.1.5 Análisis proteómico de un gradiente base a copa temporadas 2005 y 2006.**

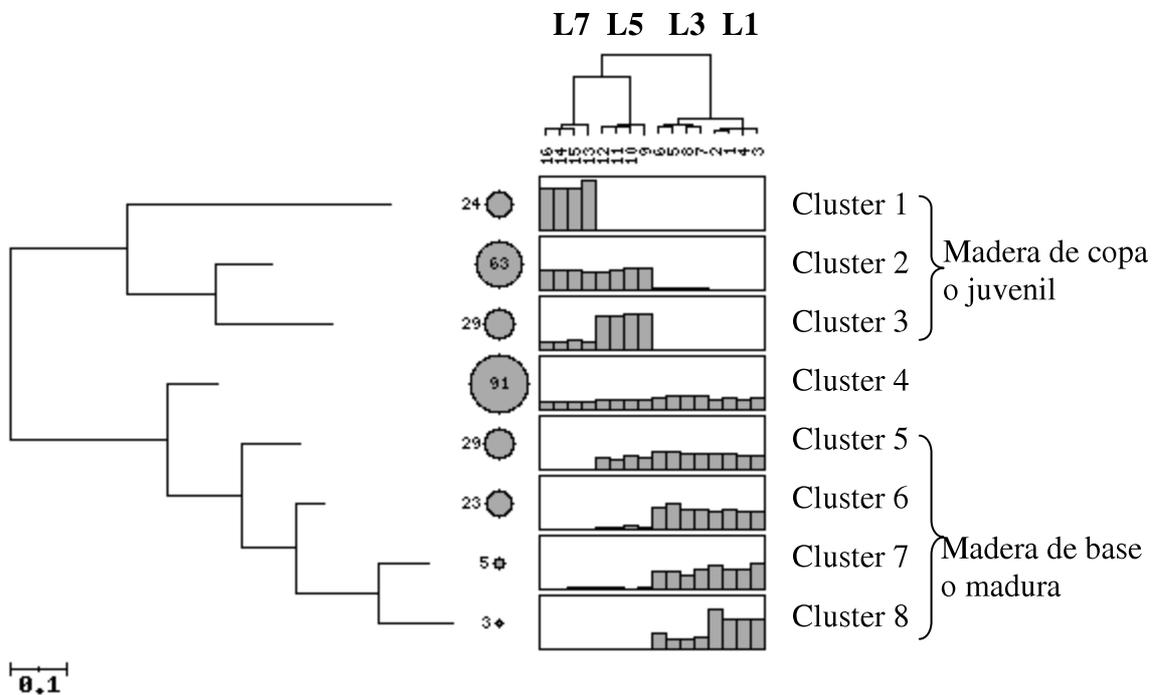
Luego del estudio de un gradiente base a copa para la temporada 2003 (Sección 4.1.4), donde en los análisis proteómicos se identificaron 33 proteínas sobreexpresadas, se analizó un gradiente similar en la temporada 2005, identificando 67 proteínas sobreexpresadas. A pesar de haber optimizado el número de proteínas identificadas, el estudio de solo un árbol en la temporada 2005 limitó el análisis, ya que no es posible separar el efecto base o copa de un efecto clonal/genotipo, sin embargo, se obtuvo información relevante que es incluida en la presente sección.

En la temporada 2006 se analizó el proteoma expresado en tejido formador de madera de base o copa proveniente de dos árboles, para fijar el factor genotipo y el factor nivel, se identificaron 65 proteínas sobreexpresadas solamente debido al efecto nivel. (Sección 4.1.6).

##### **4.1.5.1 Mapas proteicos bidimensionales de tejido formador de madera, gradiente desde base a copa 2005.**

Se construyeron mapas proteicos bidimensionales, para realizar estudios proteómicos en un gradiente base a copa de tejido formador de madera en pino marítimo, se seleccionó un árbol recto de genotipo desconocido, proveniente de Aquitania. Las muestras fueron obtenidas desde 4 niveles en el gradiente base a copa (L1, L3, L5, L7) (desde mayor a menor edad cambial). Se analizó por ANOVA la expresión de 743 spots, identificando 267 proteínas con una diferencia significativa ( $p < 0,001$ ) entre los niveles de edad cambial. Previo al análisis de agrupamiento “clustering”, para agrupar las proteínas según su perfil de expresión, la intensidad de cada spot fue centrada, dividiendo su intensidad por la intensidad promedio en todos los geles. El análisis de agrupamiento, generó 8 clusters desde madera de copa/juvenil a madera de base/madura (**Figura 28**). En el eje superior de la **Figura 28**, se representa el clustering transversal de los 16 geles analizados, 4 niveles (L1, L3, L5, L7), con 4 replicas numerados del 1 al 16. Este análisis de agrupamiento muestra la alta reproducibilidad de los replicados técnicos. En el eje vertical, los clusters 1, 2 y 3, con 116 spots, fueron considerados como sobre expresados en madera de juvenil, clusters 5, 6, 7 y 8, con 60 spots, fueron considerados como sobre expresados en madera de madura. El cluster 4 (91 spots) no fue considerado para análisis posteriores, por su débil cambio de expresión a lo largo del gradiente. Los spots sobre expresados en madera de

copa/juvenil y de base/madura fueron seleccionados para su identificación con espectrometría de masas ESI MS/MS (Electrospray masa/masa).



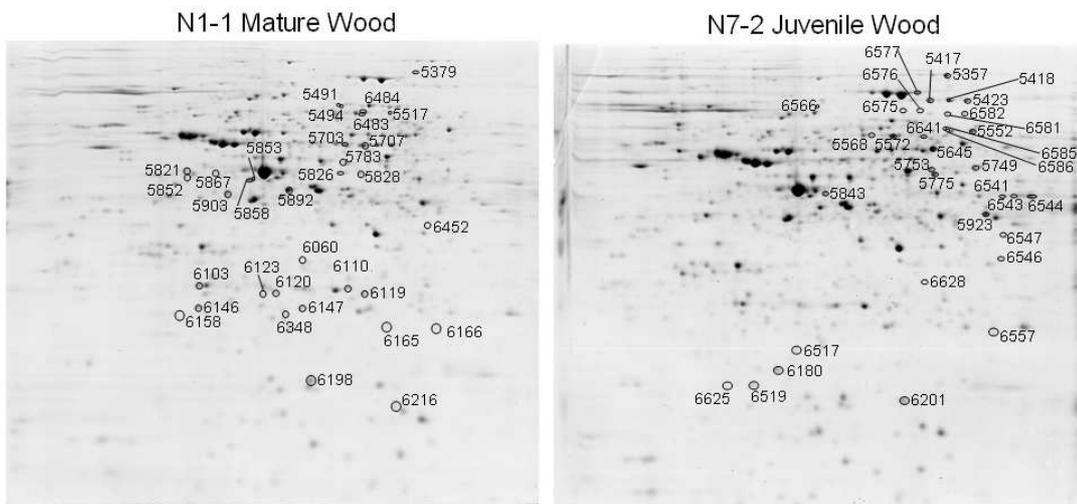
**Figura 28:** Agrupamiento de las proteínas con  $p < 0,001$ , luego de análisis estadístico se generaron 8 clusters.

#### 4.1.5.2 Identificación de proteínas por espectrometría de masas

Como criterio adicional se seleccionaron spots con variaciones constantes, de aumento o disminución a lo largo del gradiente. La lista de proteínas pre-seleccionados luego del análisis ANOVA y el agrupamiento, se verificó manualmente en cada gel, de este modo, se aseguró su resolución y reproducibilidad, al mismo tiempo se eligió el gel para extraer cada uno de ellos. Como análisis preliminar, se ensayó el límite de detección de la espectrometría de masas ESI MS/MS, analizando un gradiente de intensidad de tinción de spots, desde 0,03 a 0,3 en unidades de intensidad (desde proteínas apenas perceptibles a la vista, a proteínas fuertemente teñidas con Coomassie Blue). Se determinó 0,05 como el límite menor de detección, es decir, que proteínas débilmente teñidas con Coomassie Blue son capaces de generar péptidos detectables por espectrometría de masas ESI MS/MS. Estos spots/proteínas de menor intensidad fueron descartados del análisis. Por último, los spots fueron ordenados en intensidad creciente de tinción para su inyección en el espectrómetro de masas, con el fin de evitar la saturación de la columna al paso de las

muestras. Gracias a las modificaciones introducidas, se logró un muy alto porcentaje de identificación en la base de datos TIRG para las proteínas inyectadas (95 %).

La **Figura 29** muestra las proteínas seleccionadas en madera de base/madura (panel izquierdo) y madera de copa/juvenil (panel derecho), 33 y 34 proteínas fueron identificadas, respectivamente. Las proteínas identificadas en la **Tabla 5** se encuentran agrupadas por categorías funcionales siguiendo el catalogo MIPS (Munich Information Center for Protein Sequences, <http://mips.gsf.de/projects/funcat>).



**Figura 29:** Geles bidimensionales para tejido formador de madera en pino marítimo 2005. Panel izquierdo madera de base/madura, panel derecho madera de copa/juvenil

Table 4: Over-expressed proteins on mature and juvenile wood forming tissue of maritime pine 2005.

Mature Wood	spot ID	scans (a)	coverage	relative quantity	assignment	UniProtKB Entry	pl experimental/theoretic (kDa)	MW				Fold change			
								(b)	(c)	(%)	(g)	L1	L2	L3	L4
01 METABOLISM	01.01 amino acid metabolism														
	6120	4/4	100	TC64987	divicine hydroxymethyltransferase	Q8LBY1	01.01.09	6.6	31.77	51.71	27.6	25.0	10.4	1.0	6
	6452	11/11	18.8	TC66633	Aspartate aminotransferase, chloroplast [Precursor]	P46Z48	01.01.06.01.01	6.2	43.95	44.24	10.9	6.6	4.7	1.0	5
	5826	3/3	4.5	TC57104	Phosphoacacate dehydrogenase-like protein	Q8LQJ6	01.01.09.02.01	6.3	55.33	63.31	13.4	16.4	7.0	1.0	5
	5858	9/6	11.1	TC57104	Phosphoacacate dehydrogenase-like protein	Q94985	01.01.09.02.01	6.2	53.45	63.32	3.0	3.3	0.7	1.0	6
		2/2	6.9	TC57561	CBL-interacting protein kinase 23)	Q83VD3	01.01.09.02.01	9.3	53.45	53.51					
01.05 C-compound and carbohydrate metabolism	5703	11/9	10.4	TC73526	Adenosylhomocysteine	P68172	1.05	62.71	53.10	42.2	48.4	1.0	1.0	6	
		4/4	10.8	TC57052	Phosphopyruvate hydratase	Q8LEL0	1.05	62.71	47.83						
	6483	6/6	7.9	TC72880	Cobalamin-independent methionine synthase	Q42699	1.05	61.15	84.86	18.3	12.2	1.0	1.0	8	
		2/2	2.8	TC72976	Pyruvate decarboxylase 1	Q84V95	5.9	71.55	65.13						
	6484	5/5	6.1	TC72880	Cobalamin-independent methionine synthase	Q42699	1.05	61.15	84.86	10.8	12.2	4.7	1.0	6	
		3/3	2.8	TC72976	Pyruvate decarboxylase 1	Q84V95	5.9	72.38	65.13						
	5517	7/7	8.0	TC72880	Cobalamin-independent methionine synthase	Q42699	1.05	61.15	84.86	10.1	19.7	1.0	1.0	6	
	6348	3/3	16.28	TC72880	Cobalamin-independent methionine synthase	Q8W0Q7	1.05	29.88	83.79	14.9	21.8	1.0	1.0	6	
	5707	13/11	11.7	TC73526	Adenosylhomocysteine	P68172	1.05	62.56	53.10	12.8	11.8	3.1	1.0	6	
		9/8	21.8	TC57052	Phosphopyruvate hydratase	Q8LEL0	5.9	5.6	62.56	47.83					
02 ENERGY	5494	11/9	14.8	TC74153	NADP-dependent malic enzyme (NADP-ME)	P51615	1.05	61.15	74.25	65.23	4.2	4.1	4.9	1.0	5
		2/2	3.5	TC59260	Alpha-L-arabinofuranosyl-beta-D-xylosidase isoenzyme	Q8W012	5.8	5.6	74.25	81.99					
02.01 glycolysis and gluconeogenesis	6103	5/5	14.6	TC58637	Triosephosphate isomerase, cytosolic	P48491	2.01	32.94	27.17	1.3	1.2	1.0	1.0	4	
	6060	2/2	4.21	TC74478	ATP synthase gamma chain, mitochondrial precursor	P26360	02.45.15	37.06	30.87	6.9	7.5	1.1	1.0	6	
10 CELL CYCLE AND DNA PROCESSING	5853	5/5	12.5	TC75106	Nuclear migration protein nudC	Q8Y266	10.03.04.09	5.3	53.81	38.24	1.8	1.6	1.1	1.0	4
	5828	3/3	8.1	TC58800	GLABRA2 expression modulator	Q8S8F8	11.02.03.04	5.9	55.38	32.21	2.7	2.6	1.7	1.0	4
11 TRANSCRIPTION	6119	8/5	17.3	TC65048	Alpha 2 subunit of 20S proteasome	Q8H652	14.13.01.01	5.4	31.74	25.84	1.4	1.4	1.2	1.0	4
	6123	2/2	4.7	TC73660	Wal17 protein (Fragment)	Q43661	18.02.01.02.03	5.4	31.66	n.d.	11.3	12.0	9.3	1.0	5
32 CELL RESCUE, DEFENSE AND VIRULENCE	5821	2/2	2.1	TC73262	Cell-autonomous heat shock coonate protein 70	Q8GSN2	32.07.07.03	4.9	55.47	71.26	28.9	29.0	5.4	1.0	6
	6110	27/15	20.1	TC64985	Non-cell-autonomous heat shock coonate protein 71	Q8SP17	32.07.07.05	5.8	32.53	71.44	28.3	34.3	4.8	1.0	6
18.02 target of regulation	6165	2/2	7.4	TC66579	Small heat shock protein, chloroplast precursor	P31170	32.01.05	6.0	27.65	23.94	16.7	17.6	17.3	1.0	5
	6166	3/3	9.5	TC66579	Small heat shock protein, chloroplast (precursor)	P31170	32.01.05	6.2	27.26	23.94	13.1	12.2	14.0	1.0	5
32.01 stress response	5852	10/10	9.1	TC58306	Heat shock protein 81-2 (HSP81-2)	P55737	32.01.05	4.9	54.00	80.06	9.8	13.2	1.1	1.0	6
	5867	5/4	4.9	TC58306	Heat shock protein 81-2 (HSP81-2)	P55737	32.01.05	5.1	53.40	80.06	11.3	10.6	1.2	1.0	6
32.07 detoxification	6216	4/4	9.0	TC57686	PR10 protein	Q7X9V7	32.05.03	6.1	19.44	17.79	25.8	21.7	23.7	1.0	5
	6147	11/5	12.0	TC58457	Tau class glutathione S-transferase	Q6DNM8	32.07.07.03	5.6	29.70	26.37	1.8	2.4	1.3	1.0	4
36 INTERACTION WITH THE ENVIRONMENT (Systemic)	5903	7/6	15.2	TC58524	L-aminocyclopropane-1-carboxylate oxidase	Q9M874	36.20.16.02	5.2	50.09	36.17	2.6	2.3	1.4	1.0	4
		2/2	7.4	TC67647	SAL1 phosphatase	Q42546	5.2	50.09	37.96						
42 BIOGENESIS OF CELLULAR COMPONENTS	5783	10/10	18.5	TC74839	1-deoxy-D-xylulose-5-phosphate reductoisomerase	Q6GZ43	01.06.01.07	5.8	58.13	51.78	13.4	14.1	15.0	1.0	5
		2/2	7.4	TC67647	SAL1 phosphatase	Q42546	5.2	50.09	37.96						



6625	10/7	26.7	60.5	TC57615	Low molecular mass heat shock protein Oshsp17.3	P93439	32,01.05	Heat shock response	GO:0009408	response to heat	17.57	1.0	1.0	1.0	8.6	1
	14/14	19.8	39.5	TC6522	Stress-induced protein sti1-like protein	Q8STH1					63.71					
6201	7/4	11.2	100	TC73408	Low molecular weight heat-shock protein	Q40936	32,01.05	Heat shock response	GO:0009408	response to heat	18.20	1.0	1.0	1.5	1.0	4
6180	7/3	13.5	100	TC58073	Mitochondria-localized low molecular weight heat sh	Q40847	32,01.05	Heat shock response	GO:0009408	response to heat	23.44	1.0	1.5	1.4	1.8	4
<b>42 BIOGENESIS OF CELLULAR COMPONENTS</b>																
42,01 cell wall																
6592	9/9	24.9	94.4	TC70902	UDP-sugar pyrophosphorylase	Q5W015	42.01	cell wall	GO:0007047	cell wall organization and biocer	66.18	1.0	1.0	8.3	6.5	2
	2/2	2.8	5.6	TC66322	Stress-induced protein sti1-like protein	Q8STH1					63.71					
<b>99 UNCLASSIFIED PROTEINS</b>																
6628	3/3	10.0	100	TC58411	EXPRESSED PROTEIN	Q8F391	99	UNCLASSIFIED PROTEINS	ND		21.04	1.0	1.0	1.0	5.7	1
6566	6/6	23.5	32.59	TC79033	root hair defective 3 GTP-binding protein	Q9FKK9	99	UNCLASSIFIED PROTEINS	ND		79.50	1.0	1.0	6.8	6.7	2
	6/5	7.8	67.41	TC67122	GTP-binding protein-like;root hair defective 3 protein-Hk	Q9FKK9					93.57					

(a) scans: number of matching peptides/number of unique peptides identified by SEQUEST

(b) coverage: refers to the *Pirius* TC

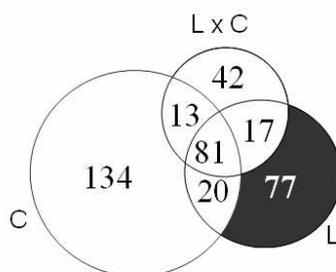
(c) relative quantity: calculated using sequest

(d) *ex*o Mw (kDa) and *pI* of the protein were estimated in comparison to a 2D gel with marker proteins

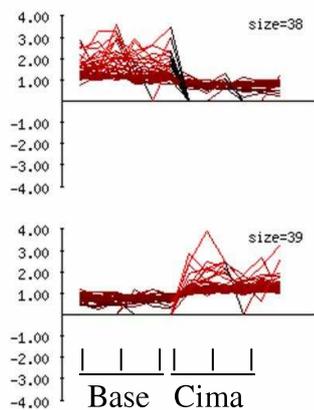
(e) Theoretic Mw (kDa) and *pI* of the homologous protein were calculated with a tool available at URL: <http://us.expasy.org/tools/peptide-mass.html>

#### 4.1.6 Mapas proteicos bidimensionales de tejido formador de madera Gradiente base a copa 2006.

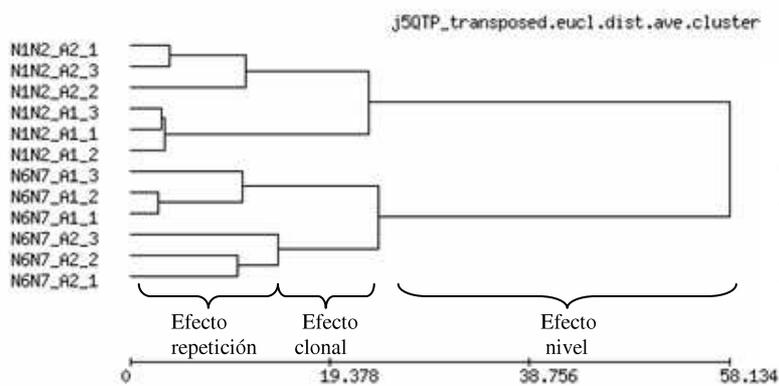
En el 2006, se utilizaron dos árboles rectos de genotipo desconocido como replicados biológicos: Árbol 1 y Árbol 2. Utilizando muestras proteicas extraídas desde la base (L1L2) o la copa (L6L7) de estos árboles, se analizó el efecto nivel, el efecto genotipo y su interacción. Se localizaron 944 spots en un mapa bidimensional de referencia, este fue analizado por ANOVA de dos factores, lo que llevó a identificar 384 proteínas expresadas diferencialmente en forma significativa ( $p < 0,001$ ), debido a por lo menos a un efecto. De estas 384 proteínas, se identificaron 77 cuya expresión diferencial se explica solamente por el efecto nivel (**Figura 30**), 42 proteínas son significativas sólo para el efecto interacción y 134 son significativas solo para el efecto clon/genotipo. De las 77 proteínas significativas sólo para el efecto nivel, 39 y 38 fueron sobre expresadas en madera de copa/juvenil o madera base/madura euclidiana de estas 77 muestras, (**Figura 30**) los niveles respectivamente (**Figura 29**). En el análisis de agrupamiento transverso basado en distancia L1L2 o L6L7 agruparon juntos, independientemente de su distinto genotipo, Arbol 1 (A1) o Arbol 2 (A2), confirmando que para estos spots el efecto nivel fue el más importante. En cada uno de los dos sub-clusters, el efecto clonal fue similar. Un total de 65 spots sobre expresados fueron seleccionados para análisis por espectrometría de masas ESI MS/MS, 33 desde madera base/madura y 32 de madera de copa/juvenil (**Figura 31**). Las proteínas identificadas en la **Tabla 6** se encuentran agrupadas por categorías funcionales siguiendo el catalogo MIPS (Munich Information Center for Protein Sequences, <http://mips.gsf.de/projects/funcat>).



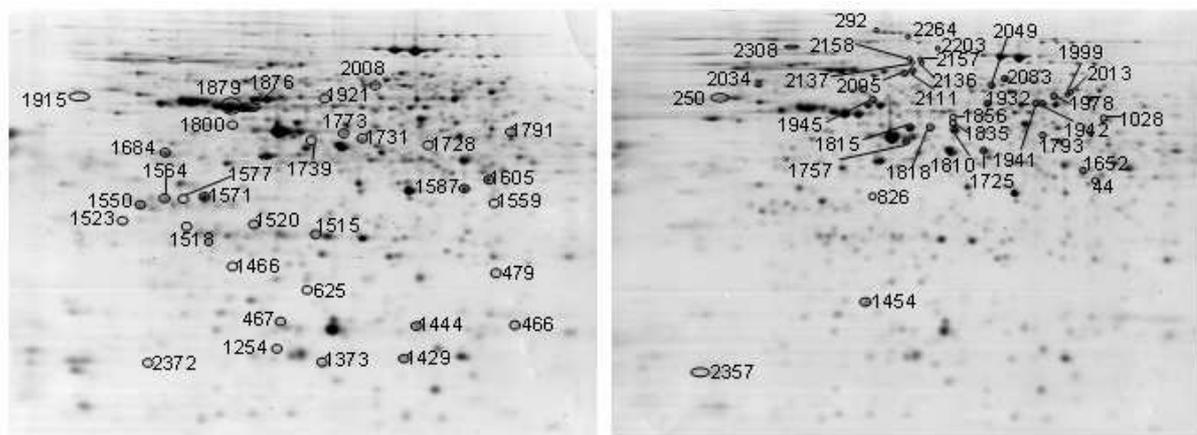
**Figura 30:** Diagrama de Venn, 384 spots significativos ( $p < 0,001$ ). C: clonal, L: nivel, LxC: interacción, 77 spots fueron significativas solo para el efecto Nivel.



**Figura 31:** Análisis clustering de la expresión de 77 proteínas, 38 y 39 fueron sobreexpresadas en madera de base/madura y de cima/juvenil, respectivamente.

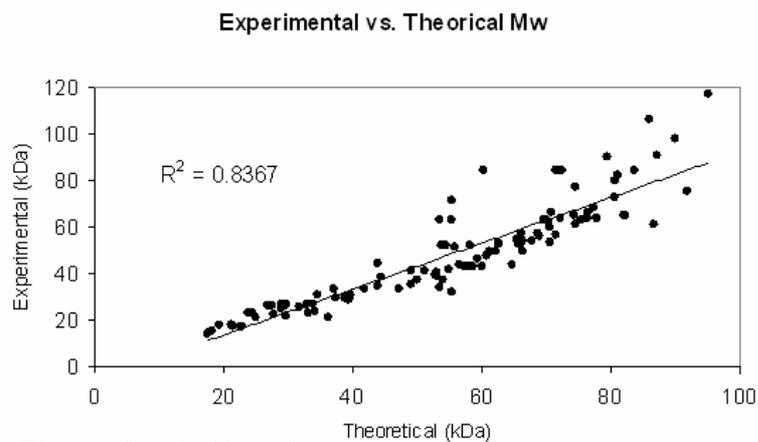


**Figura 32:** Clustering de las 77 muestras por su distancia Euclideana confirma que el efecto base a copa es el doble del efecto clonal.



**Figura 33:** Electroforesis bidimensional para tejido formador de madera en pino marítimo 2006. Proteínas sobreexpresadas en madera de base/madura se muestran en el panel izquierdo, en madera de copa /juvenil en el panel derecho. Las proteínas identificadas fueron rotuladas según la **Tabla 6**.

Una manera de verificar la correcta asignación de identidades a las proteínas identificadas, es comparar su peso molecular experimental, observado en los geles bi-dimensionales con el peso molecular teórico de la proteína asignada. Se encontró una buena correlación entre el peso molecular experimental de las proteínas en el gel vs. peso molecular teórico,  $r^2 = 0.84$ , se puede observar que la correlación es óptima hasta aproximadamente los 80 kDa, a mayor peso molecular la relación linear no continua, por efecto del porcentaje de acrilamida utilizado (**Figura 33**). Siete fragmentos proteicos no fueron incluidos en el gráfico por no tener peso molecular teórico y 5 “outliers” fueron removidos. Debido a que la información de secuencia en pino marítimo es principalmente por bases de datos de ESTs, los valores teóricos de peso molecular y pI corresponden a otras especies vegetales. Esto podría explicar la falta de correlación para algunos valores observados de ambos pesos moleculares.



**Figura 34:** Gráfico de peso molecular experimental vs. Teórico para experimentos de edad cambial 2005 y 2005

Table 2: 2006 Differentially expressed proteins between BW and CW forming tissue for the 2006 samples

spot ID	scans (a)	relative quantity % (b)	TC(c)	Assignment	UniProtKB Entry	pI (d)		MW (kDa)(e)		FOLD CHANGE (f)				ONTOLOGY			
						experimental	theoric	experimental	theoric	T1	T2	T1	T2	MIPS	GO		
<b>01 METABOLISM</b>																	
<b>01.01 amino acid metabolism</b>																	
2008	48/15	71.1	TC57104	<b>Phospholiverate dehydrogenase-like</b>	O8LGJ6	5.82	6.32	70015	63310	1.5	1.8	1.3	1.0	01.01.09.02.01	biosynthesis of serine	GO:0006564	L-serine biosynthesis
1559	13/3	29.9	TC57561	Exosome complex exonuclease, putative	A2HU12												
<b>01.02 nitrogen and sulfur metabolism</b>																	
1915	9/8	47.7	TC68943	<b>Hypothetical protein</b>	A3CHM6	6.22	5.73	39847	29831	4.6	3.9	1.0	1.0	1.02	nitrogen and sulfur metabolism	GO:0006607	GO:nitrogen compound meta
1915	12/6	30.6	TC67960	Similarity to alpha galactosidase	O9LIN8												
<b>01.03 nucleotide metabolism</b>																	
1520	7/3	39.5	TC58800	<b>Putative ABA-responsive protein</b>	O690B4	4.50	5.37	64659	43777	10.4	9.7	1.0	1.0	01.03.16.01	RNA degradation	GO:0006401	RNA catabolism
1518	22/10	31.5	TC66845	14-3-3 protein	O9XEW8		4.78		29495								
1684	8/6	20.3	TC66842	Putative DNA repair protein RAD23-2	O84L32		4.4		39842								
1739	4/2	8.8	TC57949	Carboxyl-terminal proteinase-like	O6H7H9		4.32		24839								
<b>01.05 C-compound and carbohydrate metabolism</b>																	
1520	29/4	100.0	TC74528	<b>Putative D-ribulose-5-phosphate 3epir</b>	O9AKI3	5.28	5.15	34097	24042	1.7	1.6	1.0	1.2	1.05	C-compound and carbohydrate m	GO:0005975	carbohydrate metabolism
1518	9/5	100.0	TC66413	<b>Putative ribose-5-phosphate isomeras</b>	O6ZEZ1	4.96	4.91	38897	26973	6.9	7.5	1.0	1.0	1.05	C-compound and carbohydrate m	GO:0005975	carbohydrate metabolism
1587	49/12	100.0	TC66124	<b>Fructokinase, putative</b>	O9C524	4.86	5.54	49084	41471	1.5	1.6	1.0	1.0	1.05	C-compound and carbohydrate m	GO:0005975	carbohydrate metabolism
<b>01.20 secondary metabolism</b>																	
1739	12/9	50.4	TC73223	<b>Flavanone 3 beta-hydroxylase</b>	O819C8	5.53	5.32	52959	40357	7.5	11.8	1.0	1.0	01.20.35.01.05	iosynthesis of stilbenes, flavonoi	GO:0009813	flavonoid biosynthesis
625	4/4	16.7	TC65368	Ethylene-forming enzyme (EFE)	O49839												
1477	5/5	10.7	TC67101	Pyruvate dehydrogenase	O24457												
1605	5/5	8.5	TC73216	Flavanone 3-hydroxylase	O2L6K2												
1587	4/4	8.2	TC58314	Plastidic glutamine synthetase	O75QN4												
1587	4/4	5.5	TC59783	Aspartate aminotransferase	O5F4K8												
1605	43/15	100.0	TC69396	<b>Phenylalanine hydroxylase</b>	O9M520	6.22	6.62	43956	34814	3.2	2.8	1.7	1.0	01.20.35.01.05	biosynthesis of stilbenes, flavonc	GO:0009813	flavonoid biosynthesis
1587	21/5	88.8	TC72883	<b>Phenylcoumaran benzylic ether reduct</b>	O81651	6.22	5.76	41853	33565	1.5	1.8	1.1	1.0	01.20.35.01.05	biosynthesis of stilbenes, flavonc	GO:0009813	flavonoid biosynthesis
1523	8/7	11.2	TC58713	26S proteasome non-ATPase regulatory	O8L0T8												
<b>02 ENERGY</b>																	
<b>02.01 glycolysis and gluconeogenesis</b>																	
625	6/4	100.0	TC67112	<b>Glyoxalase I</b>	O49818	5.53	5.26	24969	21088	5.6	3.1	1.0	1.0	02.01.01	lycolysis methylglyoxal bypass	GO:0009438	methylglyoxal metabolism
<b>02.30 photosynthesis</b>																	
<b>14 PROTEIN FATE</b>																	
<b>14.07 protein modification</b>																	
1921	19/17	42.6	TC65932	<b>Putative quinone oxidoreductase</b>	O8L5Q7	5.89	6.52	29579	21722	1.4	1.4	1.1	1.0	02.30.07	accessory proteins of photosynthid		
1564	8/8	18.6	TC57671	<b>Putative leucine aminopeptidase</b>	O6K689	5.58	5.5	65398	54653	1.9	1.6	1.0	1.5	14.07.11	protein processing (proteolytic)	GO:0016485	protein processing
1550	2/2	6.6	TC69393	Alpha-tubulin	O70ZL7												
1550	2/2	15.7	TC57318	ATP synthase subunit beta, mitochondria	P85088												
1523	5/4	13.8	TC75906	UDP-glucose 4-epimerase	O2V6K0												
1523	7/6	9.3	TC58505	UTP--glucose-1-phosphate uridylyltransf	O8LKG7												
1466	13/3	100.0	TC74819	<b>26S proteasome subunit</b>	O6XJF4	4.68	4.81	34534	30702	6.1	5.5	1.0	1.0	14.13.01.01	proteasomal degradation (ubiquitin	GO:0045161	proteasomal ubiquitin-dependent
1564	8/8	15.6	TC75417	Phosphoserine phosphatase	O8Z796												
<b>16 PROTEIN WITH BINDING FUNCTION OR COFACTOR REQUIREMENT (structural or catalytic)</b>																	
<b>16.01 protein binding</b>																	
1564	32/5	93.4	TC72878	<b>14-3-3 protein</b>	O9XEW8	4.87	4.78	39961	29495	1.4	1.4	1.0	1.2	16.01	protein binding	GO:0005515	protein binding
1550	2/2	6.6	TC58929	L-galactose-1-phosphate phosphatase	O5U789												
1550	20/7	81.6	TC66140	<b>14-3-3 protein</b>	O9XEW8	4.75	4.78	37499	29495	1.3	1.2	1.0	1.1	16.01	protein binding	GO:0005515	protein binding
1373	8/5	13.1	TC66845	Putative 14-3-3 protein	O6XPJ0												
1373	11/3	5.3	TC57949	High molecular weight glutenin	O8I284												
<b>20 CELLULAR TRANSPORT, TRANSPORT FACILITATION AND TRANSPORT ROUTES</b>																	
1466	13/3	100.0	TC65017	<b>RAB1-like</b>	O68V17	5.20	5.5	27730	22641	5.6	2.5	1.0	1.0	20.09.07.03	ER to Golgi transport	GO:0006888	ER to Golgi transport
<b>30 CELLULAR COMMUNICATION/SIGNAL TRANSDUCTION MECHANISM</b>																	
1373	11/4	90.5	TC57404	<b>14 kDa zinc-binding protein (Protein kii</b>	P42856	5.80	6.19	17586	14301	1.7	1.3	1.2	1.0	30.01.05.01	protein kinase cascades	GO:0007243	protein kinase cascade
1373	2/2	9.5	TC58791	Putative intracellular pathogenesis-relate	O9SNX7												

**32 CELL RESCUE, DEFENSE AND VIRULENCE**

32.01.05 heat shock response	TC66579	Heat shock protein 28	O41815	6.22	7.96	27384	26378	226	9.2	1.0	32.01.05	Heat shock response	GO:0009408	response to heat
	TC73490	Low molecular weight heat-shock prot	O40936	6.22	6.19	21248	18203	20.2	13.1	12.2	32.01.05	Heat shock response	GO:0009408	response to heat
	TC57297	Cytosolic class I small heat shock prot	O55E47	5.39	no data	18843	no data	15.2	16.6	12.0	32.01.05	Heat shock response	GO:0009408	response to heat
	TC57030	unknown	no hit											
	TC73490	Low molecular weight heat-shock prot	O40936	6.03	6.19	21210	18203	2.8	1.8	1.3	32.01.05	Heat shock response	GO:0009408	response to heat
	TC65958	Alcohol dehydrogenase	O43027	5.54	8.72	33206	23537	5.6	3.7	3.1	32.01.05	Heat shock response	GO:0009408	response to heat
	TC77713	Small heat shock protein chloroplast p	P31170											
	TC65955	Ascorbate peroxidase	O6RY68											
	TC79231	Triosephosphate isomerase chloroplast p	O9M4S8											
	TC98721	Small heat shock protein, chloroplast	P30222	5.41	no data	21497	no data	2.6	1.8	1.0	32.01.05	Heat shock response	GO:0009408	response to heat
	TC73213	Puative thioredoxin-dependent peroxidat	A1ECJ6											
<b>32.05 disease, virulence and defense</b>														
1800	TC74273	Gene induced upon wounding stress	O39171	5.16	5.21	56627	43874	14.7	8.5	5.0	32.05.03	Defense related proteins	ND	
	TC58452	Transaldolase	O04894											
	TC73522	Actin	O6S9W9											
<b>32.07 detoxification</b>														
1429	TC66210	Superoxide dismutase [Cu-Zn]	P24689	5.96	5.75	18147	15376	1.8	1.1	1.0	32.07.07.07	superoxide metabolism	GO:0006801	superoxide metabolism
1791	TC65134	Monodehydroascorbate reductase	O93YG1	6.22	6.38	55881	51748	1.5	1.4	1.0	32.07.07	Oxveen and radical detoxification	ND	
<b>36 INTERACTION WITH THE ENVIRONMENT (Systemic)</b>														
36.20 plant / fungal specific systemic sensin and response														
			phosphatase	Q42546				5.2	5.0	50.09	37.56			
<b>42 BIOGENESIS OF CELLULAR COMPONENTS</b>														
<b>42.01 cell wall</b>														
1773	TC66403	Caffeic acid ortho-methyltransferase	O9SYU3	5.67	5.6	54884	41943	2.0	2.1	1.5	42.01	cell wall	GO:0007047	cell wall organization and bio
1573	TC60933	Caffeoyl-CoA O-methyltransferase	O9ZT15	4.94	5.44	39461	29146	8.8	11.3	1.1	42.01	cell wall	GO:0007047	cell wall organization and bio
	TC73846	Puative ankyrin (A12a03430)	O6GZ22											
	TC72878	14-3-3 protein	O9LKK9											
1571	TC73846	Caffeoyl-CoA O-methyltransferase	O9ZT15	5.04	5.44	39230	29146	1.7	1.8	1.2	42.01	cell wall	GO:0007047	cell wall organization and bio
<b>42.04 cytoskeleton</b>														
1876	TC57031	Tubulin alpha-1 chain (Alpha-1 tubulin)	O6VAG1	5.21	4.93	61318	49738	1.4	1.4	1.1	42.04.05	microtubule cytoskeleton	GO:0000226	microtubule cytoskeleton org
1879	TC65872	Tubulin alpha-1 chain (Alpha-1 tubulin)	O6VAG1	5.15	4.93	61260	49738	1.4	1.3	1.1	42.04.05	microtubule cytoskeleton	GO:0000226	microtubule cytoskeleton org
<b>70 SUBCELLULAR LOCALIZATION</b>														
70.16 mitochondrion														
2372	TC66482	Cytochrome b5	P49097	4.84	4.88	17681	14745	17.3	13.5	1.0	70.16	mitochondrion	GO:0005739	mitochondrion
<b>99 UNCLASSIFIED PROTEINS</b>														
1751	TC65368	Hypothetical protein	O4C4O1	5.76	4.75	53508	34067	2.7	2.4	1.1	nd		nd	
	TC59783	Aspartate aminotransferase	O5F4K8											
	TC57059	Phosphotyrosine kinase, chloroplast [P	P41798											

**CROWN WOOD OVER-EXPRESSED PROTEINS**

<b>01 METABOLISM</b>															
<b>01.01 amino acid metabolism</b>															
1941	TC58434	Glutamate decarboxylase	A0EL89	6.08	5.74	66083	57230	1.0	1.0	1.4	01.01.03.02	metabolism of glutamate	allatamate metabolism		
	TC73623	Adenosylhomocysteinase 1	O23255												
2013	TC72904	Puative keto-acid reductoisomerase	O65XK0	6.24	6.01	70414	62377	1	1.2	2.3	01.01.1.1	metabolism of the pyruvate famli	GO:0009081	branched chain family amino	
	TC73142	Keto-acid reductoisomerase	O1M2Z5												
	TC74151	Mitochondrial processing peptidase beta	O9AXQ2												
	TC66524	Puative dehydroquinate shikimate dehydr	O9SQT8												
<b>01.05 C-compound and carbohydrate metabolism</b>															
1810	TC68204	S-adenosylmethionine synthetase	O9FVG7	5.88	5.55	57329	43069	1.0	1.1	1.8	1.7	1.05	C-compound and carbohydrate n	GO:0006730	one-carbon compound metat
1815	TC59198	S-adenosylmethionine synthetase	O9FVG7	5.46	5.55	57601	43069	1.0	1.0	1.2	1.2	1.05	C-compound and carbohydrate n	GO:0006730	one-carbon compound metat
1816	TC66027	S-adenosylmethionine synthetase	O9FVG7	5.56	5.55	57930	43069	1.0	1.0	1.4	1.4	1.05	C-compound and carbohydrate n	GO:0006730	one-carbon compound metat
	TC65872	Tubulin alpha-3 chain	P22275												
<b>02 ENERGY</b>															
<b>02.01 alcohols and alliconeogenesis</b>															
2083	TC58658	2,3-bisphosphoglycerate-independent	O42908	5.92	5.39	74532	61183	1.1	1.0	1.3	1.4	2.01	alcoholysis and alliconeogenesis	GO:0006096	alcoholysis
2203	TC58658	2,5-bisphosphoglycerate-independent	O42908	5.59	5.39	86731	61183	1.0	1.5	3.1	2.6	2.01	alcoholysis and alliconeogenesis	GO:0006096	alcoholysis
	TC74770	Hypothetical protein	AZT127												
2095	TC58983	Phosphoglucosylase	O9AUC4	5.29	5.4	75384	62949	1.0	1.0	1.5	1.7	2.01	alcoholysis and alliconeogenesis	GO:0006096	alcoholysis
2049	TC72876	Pyruvate decarboxylase 2	O5BN15	5.86	5.86	72101	63902	1.5	1.0	1.6	1.7	2.01	alcoholysis and alliconeogenesis	GO:0006094	alliconeogenesis

1652	38/10	100.0	TC74340	Malate dehydrogenase	P17783	6.22	6.26	47264	33235	1.0	1.0	1.9	1.8	2.10	tricarboxylic acid pathway (citrate)	GO:0006099	tricarboxylic acid cycle
02.13 respiration																	
1793	37/15	86.6	TC65967	Alcohol dehydrogenase	O43025	6.11	no data	56147	no data	1.0	1.0	1.4	1.7	02.13.01	anaerobic respiration	GO:0009061	anaerobic respiration
11 TRANSCRIPTION	15/15	13.4	TC65134	Monodehydroascorbate reductase	O83YG1												
11.02 RNA synthesis																	
1635	10/9	100.0	TC68732	Eukaryotic initiation factor 4A	O41741	5.67	5.38	59282	46536	1.0	1.1	1.4	1.6	11.02.03.01.01	transcription initiation	GO:0006952	transcription initiation
2264	8/7	57.1	TC75960	Hypothetical protein	A3C316	5.44	8.04	91891	75709	1.0	4.8	8.0	6.2	11.02.02	RNA synthesis	GO:0009304	RNA transcription
76	42.9		TC58983	Phosphodiesterase	O9AUO4												
14 PROTEIN FATE																	
14.01 protein folding and stabilization																	
2034	29/19	51.8	TC57696	Protein disulfide isomerase	O5EUE0	4.73	4.91	71489	56683	1.0	1.9	2.5	3.3	14.01	protein folding and stabilization	GO:0050821	protein stabilization
15/10	38.9		TC66264	UDP-glucose dehydrogenase	O6RK07												
17/10	7.4		TC65783	Methylenetetrahydrofolate reductase 1	O9SE94	5.44	5.29	54012	37324	1.0	1.2	1.6	1.8	14.01	protein folding and stabilization	GO:0050821	protein stabilization
1757	26/9	89.2	TC58578	Probable protein disulfide-isomerase 4	P38661												
12/9	10.5		TC68668	3-Isopropylmalate dehydrogenase 2, chloro	P93832	4.57	4.6	66397	49628	1.0	1.0	4.2	2.1	14.01	protein folding and stabilization	GO:0006457	protein folding
250	38/15	79.1	TC74181	Calreticulin	O9FYV2												
18/8	20.9		TC67305	Nucleosome/chromatin assembly factor c	O8L8G5												
14.07 protein modification																	
1965	16/7	93.4	TC66027	OSJNBa0085B14.13 protein (Hypohep1)	O7XV42	5.55	5.47	67860	54069	1.0	1.4	2.5	2.8	14.07.02	modification with sugar residues	GO:0006486	protein amino acid glycosylation
5/4	6.6		TC66314	Putative glucosyltransferase	O9SKC1												
292	3/3	35.6	TC59106	Ubiquitin-activating enzyme E1 3	O10MU6	5.30	5.35	95079	117453	1.0	6.8	12.0	8.8	14.07.05	ubiquitination, deubiquitination	GO:0016567	protein ubiquitination
2/2	39.0		TC65134	Monodehydroascorbate reductase	O93YG1	6.38	6.38	51748	51748								
2/2	25.4		TC58056	Putative heat shock -protein	O9S7C0												
1999	20/11	90.7	TC74151	Putative mitochondrial processing pep	OQWWT6	6.23	no data	66684	no data	1.0	1.0	1.4	1.4	14.07.11	protein processing (proteolysis)	GO:0016485	GO:protein processing
7/7	9.3		TC66524	Putative dehydroquinase shikimate dehy	O9SQ18												
16 PROTEIN WITH BINDING FUNCTION OR COFACTOR REQUIREMENT (structural or catalytic)																	
16.13.03 fatty acid binding																	
2137	11/9	84.3	TC58038	F22F7.13 protein	O9MA55	5.46	5.16	80706	73075	1.5	1.0	2.0	2.5	16.13.03	fatty acid binding (e.g. acyl-carrier)	GO:0005504	fatty acid binding
3/3	15.7		TC73142	Putative keto-acid reductoisomerase	O65XK0	5.51	5.08	82094	65246	1.0	1.0	1.7	2.1	16.13.03	fatty acid binding (e.g. acyl-carrier)	GO:0005504	fatty acid binding
2157	35/18	100.0	TC61078	Hypothetical protein	A2XNR7	5.45	5.08	82249	65246	1.1	1.0	1.4	1.7	16.13.03	fatty acid binding (e.g. acyl-carrier)	GO:0005504	fatty acid binding
2158	19/14	93.2	TC61078	Hypothetical protein	A2XNR7	5.45	5.08	82249	65246	1.1	1.0	1.4	1.7	16.13.03	fatty acid binding (e.g. acyl-carrier)	GO:0005504	fatty acid binding
2/2	6.8		TC58434	Glutamate decarboxylase	A0LE89												
18 PROTEIN ACTIVITY REGULATION																	
18.01 mechanism of regulation																	
2136	6/6	73.0	TC58038	Homology to unknown gene	OQWV34	5.52	no data	80629	no data	1.0	1.0	1.6	2.2	18.01.07	binding / dissociation	GO:0005515	protein binding
3/3	27.0		TC80063	Hypothetical protein	A3B9M8												
20 CELLULAR TRANSPORT, TRANSPORT FACILITATION AND TRANSPORT ROUTES																	
20.08 transport routes																	
2111	126/30	100.0	TC74448	Vacuolar ATP synthase catalytic subunit	P13548	5.47	5.3	77337	68550	1.2	1.0	1.4	1.5	20.09.07	vesicular transport	GO:0016192	vesicle-mediated transport
1945	101/21	100.0	TC65063	Vacuolar ATP synthase subunit B	P11574	5.29	4.98	66271	54108	1.0	1.0	1.2	1.2	20.09.07	vesicular transport	GO:0016192	vesicle-mediated transport
32 CELL RESCUE, DEFENSE AND VIRULENCE																	
32.01.05 heat shock response																	
2208	21/15	100.0	TC66390	Putative heat-shock protein	O9S7E7	4.88	5.26	87060	90956	1.0	1.1	2.0	1.8	32.01.05	Heat shock response	GO:0009408	response to heat
1454	28/6	100.0	TC58057	Small heat-shock protein	O8L5D7	5.26	8.74	23768	23421	1.4	1.0	1.6	1.5	32.01.05	Heat shock response	GO:0009408	response to heat
42 BIOGENESIS OF CELLULAR COMPONENTS																	
42.01 cell wall																	
1932	90/21	100.0	TC66264	UDP-glucose dehydrogenase	O6RK07	5.84	5.99	65709	52933	1.0	1.0	1.4	1.6	42.01	cell wall	GO:0007047	cell wall organization and biogenesis
1942	67/21	93.6	TC66264	UDP-glucose dehydrogenase	O6RK07	6.11	5.99	66093	52933	1.2	1.0	1.7	1.7	42.01	cell wall	GO:0007047	cell wall organization and biogenesis
10/9	10.9	6.4	TC73526	Adenosylhomocysteinase	P68172												
1978	23/9	69.2	TC64990	Inositol-3-phosphate synthase	O64437	6.16	5.44	68831	56264	1.8	1.0	2.7	2.0	42.01	cell wall	GO:0007047	cell wall organization and biogenesis
22/10	19.2		TC75286	Haloacid dehalogenase-like hydrolase family	O2R483												
13/11	11.6		TC57058	Aldehyde dehydrogenase 1	O8AV96												
42.04 cytoskeleton																	
1028	30/13	75.2	TC65872	Alpha-tubulin-like protein	O2XPX0	6.22	4.85	62137	49871	4.2	1.0	8.4	9.0	42.04.05	microtubule cytoskeleton	GO:0000226	microtubule cytoskeleton organization
23/16	24.8		TC58922	26S proteasome ATPase 2 subunit	O550V2	5.68	5.43	60855	47860	1.0	1.3	1.5	1.7	42.04.03	actin cytoskeleton	GO:0030036	actin cytoskeleton organization
1856	8/7	34.4	TC78099	F3F19.20 protein (Actin-related protein 9)	O9SAF1												
8/8	25.7		TC75195	3-Hydroxy-3-methylglutaryl-CoA-synthase	P93773												
7/7	23.5		TC74803	Dihydroisoamide acetyltransferase	O6RFW9												
5/5	16.4		TC74840	Malate dehydrogenase	P46487												

(a) scans: number of matching peptides/number of unique peptides identified by SEQUEST  
(b) relative quantity - calculated using seqest software  
(c) Tentative Contig  
(d) exp Mw (kDa) and pI of the protein were estimated with LMW calibration kit for SDS Electrophoresis (Amersham Biosciences)  
(e) Theoretic Mw (kDa) and pI of the homologous protein were calculated with a tool available at URL: <http://us.expasy.org/tools/peptide-mass.html>  
(f) Fold change between the 2 levels: base or crown (L1L2, L6L7) for two trees (T1, T2)

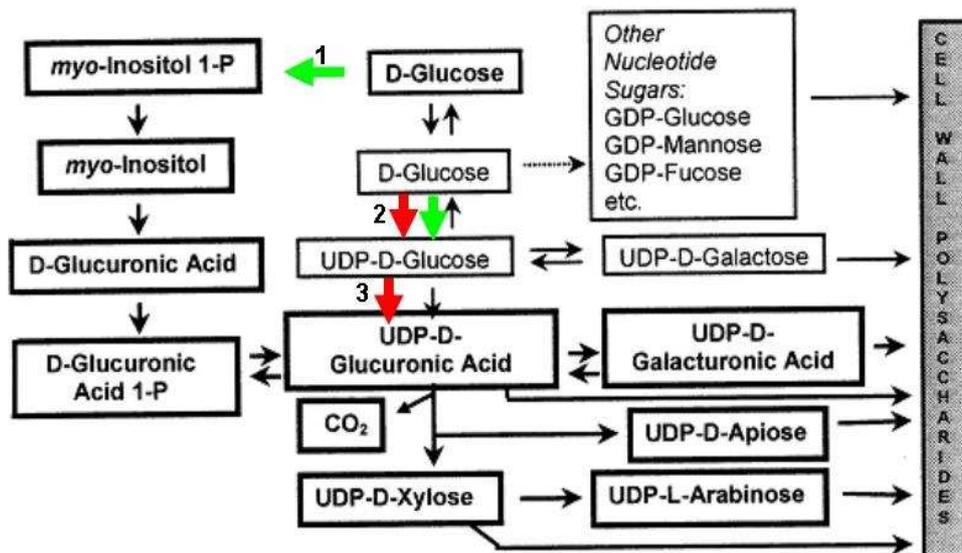
#### 4.1.8 Expresión de genes y proteínas a lo largo del gradiente base a copa

##### Genes y Proteínas relacionados con síntesis de pared celular.

Cato et al. (2006), reportó que sobre un 30% de traqueidas en diferenciación de madera de base produjo paredes secundarias, comparado con solo un 3% en madera de copa. Ya que las paredes secundarias se componen principalmente de celulosa, el mayor grosor de la pared secundaria en la base del árbol se correlaciona con el mayor contenido de celulosa en la base. La caracterización anatómica de muestras contrastantes de base o copa de la temporada 2006 (**Figura 26**), mostró las características típicas de madera de base y madera de copa. El contenido de celulosa a lo largo de un gradiente base a copa en tejido de xilema en formación, fue similar a la variación encontrada en madera totalmente desarrollada, es decir el contenido de celulosa fue mayor en la base del árbol.

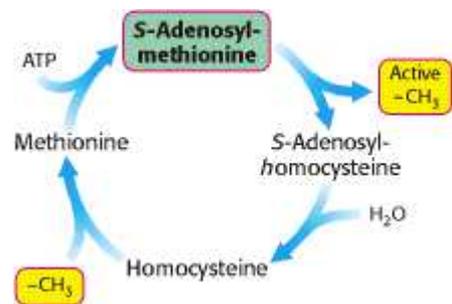
En el análisis transcriptómico de la temporada 2003, se identificaron dos celulosa sintasa asociadas con madera de base: PpinCesa3 (1.8 v.c. (veces de cambio)) y PpinCesa1 (2.8 v.c.). El gen PpinCesa3 es 98% idéntico a PtCESA3, ortólogo de AtCESA7(IRX3) y PpinCesa1 es 97% idéntico a PtCESA2 ortólogo de AtCESA8. Ambos genes de *Arabidopsis thaliana* son conocidos por estar involucrados en la formación de xilema secundario.

Proteínas relacionadas al metabolismo de polisacáridos de la pared celular fueron sobreexpresadas en madera de base (**Figura 35**). En el 2003, se identificó una UDP-glucosa-pirofosforilasa (**spot#225, 5,1 v.c.**) (**2 en Figura 35**) y una UDP-glucosa deshidrogenasa (**spot#276, 75 v.c.**) (**3 en Figura 35**). Ambas Enzimas cumplen un importante rol en la formación de paredes celulares de plantas superiores. UDP-glucosa-pirofosforilasa produce UDP-glucosa, que puede ser utilizada en síntesis de celulosa (revisado por Kleczkowski et al., 2004), o puede ser utilizada por la UDP-glucosa deshidrogenasa para formar UDP-glucuronato. La UDP-glucosa deshidrogenasa es una enzima clave en la regulación el flujo de UDP-glucosa y otros azúcares nucleotídicos hacia la formación de pectina y hemicelulosas en el aparato de Golgi (Seitz et al., 2000). En la temporada 2005 se identificó 1 proteína relacionada con pared celular sobreexpresada en madera de base, una  $\alpha$ -L-arabinofuranosidasa/ $\beta$ -D-Xilosidasa (**#5491: 6,0 v.c.**), que hidroliza unidades de xilopiranosas, galactopiranosas y arabinopiranosas, probablemente involucrada en la remodelación de la pared celular (Lee et al., 2003).



**Figura 35 :** Biogénesis de polisacáridos de la pared celular vía el metabolismo de oxidación de mio-inositol ( en negrita) y por la oxidación de nucleótidos de azúcares. 1: mio-inositol-1-fosfato sintasa , 2:UDP-D-glucosa pirofosforilasa 3:UDP-glucosa deshidrogenasa. En rojo proteínas sobreexpresadas en madera de base y en verde madera de copa. Modificado de Loewus et al, (2000)

A pesar de la mayor proporción de hemicelulosas determinada químicamente en la madera de copa, se identificó sólo 1 proteína involucrada en la síntesis de hemicelulosas sobreexpresada en madera de copa en la temporada 2005: UDP-azúcar pirofosforilasa (#6582: 8,3 v.c.) (2 en Figura 35), posteriormente en la temporada 2006 se identificó una Inositol-3-fosfato sintasa (#1978: 2,7 v.c.) (1 en Figura 35), enzima inicial del ciclo del mio-inositol, que puede aportar a la formación de ácido glucurónico para la formación de pared celular. (Loewus et al., 1962; Loewus and Murthy, 2000)



**Figura 36 :** Ciclo del metil activado

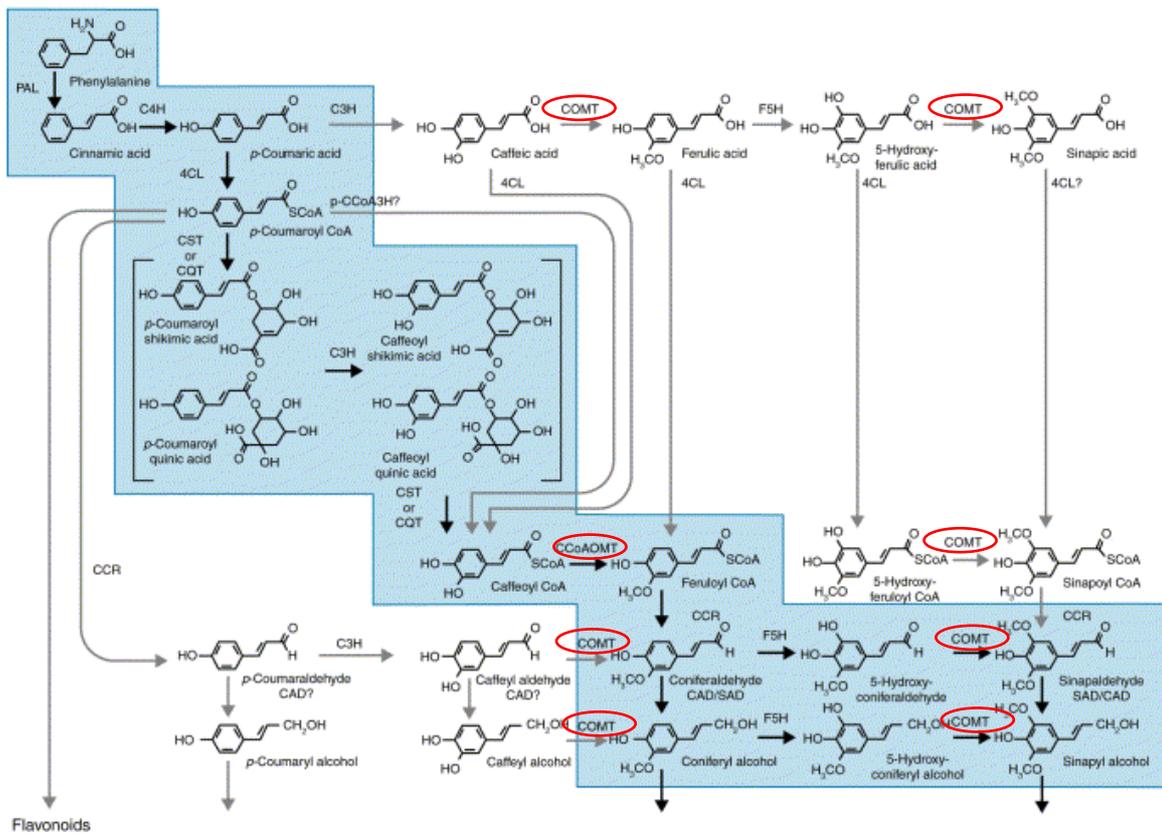
Las enzimas involucradas en el ciclo del metil activado (Figura 36), fueron altamente sobreexpresadas en madera base: por ej., el 2003, Metionina sintasa (#1057: 44 v.c.), Adenosilhomocisteinasa (#352: 6,4 v.c.), S-Adenosilmetionina sintetasa (SAM-S, #405: 5,4 v.c.) y metionina sintasa -Cobalamina

independiente (**#39: 7,1 v.c.**). En el 2005, 6 proteínas del ciclo del metil activado se identificaron en madera de base: 2 Adenosilhomocisteinasas (**#5703: 48,4 v.c.**, **#5707: 12,8 v.c.**), 3 Metionina sintasa independientes de Cobalamina (**#6483: 18,3 v.c.**, **#6484: 10,8 v.c.**, **#5517: 19,7 v.c.**), y 1 Metionina sintasa (**#6348: 21,8 v.c.**).

La abundancia de proteínas que participan en el ciclo de reacciones de transferencia del grupo metil activado en el tejido formador de madera de la base del árbol, refleja la alta demanda por reacciones de transferencia de metilo, requeridas para la formación de monolignoles en xilema. La importancia de la metilación, asociada con biosíntesis de lignina (**Figura 37**) es corroborada la temporada 2003 por la sobreexpresión de una enzima de lignificación Cafeoil-CoA O-metiltransferasa (CCoAOMT) (**#641, 2,5 v.c.**), en tejido formador de madera de base. En la temporada 2006 se identificaron 2 enzimas claves de la síntesis de lignina: ácido Cafeico orto-metiltransferasa (COMT) (**#177, v.c. 2.1**) y nuevamente CCoAOMT (**#1573, v.c. 11.3**). En madera de copa se identificaron proteínas del ciclo del metil activado, aunque en menor número y con una menor expresión de “veces cambio”, en comparación con madera de base, el 2003, se identificaron 2 S-Adenosilmetionina sintetasa (SAM-S, **#505, #483: 14,5; 3,4 v.c.**), el 2005 se identificaron 5 proteínas : 1 Metionina sintasa independiente de Cobalamina (**#6577: 59,5 v.c.**), 3 S-adenosilmetionina sintetasa (**#5749: 3,4 v.c.**, **#5753: 2,7 v.c.**, **#5775: 2,6 v.c.**) y 1 Adenosilhomocisteinasa (**#5645: 2,2 v.c.**).

A pesar que se ha descrito una mayor proporción de hemicelulosas y lignina en madera de copa, en términos de valor absoluto, se ha detectado a nivel proteómico una mayor expresión de proteínas de biosíntesis de componentes de pared celular en tejido de la base del árbol, por el mayor espesor de las paredes celulares de las traqueidas.

La sobreexpresión de estas proteínas involucradas en la síntesis de hemicelulosas, lignina y de genes relacionados con la síntesis de celulosa, aporta por primera vez evidencias proteómicas a la observación previa que caracteriza la madera madura como un tejido con una extendida fase de engrosamiento de la pared celular.



**Figura 37:** Metabolismo de precursores de la lignina (modificado de Humphreys and Chapple (2002))

**La prolongación del depósito de pared celular secundaria en tejido formador de madera de base, involucra la sobreexpresión de genes relacionados con defensa celular.**

Una de las características de las traqueidas de la base es la presencia de una gruesa pared celular, en comparación con aquellas de la copa (Zobel and Sprague, 1998). Esta pared celular más gruesa es generada por una extensión de la duración del depósito de pared celular más que por un aumento de la velocidad de síntesis de la pared (Larson et al., 2001; Ugglá et al., 2001; Cato et al., 2006).

En el análisis transcriptómico realizado para la temporada 2003, se encontraron 5 genes de respuesta a estrés, acumulados en tejidos formadores de madera de base, una proteína de estrés por calor (HSP) de 17,6 kDa (2,6 v.c.) , un precursor de gamma theonina/defensina (2,2 v.c., una LEA (late embriogenesis protein) (1,9 v.c.) un gen de regulación de daño a ADN (DDR48-stress protein) (1,5 v.c.) y una proteína homóloga a resistencia a

enfermedades (1,5 v.c.). Proteínas de embriogénesis tardía (LEA) son un grupo importante de proteínas extremadamente estables e hidrofílicas, se acumulan típicamente durante las últimas etapas de embriogénesis o en respuesta a hidratación, baja temperatura, salinidad tratamientos exógenos con ácido abscísico (ABA), indicando con esto su respuesta a la deshidratación (Ramanjulu and Bartels, 2002). Se ha propuesto que las LEA estabilizan membranas y previenen la cristalización de componentes celulares (Dure et al., 1989) (Garay-Arroyo et al., 2000). Tioninas son proteínas pequeñas, básicas, ricas en cisteína, que pueden funcionar como moléculas pequeñas en contra de diversos patógenos vegetales (Florack and Stiekema, 1994 ; Broekaert et al., 1995) . En este estudio se ha reportado por primera vez la sobreexpresión de estos dos genes en tejidos formadores de madera de la base.

En concordancia con los datos transcriptómicos, en la temporada 2003 encontramos 8 proteínas relacionadas con defensa y estrés sobreexpresadas en tejido formador de madera de base, 3 HSP (**#974, #1163, #995, veces de cambio 85, 18 y 10 , respectivamente**), 3 peroxidasas (**#1033, #1107 y #445, veces de cambio 26, 5,1 y 2,4 v, respectivamente**), una proteína universal de estrés (**spot #805; 13,0 v.c.**) y una chaperonina cnp60 (**spot #184; 2,7 v.c.**).

El 2005 y 2006, 17 proteínas correspondientes a esta categoría fueron sobreexpresadas en madera de base. 12 proteínas corresponden a la subcategoría respuesta a estrés: 2 HSP70 (**#6110: 34,3 v.c.; #5821: 29,9 v.c.**), 2 HSP90 (**#5867: 11,3 v.c.; #5852: 13,2 v.c.**) y 8 pequeñas HSP (**#479: 22,6 v.c.; #466: 20,2 v.c.; #6165: 17,6 v.c.; #1254: 15,2 v.c.; #6166: 13,1 v.c.; #1515: 5,6 v.c.; #1444: 2,8 v.c y #467: 2,6 v.c.**), 2 proteínas corresponden a la subcategoría enfermedad, virulencia y defensa: una PR10 homóloga con una ribonucleasa clase 2 (**#6216: 25,8 v.c.**) y una proteína inducida por estrés (**#1800: 14,7 v.c.**). 3 proteínas corresponden a la subcategoría detoxificación, 2 ascorbato peroxidasas (**#6158: 14,9 v.c. y #6146: 7.2 v.c.**) y una Glutathion-S transferasa del tipo Tau (**#6147: 2,4 v.c.**).

En plantas, HSP de bajo peso molecular, son acumuladas en respuesta a varios estrés (Vierling, 1991; Waters et al., 1996; Costa et al., 1998) y también parecen tener roles específicos en procesos de desarrollo, incluyendo maduración de semilla, embriogénesis somática y formación de madera (Puigderrajols et al., 1996; Pla et al., 1998; Le Provost et al., 2003). En animales, la expresión de HSP de bajo peso molecular, durante la transición división celular a diferenciación, ha sido relacionado con un rol preventivo en las células en diferenciación, para prevenir apoptosis (Arrigo, 2005). Gion (2005) observo también

que HSP de bajo peso molecular se acumularon en madera de base y sugirió que jugaban un rol en prolongar la fase de engrosamiento de la pared celular de la xilogénesis, atrasando la entrada en muerte programada celular (PCD).

En madera de copa se identificaron 2 proteínas de la categoría defensa el 2003, una HSP 22 (#1116: 9,0 v.c.) y una HSP 70 (#96: 2,8 v.c.), el 2006 se sobreexpresó solo una probable HSP(#2208: 2,0 v.c.). La temporada 2005 fue excepcional, se identificó la sobreexpresión de 8 proteínas de defensa e inducidas por estrés: 5 corresponden a proteínas inducidas por estrés (#6544: 112,9 v.c.; #6543: 25,8 v.c.; #6581: 19,2 v.c.; #6575: 10,5 v.c., #6576: 8,0 v.c.) y 3 a HSP de bajo peso molecular (#6519: 23,0 v.c., #6517: 13,5 v.c., #6625: 8,6 v.c.). El alto número de proteínas inducidas por estrés, se relaciona en esta temporada con la exposición a estrés hídrico del árbol entre las semanas 10 a 18, antes de la recolección de tejido de madera en formación. Este dato nos ilustra la importancia del factor ambiental en la expresión de proteínas y demuestra que el ambiente es capaz de producir una remodelación del proteoma, cambiando las características típicas del gradiente base a copa.

En árboles no sometidos a estrés, como en las temporadas 2003 y 2006, se identificó la sobreexpresión de proteínas de respuesta a estrés y de detoxificación en tejido formador de madera base (17 en madera base vs 3 en madera de copa). Estos resultados son coincidentes con los obtenidos previamente a nivel proteómico por Gion et al. (2005). Además esto es consistente con la hipótesis que estos genes están involucrados en prolongar el depósito de pared celular secundaria, evitando la muerte celular programada, resultando en el mayor espesor de la pared celular secundaria característica de madera base.

### **Genes relacionados con “síntesis de proteínas” y “energía” son sobreexpresados en madera formación de copa.**

El mayor contenido de proteínas observado en la copa del árbol puede ser resultado de una mayor tasa de división celular, con más células en las etapas de división y expansión, como se observó en observaciones microscópicas (**Figura 26**). Este resultado es consistente con lo reportado en *P. radiata*, donde la división celular fue 3,3 veces mayor en el cambium de la copa, comparado con el de la base, hasta en el periodo de formación de madera tardía (Cato et al., 2006). En pino silvestre (*P. sylvestris*), Ugglá et al. (2001), también reportó que

la tasa y el tiempo de división y expansión celular fueron mayores en la copa del árbol que en su base.

A nivel proteómico, la temporada 2003 encontramos dos componentes esenciales del citoesqueleto, actina (**spot #495, 6,2 v.c.**) y tubulina, (**spot#718, 2,6 v.c.**), sobreexpresados en tejido formador de madera de copa. Estas proteínas asociadas a citoesqueleto, pueden estar involucradas en el control de la división celular. En particular, los filamentos de actina son responsables por muchos aspectos del comportamiento celular, como división celular, movimiento intracelular y expansión celular.

La división celular y expansión son altamente demandantes en términos de síntesis de proteínas y requerimientos de energía, como esta demostrado por el alto contenido de proteína (Figura 1 Paiva et al.).

Interesantemente, en madera de copa se identifico la sobreexpresión de proteínas relacionadas con transcripción y modificación de proteínas

En las temporadas 2005 y 2006, se identificaron 5 proteínas de la categoría “destino de proteínas” en madera de copa: 2 proteínas relacionadas con proteólisis por ubiquitinación (**#6586: 8,4 v.c.** y **#292: 12,0 v.c.**) y 3 proteínas relacionadas a modificación de proteínas, 1 proteína hipotética de glicosilación (**#1965: 2,8**)y 2 proteínas involucradas en plegamiento y estabilización de proteínas: 1 Calreticulina (**#250: 4,2 v.c.**) y 1 Proteína disulfuro isomerasa (**#2034: 3,3 v.c.**).

En madera de copa, también se identificaron 3 proteínas sobreexpresadas de la categoría funcional “transcripción”: El 2003,se identifico un complejo asociado a polipéptido naciente (**spot #1022, 24 v.c.**), demostrando la importancia de la síntesis de proteína en madera de copa, el 2005 y 2006 se identificaron 1 factor de iniciación de translación Eucariótico (**#6557: 15,3 v.c.**) y una proteína hipotética relacionada con síntesis de tRNA (**#2264: 8,0 v.c.**).

En Madera madura, en la categoría funcional “destino de proteínas”, se identifico sólo 1 proteína relacionada con degradación proteosomal (**#1523: 6,1 v.c.**) y una proteína de control transcripcional (**#5828: 2,7 v.c.**).

La sobreexpresión de categorías funcionales, y las observaciones microscopicas del tejido de copa, soporta el tejido de madera en formación de copa , como un lugar de activa división celular y transcripción de RNAs y expresión de proteínas que dan lugar a nuevas células.

### **Isozimas diferentes son reclutadas en madera de copa/juvenil o de base/madura.**

Se observó que isozimas diferentes de la misma proteína fueron reclutadas en tejido formador de madera de juvenil o madura, por ejemplo, en 2005, se identificó 1 spot correspondiente a Adenosilhomocisteinasa (**Fig. 38A**) en madera de madura (**#5703** Mw 62.71, pI 5.8), y una diferente localización en madera juvenil (**#5645**: Mw 65.56, pI 6.2), También, Fosfoglicerato deshidrogenasa mostró 2 spots en madera madura (**Fig. 38B**) (**#5826**, **#5858** Mw 55.33, 53.45 y pI 5.8, 5.3) y 2 spots diferentes spots en madera juvenil (**#5568**, **#5572** Mw 69.97,69.74 y pI 5.8, 5.9 respectivamente). Finalmente, Metionina sintasa Cobalamina-independiente (**Fig. 38C**), mostró 3 spots en madera madura (**#6483**: **#6484**, y **#5517**, con Mw 71.55, 72.38 y 72.26 kDa y con pI 5.9, 5.9 y 6.0 respectivamente) y solo un 1 spot en madera juvenil (**#6577**: 83.74 kDa, pI 6.1)

#### **4.1.9 Registro de datos proteómicos en base de datos PROTIC.**

Los datos proteómicos reunidos durante los análisis del 2005 y el 2006, fueron puestos a disposición de la comunidad científica, con libre acceso, en la base de datos PROTIC (<http://cbi.labri.fr/outils/protic/>) (Ferry-Dumazet et al., 2005), se incluyen las imágenes digitalizadas de todos los geles analizados en el gradiente base a copa, indicando la posición de cada una de las 132 proteínas identificadas, al seleccionar alguna proteína de interés (**Figura 39**), se despliegan las coordenadas de ubicación en el gel, peso molecular, pI, y el análisis de electrometría de masas con el detalle de los péptidos identificados. Una captura de la pantalla de este sitio se muestra en la **Figura 39**. Se espera que estos datos sean de utilidad para investigadores que analicen por técnicas proteómicas xilema en formación de especies coníferas y otras especies forestales.

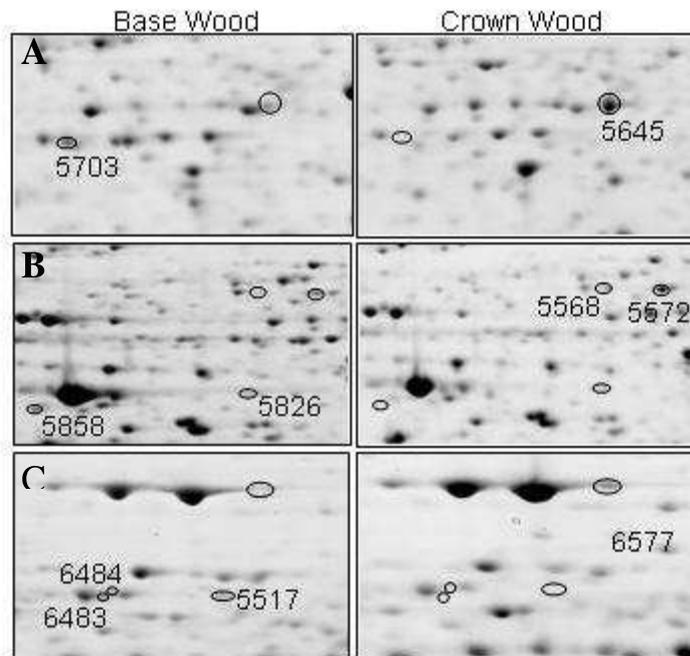


Figura 38 : Isozimas diferentes son recrutadas en madera de base y madera de copa

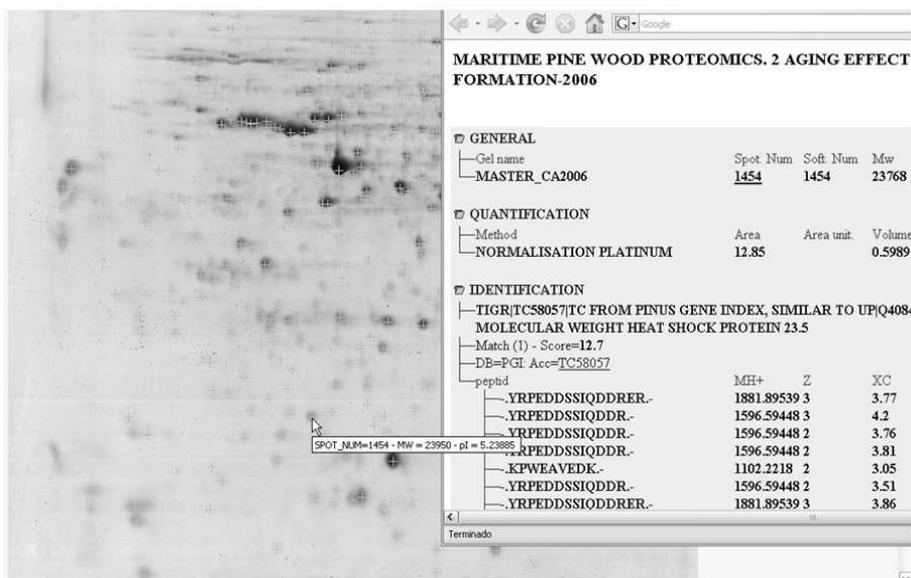
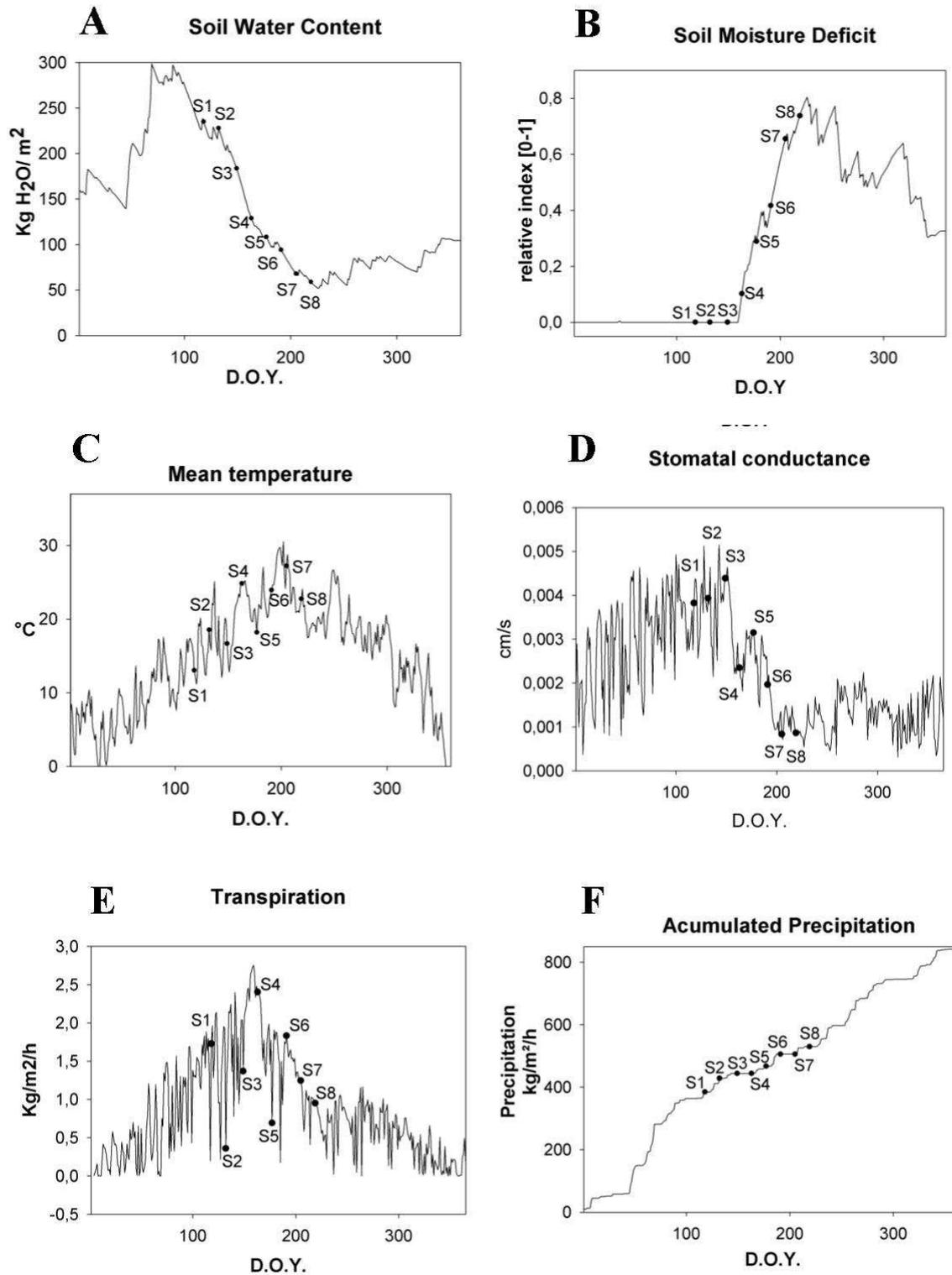


Figura 39: Captura de pantalla de la Base de datos PROTEOMICS, al seleccionar una proteína determinada, se muestra el detalle de su identificación, péptidos, pI y peso molecular

## **4.2 Análisis Fenotípico y Proteómico del Gradiente de temporada 2006**

**4.2.1 Datos Ecofisiológicos.** Las muestras de xilema secundario en formación fueron colectadas a lo largo de la temporada de crecimiento, desde la mitad de la primavera a la mitad del verano boreal. La **Figura 40** resume la variación anual de los datos Ecofisiológicos para la temporada 2006. Las fechas de recolección de muestra se indican en los gráficos desde S1 a S8, el día del año (DOY: day of the year) correspondiente. El gradiente de colección de las muestras, coincidió con una continua disminución del contenido de agua en suelo (**Figura 40A**). El déficit hídrico del suelo aumentó notablemente luego del punto S4 (**Figura 40B**). A partir de este punto, junto con la ocurrencia de las máximas temperaturas estivales (**Figura 40C**), se evidencian reacciones fisiológicas frente al estrés hídrico del verano, como la disminución de la conductancia estomática (**Figura 40D**) y la transpiración (**Figura 40E**). El déficit hídrico del suelo, si bien sigue un constante aumento a partir del punto S4, disminuye levemente por efecto de las precipitaciones registradas entre los puntos S5-S6 y S7-S8 (**Figura 40B, 40F**).



**Figura 40:** Datos ecofisiológicos para la temporada 2006 en una población de pinos marítimos adultos establecidas en la unidad INRA Pierroton.

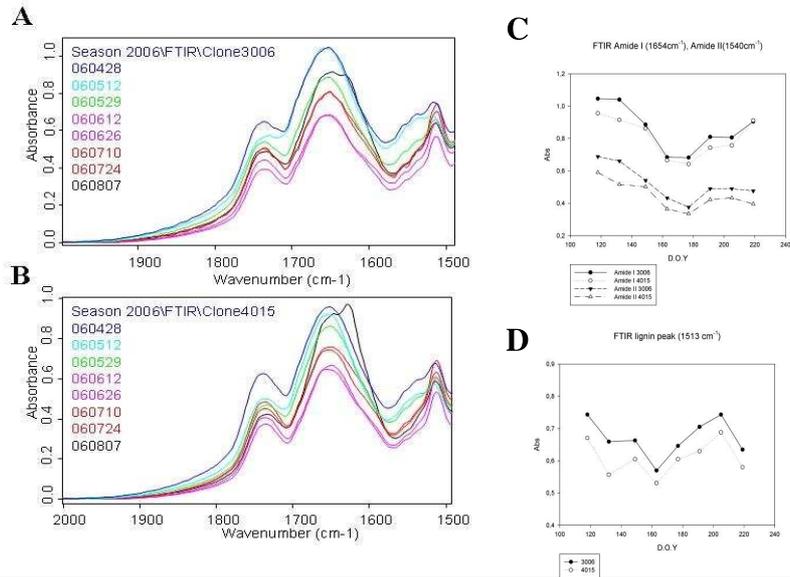
## **4.2.2 Variabilidad de la composición química de la pared celular a lo largo de la temporada**

### **4.2.2.1 Análisis de Infrarrojo con transformada de Fourier (FTIR)**

Se obtuvieron espectros de análisis infrarrojos con transformada de Fourier para todas las muestras a lo largo del gradiente de temporada (S1-S8). Los espectros infrarrojos mostraron perfiles similares para ambos clones (3006 y 4015), descartando comportamientos diferentes entre genotipos a través de la temporada de crecimiento. En la **Figura 41** una banda fuerte con un máximo a  $1654\text{cm}^{-1}$  fue atribuida al grupo amida I y a  $1540\text{cm}^{-1}$  una banda “hombro” fue atribuida al grupo amida II. A  $1513\text{cm}^{-1}$  se registró un pico agudo representativo de lignina (**Figuras 41C, 41D**).

Las bandas amida I y II (**Figura 41C**), que representan la cantidad de proteínas, mostraron su mayor nivel entre S1-S3 decreciendo hasta un nivel mínimo en S4, luego de S5, las amidas II aumentaron discretamente hasta S8, sin alcanzar los niveles iniciales, las amidas I mostraron similar perfil, sin embargo al analizar los espectros, se aprecia que la absorción de la banda amida I, para la muestra S8 se ve influenciada por una nueva absorción aproximadamente a  $1630\text{cm}^{-1}$  (**Figura 41A,B**), por lo que para este punto se deberá confirmar la cantidad de proteínas por otras técnicas.

El pico que representa el contenido de lignina, a  $1513\text{cm}^{-1}$  (**Figura 41A,B**) mostró en el espectro una forma irregular desde S1 a S3, posiblemente debido a interferencia con vibraciones del grupo amida II, lo que hace poco confiable las mediciones entre estos puntos, sin embargo desde S4 a S7 mostró mayor resolución y aumento sostenido hasta S7, donde alcanzo su máximo nivel.



**Figura 41:** A, B: Espectros FTIR para clones 3006 y 4015 a lo largo de la temporada 2006. C, D: Comparación de datos de absorción para Amina I y II y lignina.

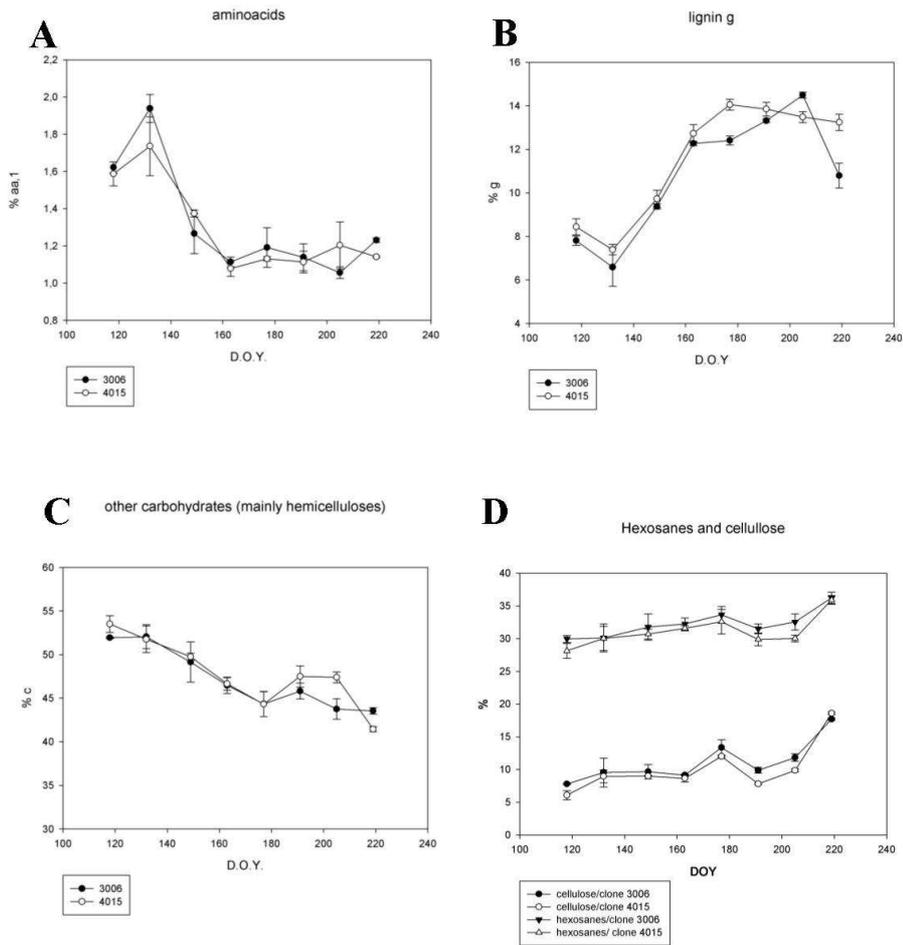
#### 4.2.2.2 Pirólisis analítica.

Ambos clones (3006 y 4015), mostraron perfiles similares a través de la temporada de crecimiento, no registrándose variaciones debido al genotipo. El contenido de proteínas estimado por productos de pirólisis proveniente de aminoácidos (**Figura 42A**), mostró una tendencia a disminuir a lo largo de la temporada desde S1, S2 ( $1.7 \pm 0.16\%$ ) hasta S3, manteniendo un bajo nivel hasta S8 ( $1.2 \pm 0.08\%$ ). Los resultados de la pirólisis confirmaron la mayor cantidad de proteínas detectada por FTIR al inicio de la temporada.

La cantidad de lignina, estimada por las unidades G, aumentó significativamente a lo largo de la temporada de crecimiento, desde S1, S2 ( $7.6 \pm 0.8\%$ ), S3 ( $9.5 \pm 0.2\%$ ) hasta S4-S8 ( $13.1 \pm 0.6\%$ ). Niveles similares fueron reportados previamente ( $12.5 \pm 3.7\%$  unidades G) para tejido formador de madera (Paiva, 2006). Esta cuantificación viene a confirmar el aumento en la cantidad de lignina a través de la temporada detectado por FTIR y se ajusta a las características de la madera tardía, donde existe un mayor porcentaje de lignina con respecto a la madera temprana, por el mayor grosor de la pared celular.

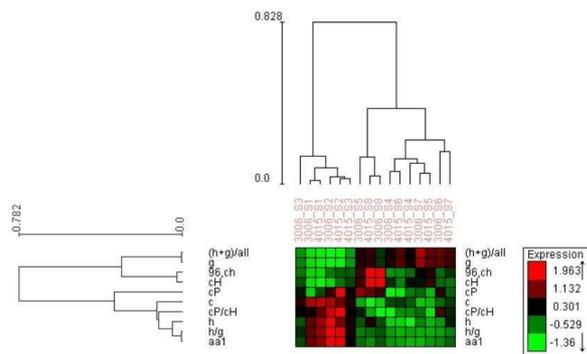
Los productos derivados de carbohidratos que no provienen de hexosanos o pentosanos (principalmente provienen de hemicelulosas) disminuyeron a lo largo de la temporada de crecimiento (**Figura 42C**), desde S1 ( $52.7 \pm 1.1\%$ ), gradualmente hasta S5 ( $44.3 \pm 0.0\%$ ), y luego aumentaron en S6 ( $46.6 \pm 1.2\%$ ), para continuar disminuyendo hasta S8 ( $42.5 \pm 1.5\%$ ), este hecho es coincidente con la mayor proporción de pared celular primaria presente en la madera temprana, que se va reduciendo a lo largo de la temporada. Hexosanos y celulosa siguen un perfil opuesto a las hemicelulosas (**Figura 42D**), los hexosanos aumentaron desde S1 ( $29.1 \pm 1.3\%$ ), hasta S5 ( $33.1 \pm 0.8\%$ ), luego disminuyeron en S6 ( $30.7 \pm 1.2\%$ ), para aumentar hasta S8 ( $36.1 \pm 0.2\%$ ). Celulosas aumentaron desde S1 ( $6.9 \pm 1.2\%$ ), hasta S5 ( $12.7 \pm 1.0\%$ ), disminuyeron en S6 ( $8.9 \pm 1.5\%$ ), para aumentar nuevamente hasta S8 ( $18.2 \pm 0.6\%$ ).

Hemicelulosas y Hexosanos (entre ellos celulosa), siguen una tendencia constante de disminución o aumento, respectivamente a lo largo de la temporada, de manera interesante, cambios ambientales como la precipitación registrada entre S5 y S6 (**Figura 40B**), pueden influir temporalmente en la composición de polisacáridos de la pared celular, por ejemplo en el punto S6, luego de este punto la tendencia anual es retomada en ambos genotipos (**Figura 42 C y D**). Esto demuestra la importante respuesta ambiental en la expresión de polisacáridos de pared celular para ambos genotipos, 3006 y 4015.



**Figura 42:** Pirólisis analítica de madera en formación a través de un gradiente de temporada 2006



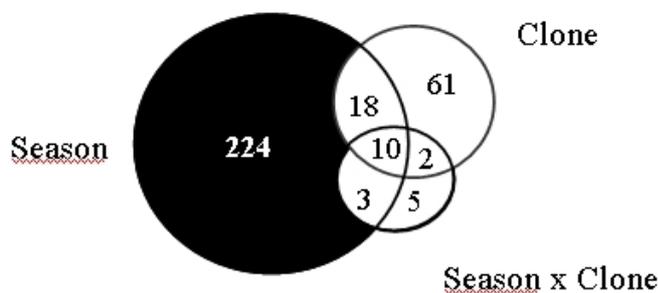


**Figura 44 :** Clustering jerárquico de los productos de pirólisis analítica.

### 4.2.3 Mapas bidimensionales de tejido formador de madera.

Se construyeron mapas bidimensionales, para estudios proteómicos comparativos de tejido formador de madera a través de un gradiente de temporada.

El volumen de 1052 spots fue analizado por ANOVA, esto llevó a la selección de 323 spots que mostraron una expresión diferencial significativa ( $P < 0.001$ ), para el efecto temporada, efecto clonal o el efecto interacción (temporada x clon), de estas 323 spots, 224 fueron significativas solo para el efecto temporada (**Figura 45**). Las 224 proteínas fueron posteriormente analizadas, para la identificación de aquellas proteínas sobrepresadas en madera temprana y madera tardía.

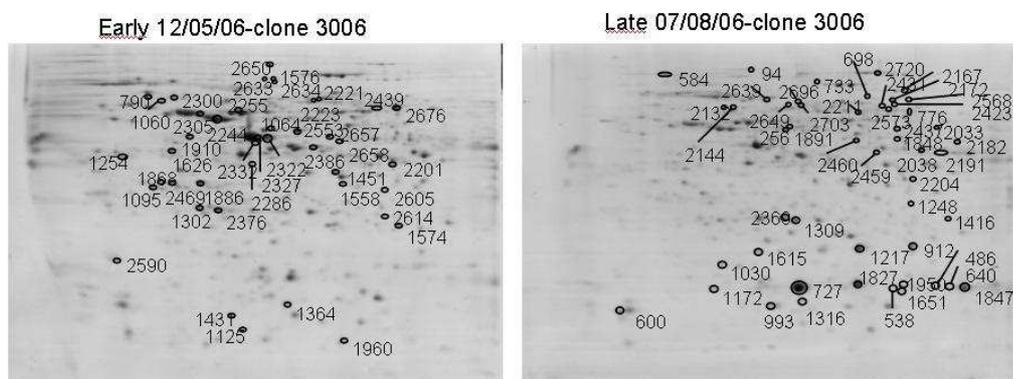


**Figura 45:** Diagrama de Venn, mostrando la proteínas diferencialmente expresadas que son significativas ( $p < 0,001$ ) para los efectos temporada, clon e interacción.

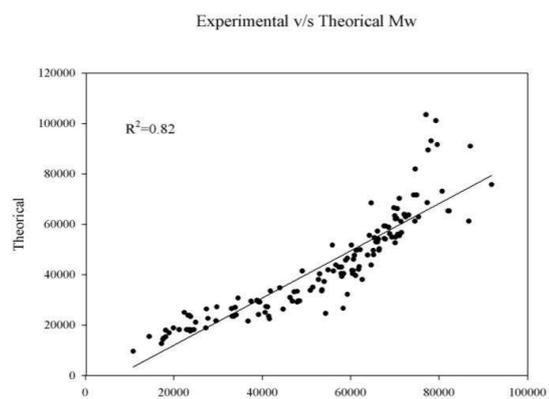
#### 4.2.4 Identificación de proteínas en tejido formador de madera y clasificación funcional.

Como criterio adicional de selección, se eligieron proteínas sobrepresadas siguiendo un perfil continuo creciente o decreciente a través de la temporada, las imágenes digitalizadas de los geles fueron revisadas manualmente, para asegurar la selección de proteínas con una alta expresión, resolución y reproducibilidad entre los replicados de geles. Luego de la extracción manual de las proteínas seleccionadas desde los geles, se analizaron 94 proteínas por LC ESI MS/MS y se identificaron utilizando la base de datos TIGR (TIGR Pinus Gene Index database). 93 proteínas lograron ser identificadas (99%), 42 y 51 proteínas fueron sobrepresadas en tejido formador de madera temprana y tardía, respectivamente. Geles representativos de madera temprana y madera tardía se muestran en la **Figura 46**, las proteínas identificadas en cada gel fueron finalmente agrupadas en categorías funcionales en la **tabla 7**.

Una manera de verificar la correcta asignación de identidades a las proteínas es analizar la regresión lineal entre el peso molecular teórico vs. peso molecular experimental, (**Figura 47**) el coeficiente de correlación alcanzó un alto valor  $r^2 = 0.82$ , sin embargo luego de los 78 kDa la regresión lineal no explica la relación entre los pesos moleculares, debido al porcentaje de acrilamida utilizado para la electroforesis, ya que a muy altos pesos moleculares no se sigue una relación lineal entre la distancia de migración y el peso molecular. En algunos casos la aparente falta de correlación se puede deber a que los valores de peso molecular y pI son estimados en base a proteínas de otras especies.



**Figura 46:** Geles representativos de madera temprana y madera tardía, mostrando las proteínas identificadas, las etiquetas siguen la numeración de la tabla 7



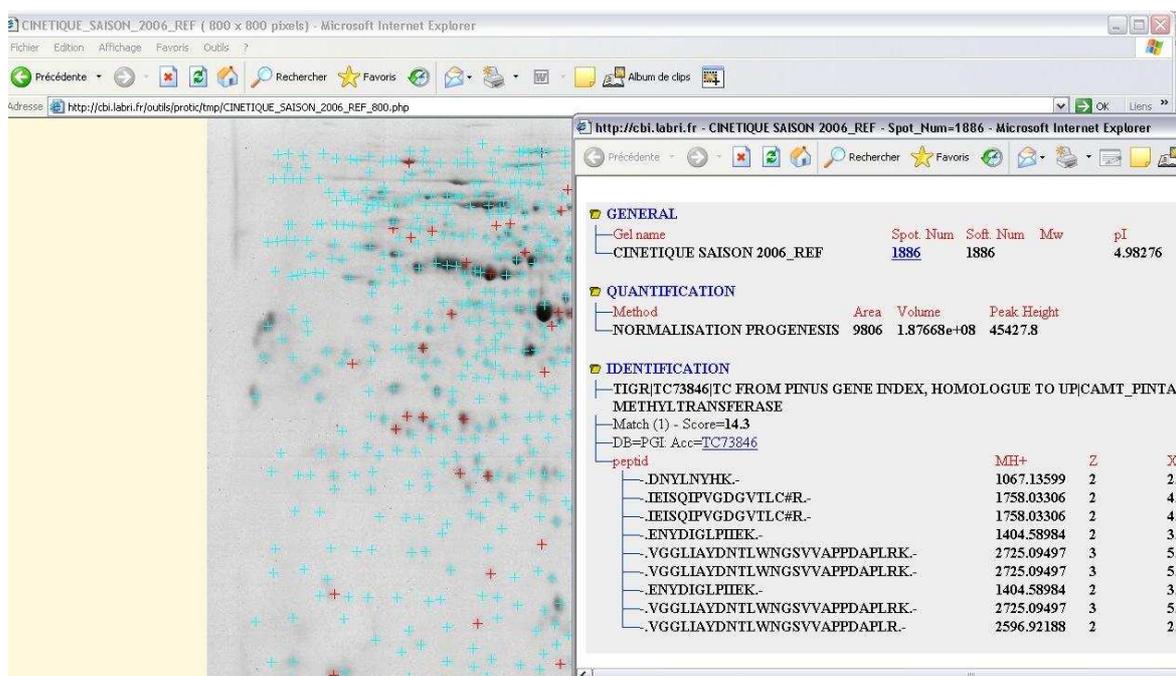
**Figura 47:** Regresión lineal del peso molecular experimental vs. el peso molecular teórico, el  $R^2$  obtenido es 0,82

Early Wound 01.01 amino acid metabolism overexpressed	spots	ID	scans	relative abundance	TC	Assignment	UniProt KB Entry	MIFS	Gene	Ontology GO	exp pi	exp Mw	Fold change											
													12/20/08	12/24/15	12/30/08	12/40/15	15/30/08	15/40/15	18/30/08	18/40/15				
01.03 nucleotide metabolism	2386	66	57.61	TC58314	Glutamine synthase	G292592	01.03.03.01.01 biosynthesis of glutamine	5.67		GO:000542	glutamine biosyn	5.74	39293	1.3	1.9	1.2	1.2	1.1	1.0	1.0				
		67	15.09	TC58319	UDP-glucosyl transferase 1	G292591		5.63				glutamine biosyn	4.84	39294										
		57	15.09	TC58319	UDP-glucosyl transferase 1-like protein SUPTG1	G292591		5.63				glutamine biosyn	4.84	39294										
		22	3.22	TC57711	Fructose kinase 3	Q8VW45		5.57				carbohydrate m.	39137											
		22	3.19	TC57711	Probable mannitol dehydrogenase	Q9ZRF1		6.46				carbohydrate m.	39137											
		1910	4.3	81.6	TC64867	Serine hydroxymethyltransferase	Q23254	01.01.09.02.01 biosynthesis of serine	6.8	4.91	GO:0005654	L-serine biosynt	60156	2.0	1.7	1.6	1.5	1.8	1.0	1.2				
			30	18.4	TC73846	Citruoyl-CoA methyltransferase	Q9ZT75		5.44				carbohydrate m.	29145										
	01.05 C-compound and carbohydrate metabolism	1302	22	44.79	TC65720	Uridylate kinase (UK) (Uridine monophosphate kinase) (UMP-kinase)	Q04065	01.03.04.04	5.79		GO:0005220	pyrimidine nucleot metabolism	22481	41642	1.7	1.7	1.4	1.2	1.3	1.2	1.3	1.0		
			44	55.21	TC65637	Triosephosphate isomerase, cytosolic (TIM)	Q04066		5.39				pyrimidine nucleot metabolism	27169										
			2200	60.1	TC65659	Beas-13-oligomannose 6-phosphate 4-epimerase	Q8TAD7		4.79				carbohydrate m.	70089	2.2	1.7	1.8	1.9	1.8	1.6	1.0	1.4		
			22	2.41	TC74186	Rubisco subunit protein alpha subunit, chloroplast	P08262		4.79				carbohydrate m.	56792										
			22	2.41	TC74234	TSE1.7	Q9M226		4.63				carbohydrate m.	56430										
			2439	1613	100	TC57058	Aldehyde dehydrogenase	Q8LLR2	1.05	5.920		GO:0005975	carbohydrate m.	67836	2.4	2.2	2.1	2.0	1.7	1.7	1.0	1.2		
			2576	1611	100	TC57058	Aldehyde dehydrogenase	Q8LLR2	1.05	6.33		GO:0005975	carbohydrate m.	67827	2.1	1.9	2.2	2.1	2.0	1.6	1.0	1.1		
			2553	88	96.44	TC82024	S-adenosyl-L-methionine synthetase	Q94404	1.05	5.42		GO:0005970	one-carbon corr	61939	2.4	3.1	2.0	1.8	1.8	2.0	1.0	1.2		
			1626	33	1.55	TC92529	Actin	Q5S9Y7		5.3			carbohydrate m.	41390										
			1502	139	100	TC68124	Fructose kinase 3	Q8VW45		5.57			carbohydrate m.	56145	1.8	1.8	1.6	1.5	1.3	1.6	1.0	1.1		
		02 ENERGY	2221	1212	92.12	TC71568	Dihydroxyacetone kinase 1	P54838	01.06.01.03	5.25		GO:0009247	glycolipid biosynth	62206	70373	1.9	2.2	1.9	1.6	1.5	1.7	1.0	1.3	
2223			1211	85.86	TC72890	5-methyltetrahydropteroyldiamate-homocysteine methyltransferase	Q42699	01.06.01.03	6.1		GO:0009247	glycolipid biosynth	84657	70373	1.9	2.2	1.9	1.6	1.5	1.7	1.0	1.1		
2223			1211	85.86	TC70006	Dihydroxyacetone kinase 2	P43550	01.06.01.03	5.59		GO:0009247	glycolipid biosynth	84657	70373	1.9	2.2	1.9	1.6	1.5	1.7	1.0	1.1		
2286			33	5.38	TC58763	Phosphoglycerate dehydrogenase-like protein	Q8LJG6		6.32				glycolipid biosynth	63309										
2332			1612	96.12	TC58276	Enoyl-ACP reductase precursor	Q05046	01.06.01.05	5.31			GO:0006833	fatty acid biosyn	53395	1.8	1.8	1.5	1.3	1.1	1.4	1.0	1.4		
			1612	96.12	TC58276	Enoyl-ACP reductase precursor	Q05046	01.06.04.05	6.05			GO:0005962	fatty acid catabol	45740	58857	2.2	1.8	1.7	1.3	1.6	1.5	1.0	1.2	
			22	1.22	TC73003	Malate dehydrogenase 2, mitochondrial precursor	Q8LJ43		5.59				glycolysis and gluconeogenesis	33366										
02.01 glycolysis and gluconeogenesis			2201	715	78.12	TC73572	Fructose-bisphosphate aldolase, chloroplast precursor	Q40877	2.01	5.35		GO:0006096	glycolysis	52716	2.1	1.5	1.8	1.7	1.6	1.1	1.3	1.0		
	2376		115	93.24	TC89527	Triosephosphate isomerase, cytosolic (TIM)	P48491	2.01	5.07		GO:0006094	gluconeogenesis	41196	1.7	1.8	1.7	1.3	1.5	1.0	1.2	1.1			
			30	2.41	TC57200	Uridylate kinase (UK) (Uridine monophosphate kinase) (UMP-kinase)	Q04065		5.79				gluconeogenesis	22481										
			30	1.93	TC58720	Phosphoglycerate dehydrogenase-like protein	Q8LJG6		4.3				gluconeogenesis	63309										
			30	1.63	TC58720	Phosphoglycerate dehydrogenase-like protein	Q8LJG6		4.3				gluconeogenesis	63309										
			22	0.87	TC57750	Alpha tubulin	Q8H633		4.88				glycolysis and gluconeogenesis	46513										
	02.02 tricarboxylic-acid pathway (citrate cycle, Krebs cycle, TCA cycle)		1451	107	88.13	TC73537	NAD-malate dehydrogenase precursor	Q9XP42	2.10	5.67		GO:0006099	tricarboxylic acid	50899	1.7	1.9	1.6	1.3	1.4	1.1	1.1	1.0		
				77	10.87	TC57058	Aldehyde dehydrogenase	Q8LLR2		6.33				tricarboxylic acid pathway (citrate cycle)	59920									
			2657	2115	96.25	TC65986	Alcohol dehydrogenase (Fragment)	Q43923	02.13.01	5.68		GO:0009361	anaerobic respir	40480	6.044	1.9	2.4	1.8	1.6	1.7	1.4	1.1	1.0	
				33	3.75	TC73534	Actin 2	Q92564		5.31				anaerobic respir	41626									
			1574	64	100	TC65998	Putative quinone-oxidoreductase OR2	Q6ZKJ0	02.30.07	6.09			accessory proteins of photosyntheloid	36908	1.7	1.5	1.6	1.6	1.2	1.2	1.1	1.0		
		11 TRANSCRIPTION 11.02 RNA synthesis	1254	99	96.47	TC57365	Putative elongation factor	Q1XG58	11.02.03.01.04	4.56		GO:0006354	RNA elongation	54362	2.0	2.2	1.6	1.4	1.2	2.0	1.1	1.0		
			1451	63	100	TC72911	TGB12K interacting protein 2	Q8HP56		4.46			transcription	37342										
			1576	99	80.5	TC58135	Glycine-rich RNA-binding protein	Q9XFL4	11.02.03.01	7.85		GO:0006350	transcription	14420	2.1	3.0	1.3	1.3	1.4	1.3	1.0	1.0		
			2633	87	68.22	TC51078	Putative transcription factor (Acyl-CoA binding family) pQ7SL4	Q92344	11.02.03.04	5.23		GO:0045449	regulation of tra	71543	74867	2.2	2.3	1.7	1.4	1.7	1.4	1.0	1.1	
			2633	44	31.78	TC57694	Putative transcription factor (Acyl-CoA binding family) pQ7SL4	Q92344	11.02.03.04	5.41		GO:0045449	regulation of tra	71543	74867	1.7	1.8	1.6	1.3	1.6	1.3	1.2	1.0	
			2634	99	93.95	TC58038	Putative transcription factor (Acyl-CoA binding family) pQ7SL4	Q9M555	11.02.03.04	5.16		GO:0045449	regulation of tra	73974	74867	1.8	1.8	1.4	1.1	1.3	1.0	1.0	1.1	
				22	6.05	TC78636	Acyl-peptide hydrolase-like	Q9FG66		5.08				regulation of tra	75422									
14 PROTEIN FATE 14.01 protein folding and stabilization			780	1512	100	TC57686	Protein disulfide-isomerase precursor (PDI)	Q9XFE1	14.01	4.84		GO:005921	protein stabilizer	70130	2.0	2.3	1.3	1.5	1.3	2.1	1.0	1.0		
			1060	33	100	TC57686	Protein disulfide-isomerase precursor (PDI)	Q9XFE1	14.01	4.75		GO:005921	protein stabilizer	69310	2.2	2.2	1.8	1.6	1.6	1.0	1.2	1.0		
			1060	33	100	TC57686	Protein disulfide-isomerase precursor (PDI)	Q9XFE1	14.01	6.23		GO:005921	protein stabilizer	35173	1.8	1.7	1.5	1.3	1.4	1.2	1.1	1.0		
			16 PROTEIN WITH BINDING FUNCTION OR COFACTOR REQUIREMENT 16.01 protein binding	1095	77	83.68	TC66140	14-3-3 protein	Q9XW68	16.01	4.78		GO:0005515	protein binding	46853	1.6	1.5	1.7	1.5	1.4	1.0	1.0	1.0	
					32	6.28	TC57949	Carboxyl-terminal proteinase-like	Q8H7H9		4.32			protein binding	24338									
					5/5	10.04	TC68845	GF14 protein	Q49362		4.79			protein binding	29184									
	1865			43	52.4	TC72978	14-3-3 protein	Q9LKK9	16.01	4.73		GO:0005515	protein binding	48430	2.4	2.4	2.3	1.9	1.7	1.0	1.4			
	2469			4/4	47.5	TC58959	L-galactose-1-phosphate phosphatase	Q5U789		5.22			protein binding	28048										
				7/4	71.9	TC58959	L-galactose-1-phosphate phosphatase	Q5U789		5.22			protein binding	28048										
				8/6	71.5	TC74635	14-3-3 protein (Acyl-CoA binding family) pQ7SL4	Q9M555	16.01	6.09		GO:0005515	protein binding	46156	1.9	1.7	2.1	2.2	1.8	1.6	1.0	1.3		
				22	5.16	TC58929	L-galactase-1-phosphate phosphatase	Q5U788		5.22			protein binding	45740										
	18 PROTEIN ACTIVITY REGULATION 18.02.01 enzymatic activity regulation	1364		33	100	TC57404	14 kDa zinc-binding protein (Protein kinase C inhibitor)	P42555	18.02.01.02.05	6.56		GO:0019210	kinase inhibitor	17197	1.8	2.0	1.1	1.4	1.1	1.3	1.1	1.0		





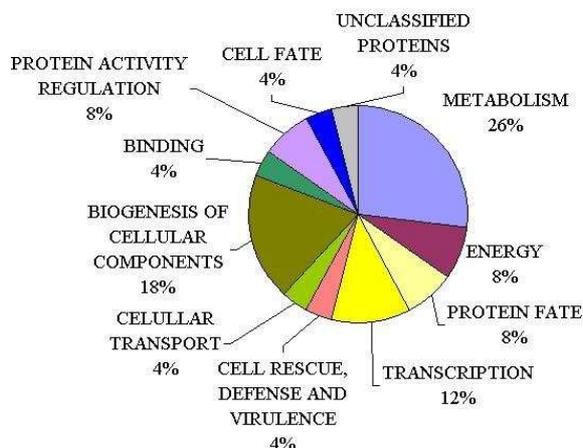
Los datos obtenidos desde el análisis proteómico, fueron hechos disponibles en la base de datos PROTIc del centro Bioinformático de Burdeos (<http://cbi.labri.fr/outils/protic/>), en esta base de datos se encuentran las imágenes digitalizadas de los geles analizados, con el detalle de los datos de espectrometría de masas para cada una de las proteínas identificadas. Una captura de pantalla del sitio de puede ver en la **Figura 48**.



**Figura 48** : Captura de Pantalla de la base de datos PROTIc, que contiene la información de los geles analizados con las proteínas y el detalle de los péptidos identificados.

La clasificación de las proteínas identificadas se realizó asignándolas a categorías funcionales siguiendo el catalogo funcional MIPS (Munich Information Center for Protein Sequences). Este catálogo provee un sistema general de anotación estable (sólo cuatro extensiones mayores se han requerido desde 1996). Las proteínas que aumentan su expresión en 2 o más veces (veces de cambio: v.c.) fueron consideradas significativamente sobreexpresadas. A continuación se analizan sólo las proteínas sobreexpresadas en tejido formador de madera temprana y madera tardía, agrupándolas por categorías funcionales.

## EARLY WOOD



**Figura 49** : Clasificación por categorías funcionales de las 26 proteínas sobreexpresadas en tejido formador de madera asociado a madera temprana

### 4.2.4.1 Proteínas sobreexpresadas en Madera Temprana o de Primavera.

26 Proteínas fueron sobreexpresadas significativamente ( $>2$  veces cambio “v.c.”) en tejido formador de madera asociado a madera temprana, (**Figura 49**) la categoría funcional más representada fue “metabolismo”, con 7 proteínas identificadas (26% del total de proteínas sobreexpresadas), 4 están relacionadas con el metabolismo de compuestos C: una Beta-1,3-glucanasa ácida (**#2300, v.c. 2.2**), frecuentemente identificada por su función de proteína PR, también participa de procesos de desarrollo (Doxey et al., 2007) degradando la callosa (beta-1,3 glucan) depositada transientemente durante la formación de la placa celular en la división celular (Samuels, 1995; Scherp et al., 2001).

2 Aldehído deshidrogenasas (ALDH) (**#2439, #2576; v.c. 2.4, 2.1**), homólogas a la ALDH2B7 Ath de Arabidopsis thaliana son sobreexpresadas en madera temprana, estas proteínas se expresan en mitocondria y se ha reportado su participación en detoxificación de aldehído luego de estrés por hipoxia en arroz (Kirch et al., 2004), igualmente se sugiere un rol para estas ALDH durante fermentación alcohólica (Tsuji et al., 2003).

Se identificó una S-adenosil-L-metionina sintetasa (**#2553, v.c. 2.4**), involucrada en el ciclo del metil activo, además una Serina hidroximetiltransferasa (**#1910, v.c. 2.0**), que participa en las reacciones de transferencia de metilo través de la regeneración de 5,10-metileno tetrahidrofolato, catalizando la interconversión entre serina y glicina, esta es la principal vía para la incorporación de unidades de C para la biosíntesis de numerosos metabolitos celulares. (Schich and Szebenyi,2005). Esta enzima se ha encontrado en altos

niveles en células de alta proliferación y se ha propuesto como una enzima objetivo para el desarrollo de agentes anticancerígenos y antimicrobianos (Agrawal et al., 2003 )

Las reacciones de transferencia de metil activado son importantes para distintas reacciones biosintéticas en las células vegetales, y la abundancia de estas proteínas en la madera temprana refleja la demanda de reacciones de transferencia de metilo. Además glicina es uno de los aminoácidos más abundantes en las proteínas estructurales de la pared celular (Allona et al., 1998)

2 proteínas de la categoría “metabolismo de lípidos, ácidos grasos e isoprenoides”: una Dihidroxiacetona quinasa 1 (**#2221, v.c. 2.2**) y una hidrolasa Hydroxi-isobutiril-coenzima A -like, (**#2332, v.c. 2.2**), estas proteínas son importantes para mantener la disponibilidad de lípidos para la formación de nuevas células.

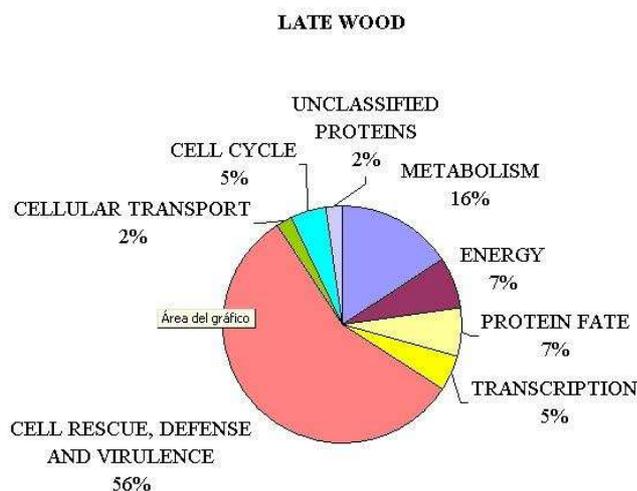
En la categoría funcional “biogénesis de componentes celulares” (18%), se identificaron 4 proteínas involucradas en formación de citoesqueleto: 2 Actinas (**#2322, #2327; v.c. 2.5, 2.8**), una cadena beta-4 Tubulina (**#2305, v.c. 2.0**) y proteína homóloga de tumor controlada translacionalmente (**#2590, v.c. 2.0**), involucrada en enlace de calcio y estabilización de microtúbulos, la sobreexpresión de estas proteínas demuestra la importancia de la organización del citoesqueleto de nuevas células en madera temprana. Además se identificó una Cafeoil-CoA O-metiltransferasa (**#1886, v.c. 2.1**) fundamental en biosíntesis de lignina para su depósito en la pared celular. La categoría “biogénesis de componentes celulares” no está presente en madera tardía.

En la categoría funcional “transcripción” (12%), se identificaron 3 proteínas relacionadas con síntesis de ARN, un factor de elongación putativo (**#1254, v.c. 2.2**), una proteína de unión a ARN rica en glicina (**#1431, v.c. 3.0**) y un factor de transcripción putativo (familia unión Acil-CoA) (**#1576, v.c. 2.3**). En la categoría “regulación de actividad de proteínas” (8%) se identificó una proteína de 14kDa de unión a zinc (inhibidora de proteína quinasa C) (**#1364, v.c. 2.0**) y una multicistatina (**#1960, v.c. 2.1**) (inhibidora de cisteín proteasas). Las categorías funcionales “transcripción” y “regulación de actividad de proteínas“, suman un 20% del total de las proteínas sobreexpresadas en madera temprana, esto demuestra la importancia de la transcripción de nuevos RNAs y la regulación de la actividad proteica en madera temprana, actividades derivadas de la intensa división celular.

En la categoría funcional “energía”, con un (8%) se identificaron 2 proteínas, un precursor cloroplástico de Fructosa bisfosfato aldolasa (**#2201, v.c. 2.1**) y una Alcohol

deshidrogenasa (#2657, v.c. 2.4), que participa en la generación de ATP de manera anaeróbica.

Otras categorías representadas por solo una proteína fueron: “rescate celular, defensa y virulencia” con una Peroxidasa (#1064, v.c. 2.9), “transporte celular” con una Probable ATPasa de transporte de H<sup>+</sup> (#2255, v.c. 2.1), la función “enlace (binding)”, con una proteína 14-3-3 (#1865, v.c. 2.4), “destino celular” con un modulador de expresión GLABRA2 (#2658, v.c. 2.1) y una “proteína no clasificada”, Sec13p (#1558, v.c. 2.0).



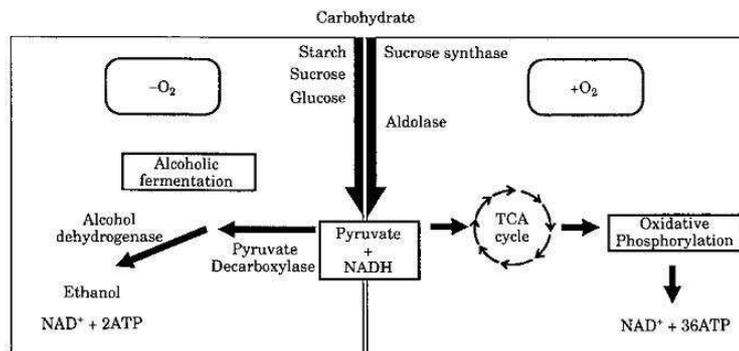
**Figura 50:** Clasificación por categorías funcionales de las 44 proteínas sobreexpresadas en tejido formador de madera asociado a madera tardía

#### 4.2.4.2 Proteínas sobreexpresadas en Madera Tardía o de verano.

Se encontraron 44 proteínas sobreexpresadas en madera tardía, (**Figura 50**) la categoría “rescate celular, defensa y virulencia” fue significativamente sobre representada (56%) esto contrasta con el 4% de proteínas de esta categoría sobreexpresadas en madera temprana. Se identificaron 19 HSP : 7 precursores cloroplásticos de pequeñas HSP, (**#912, #993, #1217, #1030, #640, #1615, #2369; v.c. 9.7, 5.7, 14.7, 2.5, 3.4, 3.4, 3.7**), 5 HSP de bajo peso molecular (**#1950, #727, #486, #538, #1847 ;v.c. 4.1, 4.6, 4.9, 4.1, 6.2**), 3 HSP 26 (tipo I) (**#2459, #1827, #94; v.c. 4.2, 7.9, 4.9**); una HSP101 (**#2720, v.c. 4.7**), una HSP putativa hsp20 (**#1172, v.c. 4.0.**), una HSP 90 (**#584, v.c. 4.2**) y una proteína similar a Taumatina (**#1316, v.c. 5.0**), en la misma categoría funcional se identificaron 5 proteínas de respuesta a estrés (**#698, #2568, #2172, #600, #776, v.c. 2.7, 3.8, 5.6, 3.0, 4.8**) y una proteína Ascorbato peroxidasa (**#1309, v.c. 2.8**), fundamental en la respuesta a estrés oxidativo.

La segunda categoría funcional en madera tardía fue “metabolismo” (16%), se identificaron 5 proteínas del metabolismo de C: precursor mitocondrial, de isoforma de 62 kDa de la enzima málica dependiente de NAD, (**#2431, v.c. 3.0**); manosa 6-fosfato reductasa dependiente de NADPH (**#2204, v.c. 3.7**), beta glucosidasa (**#2132, v.c. 2.5**); fosfoglucomutasa (**#2639, v.c. 2.3**); y GDP-manosa 3,5-epimerasa (**#1848 v.c. 2.5**); en esta categoría funcional también se identificó una Fosforibosilaminoimidazolcarboxiamida formiltransferasa/IMP ciclohidrolasa (**#2573, v.c. 4.1**), involucrada en el metabolismo de nucleótidos y una mevalonato difosfato descarboxilasa (**#2033, v.c. 3.0**), del metabolismo

En la categoría funcional “energía” (7%) se identificó: Isocitrato deshidrogenasa NADP+ (#2182, v.c. 5.6) y 2 Alcohol deshidrogenasas (#2191, #2038; v.c. 3.3, v.c. 2.1). La alcohol deshidrogenasa cumple un rol central en el metabolismo anaeróbico, (Perry and Furnier, 1996) (**Figura 51**). Bajas concentraciones de etanol se encuentran a menudo presentes en tejidos leñosos de árboles saludables (Joseph and Kelsey, 2004), cuando la cantidad de oxígeno celular es abundante, se sintetizan cantidades mínimas de etanol ya que prácticamente todo el piruvato producido en la glicólisis es metabolizado por la piruvato deshidrogenasa mitocondrial e ingresa al ciclo del ácido tricarboxílico, para la producción de ATP por fosforilación oxidativa. Si el oxígeno se vuelve limitante, la fosforilación oxidativa se detiene rápidamente y la fermentación alcohólica se inicia luego de un breve periodo de retraso. La fermentación alcohólica en condiciones anaeróbicas, parece funcionar como un sistema de respaldo para mediar la acidosis celular (la fermentación alcohólica no afecta el pH como la fermentación láctica), junto con generar NAD<sup>+</sup> y ATP para mantención celular, hasta que la fosforilación oxidativa es restablecida. La fermentación alcohólica no acidifica el citoplasma, a diferencia de la fermentación láctica. El lactato es un ácido y su acumulación en el citoplasma puede alterar el pH celular causando daño, mientras que el etanol difunde al medio extracelular y no causa mayores problemas excepto a altas concentraciones (Tadege and Kuhlemeier, 1997). La síntesis de etanol en presencia de oxígeno (fermentación aeróbica) ha sido descrita funcionando simultáneamente con la respiración aeróbica en polen de Tabaco *Nicotiana tabacum*. (Joseph and Kelsey, 1997). La familia de alcohol deshidrogenasas es mayor en pino que en angiospermas y el tamaño de los genes es mayor, en angiospermas se expresan típicamente 2 a 3 ADH, en *Pinus banksiana* se identificaron al menos 7 ADH, no existiendo hasta el momento explicación funcional a esta diferencia (Perry and Furnier, 1996). Además la acumulación de una ATP sintasa vacuolar, que transporta H<sup>+</sup> a la vacuola (#1248, v.c. 2.8), corrobora la hipótesis de acidosis citoplasmática durante la formación de madera tardía.



**Figura 51 :** Durante condiciones de hipoxia, se activan mecanismos de fermentación anaeróbica. La fermentación láctica lleva a acidificación del citosol e inhibición de la glicólisis, especies de plantas resistentes a condiciones anaeróbicas, poseen una fermentación alcohólica muy activa, esta fermentación es menos toxica y no afecta el pH de la célula (modificado de Perry, 1996)

En la categoría funcional "destino de proteínas" (7%) se expresó una proteína F13M7.3 (putativa disulfido isomerasa) (#1891, v.c. 2.5), involucrada en plegamiento de proteínas y estabilización y 2 proteínas involucradas en modificación de proteínas: peptidasa mitocondrial subunidad beta (#2211, v.c. 3.8), y proteasa Clp dependiente de ATP (#733, v.c. 3.4).

En la categoría funcional transcripción, se identificó la RNA helicasa 38 ATP-dependiente DEAD-box (#2703, v.c. 3.3), involucrada en el metabolismo de RNA y una glicil-tRNA sintetasa (#2167, v.c. 3.7).

De la categoría "ciclo celular y procesamiento de DNA" se expresaron 2 proteínas nucleares de unión a GTP RAN-A1 (#1416, #1651; v.c. 2.3, v.c. 3.5). La categoría funcional "transporte celular" expresó una subunidad E de ATP sintasa vacuolar (#1248, v.c. 2.8), involucrada en el transporte de protones hacia los organelos. Finalmente se identificó una proteína no clasificada (#2437 v.c. 2.2), en la categoría "proteínas no clasificadas"

#### **4.2.5 Discusión de los resultados del gradiente de temporada**

La mayor variación en la formación de madera ocurre a lo largo de la temporada de crecimiento, los resultados presentados en este capítulo, ponen en evidencia la plasticidad del tejido formador de madera en respuesta al cambio en las condiciones edafoclimáticas a

madera temprana a madera tardía es siempre abrupto (Ivković and Rozenberg, 2004).

Datos Ecofisiológicos muestran que después de la fecha de muestreo S4 las condiciones de estrés hídrico aumentaron, comenzando la formación de madera tardía, esto fue confirmado por análisis transcriptómicos preliminares que muestran una respuesta drástica a los cambios ambientales.

El análisis de agrupamiento jerárquico de los datos de pirólisis analítica separó dos grupos: 1) S1, S2 y S3, asociados con madera temprana, 2) S4, S5, S6, S7 y S8 asociados a madera tardía. En el análisis PCA de estos datos, el punto S8 agrupó relacionado con madera tardía, pero independientemente, por su mayor contenido de celulosa, y lignina en la pared celular, ya en el momento de colecta de muestra este tejido se mostró más lignificado que en los puntos recolectados anteriormente, igualmente la presencia de una gruesa pared celular dificultó la extracción de proteínas en madera tardía, que en algunos casos debió ser repetida. El agrupamiento obtenido por clustering jerárquico y por PCA fue válido para ambos clones, demostrando la mayor importancia del efecto ambiente sobre el efecto genotipo en este gradiente.

Los análisis químicos, mostraron que el contenido de lignina, hexosanos y celulosa se incrementaron continuamente a través de la temporada, con los hexosanos y celulosa mostrando variación en el punto S6, probablemente debido a la precipitación ocurrida antes de la colección de muestra. Otros carbohidratos, principalmente de origen hemicelulósico disminuyeron su porcentaje continuamente a lo largo de la temporada, por la acumulación de lignina y celulosa en la pared celular a lo largo de la temporada. Igualmente las hemicelulosas muestran variación en el punto S6, mostrando al igual que las hexosas, su dependencia a las condiciones ambientales.

Los análisis proteómicos demostraron sin lugar a dudas la importancia de los mecanismos de defensa celular al final de la temporada de crecimiento con un 56% de sobreexpresión de proteínas, de la categoría funcional “rescate celular, defensa y virulencia” en madera tardía, en comparación con sólo un 4% en madera temprana. Esta evidencia Proteómica

aporta sustento experimental a la hipótesis de Gion et al.( 2005), sobre la importancia de mecanismos de defensa celular en la extensión del depósito de pared celular.

Se ha identificado la fermentación etanólica como un proceso de respaldo al metabolismo celular en condiciones anaerobias, por el mayor grosor de las paredes celulares se espera mayor hipoxia en madera tardía. Es un proceso que ocurre en dos etapas, en la cual piruvato es inicialmente decarboxilado a acetaldehído, por la piruvato decarboxilasa y el acetaldehído es subsecuentemente convertido metanol por la alcohol deshidrogenasa regenerando NAD+. Lactato y Etanol son producidos en un grado variable en la mayoría de las plantas , bajo condiciones de estrés (Tadege and Kuhlemeier, 1997).

En madera temprana, las categoría mas importante fueron “metabolismo” 26%, con proteínas que participan en la formación de la placa celular, reacciones de transferencia de metil activado, biosíntesis de lípidos, detoxificación de hipoxia, ”biogénesis de componentes celulares” 18%, con una sobreexpresión de proteínas relacionadas con citoesqueleto, proteínas de pared celular, trascrición 12% con sobreexpresión de proteínas del metabolismo de RNA y “destino de proteínas” con proteínas inhibidoras de la acción de quinasas y proteasas. Todas las categorías sobreexpresadas son consecuentes con la activa división celular que ocurre en madera temprana.

#### **4.2.6 Análisis proteómico y transcriptómico de un gradiente base a copa y un gradiente de temporada en tejido formador de madera en pino marítimo. Temporada 2006**

Los resultados de estos dos gradientes, analizados el año 2006, se incorporaron al artículo “A Combined Proteome and Transcriptome analysis in a seasonal and base to crown gradient of wood forming tissue in maritime pine”. Marcelo Garcés, Phillipe Chaumeil, Grégoire Le Provost, Céline Lalanne, Jose Carlos Rodrigues, Ana Alves, Stéphane Claverol, Jorge Paiva, Alexander Bosc, Raul Herrera and Christophe Plomion \*

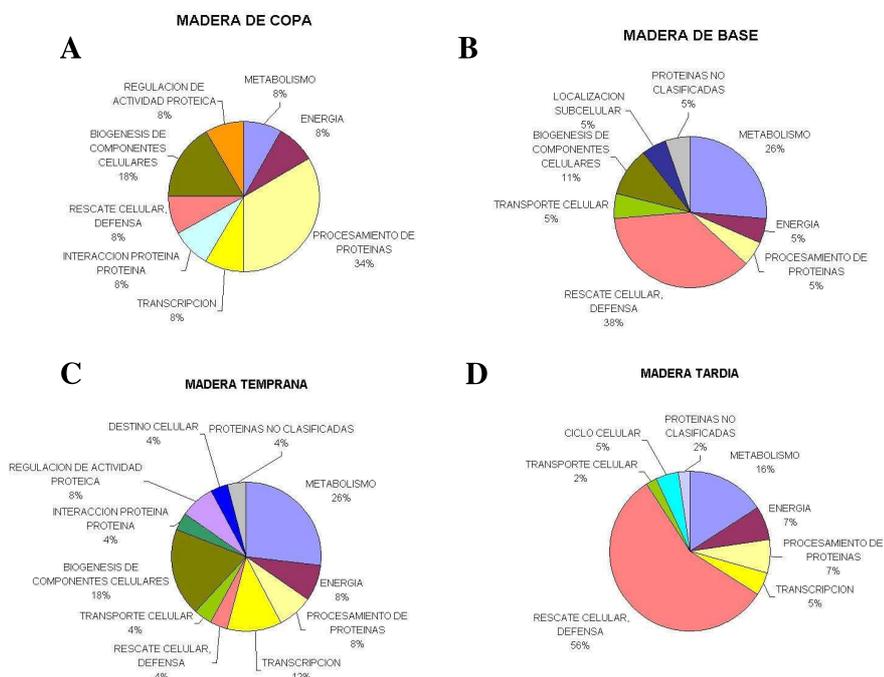
para Molecular & Cell Proteomics

### 4.3 Discusión Final

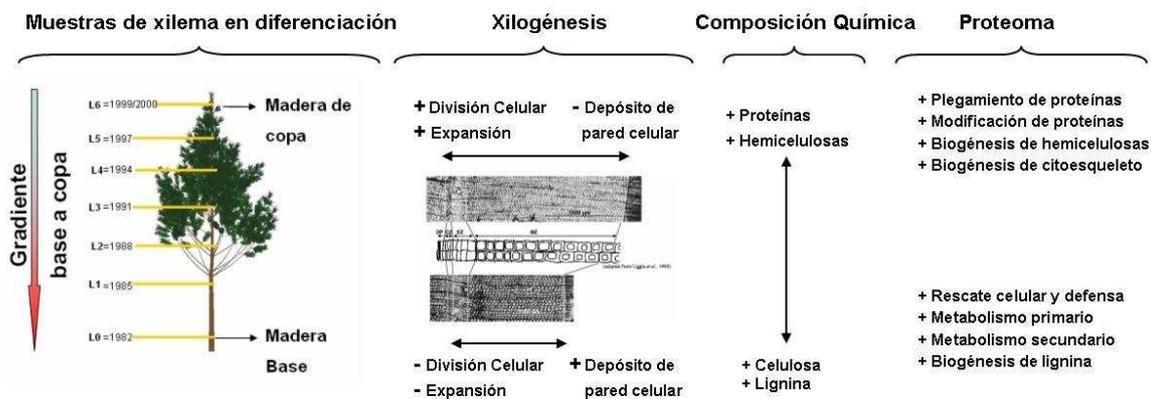
Este trabajo de tesis abordó la variabilidad observada en distintos tipos de madera analizando la sobreexpresión de proteínas que pudieran ser responsables de esta variabilidad y por lo tanto de las propiedades finales de la madera.

Se analizaron dos gradientes de tejido formador de madera: (i) un gradiente de edad cambial, desde madera de base a madera de copa, en las temporadas 2003, 2005 y 2006 (ii) un gradiente de temporada el 2006, desde madera temprana a madera tardía. Muestras recolectadas a través de estos gradientes fueron analizadas por métodos químicos, proteómicos y transcriptómicos. Los datos obtenidos para la temporada de crecimiento 2006, además fueron analizados tomando en cuenta las variables ecofisiológicas de la parcela. El 2006 se trato de fijar la variación genética entre clones, estudiando dos clones para cada gradiente. Se identificó una remodelación del proteoma, entre madera de copa y madera de base y desde madera temprana a madera tardía. La **Figura 52** resume la remodelación del proteoma entre los gradientes analizados.

En la madera de copa (**Figura 52A**), las categorías funcionales sobreexpresadas más



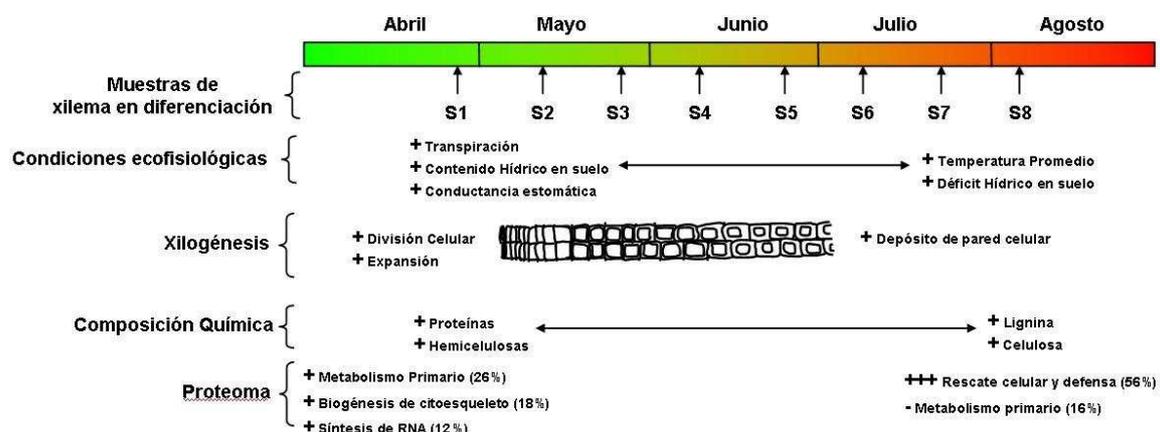
**Figura 52:** Remodelación del proteoma en dos gradientes (i) madera de copa y madera de base. (ii) madera temprana y madera tardía



**Figura 53:** Resumen de los principales resultados obtenidos en el gradiente base a copa (Figura modificada de Paiva et al, 2008)

representativas fueron: “procesamiento de proteínas” (34%) y “biogénesis de compuestos celulares” (18%). En la madera de base (Figura 52B), se sobrepresaron las categorías funcionales “rescate celular y defensa” (38%) y “metabolismo” (26%). La comparación de los resultados proteómicos con los datos morfológicos y químicos de la pared celular para el gradiente base a copa es resumida en la Figura 53.

Muestras representativas de madera de base y copa fueron caracterizadas morfológicamente, confirmando las características ya descritas (Cato et al., 2006), con una mayor división celular y expansión en la madera de copa y un extendido depósito de pared celular en madera de base.

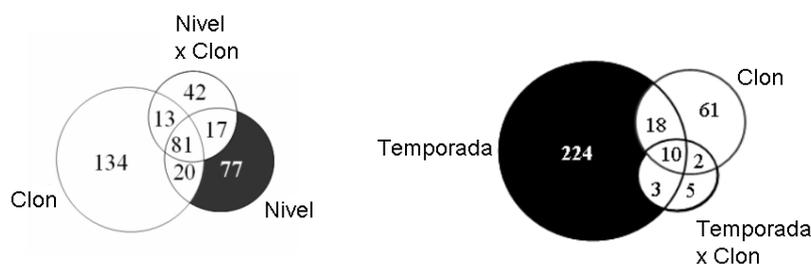


**Figura 54:** Resumen de los principales resultados obtenidos en el gradiente temporada 2006

Los análisis químicos de constituyentes de pared celular revelaron un mayor contenido de proteínas y hemicelulosas en madera de copa y una mayor proporción de Celulosa y Lignina en madera de base. En madera de copa, se sobreexpresaron proteínas involucradas en Plegamiento de proteínas, Modificación de proteínas, Biogénesis de hemicelulosas y Biogénesis de citoesqueleto, la sobreexpresión de todas estas vías metabólicas celulares es consecuente con las características antes descritas de intensa división celular en madera de copa. Por otro lado, en madera de base se identificó la sobreexpresión de proteínas involucradas en “Rescate celular y defensa”, “Metabolismo primario, secundario” y “Biogénesis de Lignina”, la sobreexpresión de estas categorías, soporta la hipótesis que en madera de base el depósito de pared celular se extiende gracias a mecanismos protectores de la célula (Gion et al., 2005), (por ejemplo por HSP identificadas en este estudio). La evasión de la muerte celular programada permite la acumulación de pared celular, probablemente mediante proteínas involucradas en biosíntesis de lignina.

Para el gradiente de temporada de crecimiento 2006, desde madera temprana a madera tardía, se evidenció una intensa remodelación del proteoma expresado, las categorías funcionales más abundantes en madera temprana (**Figura 52C**) fueron: “Metabolismo” (26%), “Biogénesis de compuestos celulares” (18%) y “Transcripción” (12%), por otro lado las categorías funcionales más abundantes para madera tardía (**Figura 52D**) fueron: “Rescate celular y defensa” (56%) y “metabolismo” (16%). La comparación de los resultados proteómicos con los datos ecofisiológicos y químicos de la pared celular son resumidos en la **Figura 54**.

Los análisis proteómicos mostraron claramente una mayor importancia del efecto clonal vs. efecto nivel para el gradiente base a copa, comparativamente, el efecto clonal tiene menor



importancia en el gradiente de temporada (**Figura 55**).

**Figura 55:** Diagramas de Venn, A) de 384 proteínas significativas del gradiente base a copa 248 son influenciadas por el efecto clonal (65%). B) de 323 proteínas significativas del gradiente temporada 91 son influenciadas por efecto clonal (28%).

### **Optimizaciones técnicas introducidas**

Durante el desarrollo de este trabajo de tesis se introdujeron optimizaciones técnicas al análisis proteómico, como testear las cantidades mínimas de proteínas detectables, e inyectar las muestras en orden de intensidad para evadir la interferencia de proteínas más intensas. Se utilizó espectrometría de masas ESI LC MS/MS, para luego buscar homología en la base de datos de ESTs de pino, como resultado se obtuvo un muy alto porcentaje de identificación (95%), comparado con un 67,9% reportado por Gion et al.(2005).

Usualmente luego de la electroforesis bidimensional los geles eran digitalizados y descartados, preparándose geles frescos de selección de spots para ESI MS/MS, se introdujo la modificación de almacenar los geles digitalizados en bolsas plásticas selladas con Ac. Acético 5% a 4°C, hasta la remoción manual de los spots seleccionados, luego de un lavado con Ac. Acético 5%, se cortaron manualmente los spots y se realizó la ESI MS/MS, no se detectaron diferencias producidas por el periodo de almacenaje de los geles, esto permitió ahorrar material recolectado, material y tiempo. El uso del software de nueva generación para análisis de imágenes “SameSpot”, también fue una modificación relevante que ayudó con una importante reducción del tiempo de análisis de imágenes. En este sentido, se redujo el tiempo de análisis para un experimento utilizando “ImageMaster 2D-Platinum” de aproximadamente 8 semanas a sólo 2 semanas con “SameSpot”.

Estudios futuros podrán identificar genes específicos involucrados en las categorías funcionales identificadas, para realizar estudios de poblaciones e identificar diferencias entre genotipos en etapas tempranas de desarrollo.

## 5.0 Conclusiones

En este estudio se han analizado las proteínas y algunos genes involucrados en la variación de las propiedades de madera en un gradiente de edad cambial y un gradiente de temporada;

Se ha utilizado un enfoque global, que incluyó análisis ecofisiológicos para analizar variables ambientales, el uso de replicas biológicas para reducir el efecto genotipo, estudios químicos para verificar el fenotipo presente en los tipos de madera analizados ,y proteómico, para identificar las proteínas sobreexpresadas.

La aproximación Proteómica confirma resultados obtenidos previamente por transcriptómica, en términos de presentar cambios en la expresión de las mismas categorías funcionales.

En el gradiente de edad cambial, en madera de copa se identificaron proteínas que participan en el plegamiento y modificación de proteínas, biogénesis de hemicelulosas y de citoesqueleto, características que soportan la intensa división celular observada (Figura 51).

En madera de base ocurre una menor división celular y un aumento de depósito de pared celular, hay una sobreexpresión de proteínas de defensa celular y al mismo tiempo una sobreexpresión de proteínas del metabolismo primario y secundario, así como proteínas del metabolismo de lignina.

En el gradiente de temporada, desde madera temprana a madera tardía, (Figura 52) en madera temprana hay una mayor actividad de división celular, por las condiciones favorables de humedad y de fotoperíodo, que inducen mayor actividad el metabolismo primario, biogénesis de citoesqueleto y síntesis de RNA.

El efecto genotipo fue mas importante para el gradiente de edad cambial que para el gradiente de temporada.

Las proteínas identificadas como sobreexpresadas en este estudio, pueden ser utilizadas como genes candidatos expresionales. Estudios poblacionales verificando polimorfismos asociados a estos genes candidatos han sido iniciados.

## 6. Referencias Bibliograficas

### A

- ABE, H., FUNADA, R., OHTANI, J. & FUKAZAKA, K. (1997) Challenges in the arrangement of cellulose microfibrils associated with the cessation of cell expansion in tracheids. *Trees*, 11, 328-332.
- AGRAWAL, S., A. KUMAR, V. SRIVASTAVA, AND B. N. MISHRA. 2003. Cloning, expression, activity and folding studies of serine hydroxymethyltransferase: a target enzyme for cancer chemotherapy. *J. Mol. Microbiol. Biotechnol.* 6: 67-75.
- ALLONA, I., QUINN, M., SHOOP, E., SWOPE, K., STCYR, S., CARLIS, J., RIEDL, J., RETZEL, E., CAMPBELL, M. M. & SEDEROFF, R. (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proc Natl Acad Sci US A*, 95, 9693-9698.
- ALONI, R. (1987) Differentiation of Vascular Tissues. *Annual Review of Plant Physiology*, 34, 179-204.
- AVILA, C., GARCÍA-GUTIÉRREZ, A., CRESPILO, R. & CÁNOVAS, F. M. (1998) Effects of phosphinotricin treatment on glutamine synthetase isoforms in Scots pine seedlings. *Plant Physiology et Biochemistry*, 36, 857-863.
- ÁVILA SÁEZ, C., MUÑOZ-CHAPULI, R., PLOMION, C., FRIGERIO, J. M. & CÁNOVAS, F. M. (2000) Two genes encoding distinct cytosolic glutamine synthetases are closely linked in the pine genome. *FEBS Letters*, 477, 237-243.
- ARRIGO, A. P. (2005) In search of the molecular mechanism by which small stress proteins counteract apoptosis during cellular differentiation. *Journal of Cellular Biochemistry*, 94, 241-246.

### B

- BAHRMAN, N. & PETIT, R. J. (1995) Genetic polymorphism in maritime pine (*Pinus pinaster* Ait.) assessed by two-dimensional gel electrophoresis of needle, bud, and pollen proteins. *Journal of Molecular Evolution*, 41, 231-237.

- BARADAT, P. & MARPEAU-BEZARD, A. (1988) Le pin maritime, *Pinus pinaster* Ait. Biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce.
- BOERJAN, W. (2005) Biotechnology and the domestication of forest trees. *Current Opinion in Biotechnology*, 16, 159-166.
- BROEKAERT, W. V., TERRAS, V. R. G., CAMMUE, B. P. A. & OSBORN, R. W. (1995) Plant defensins: novel antimicrobial peptides as components of the host defense system. *Plant Physiol*, 108, 1353-1358.

## C

- CANOVAS, F. M., DUMAS-GAUDOT, E., RECORBET, G., JORRIN, J., MOCK, H. P. & ROSSIGNOL, M. (2004) Plant proteome analysis. *Proteomics*, 4, 285-98.
- CARPENTIER, S. C., WITTERS, E., LAUKENS, K., DECKERS, P., SWENNEN, R. & PANIS, B. (2005) Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics*, 5.
- CATO, S., MCMILLAN, L., DONALSON, L., RICHARDSON, T., ECHT, C. & GARDNER, R. (2006) Wood formation from the base to the crown in *Pinus radiata*: gradients of tracheid wall thickness, wood density, radial growth rate and gene expression. *Plant Molecular Biology*, 60, 565-581.
- CERVERA, M. T., STORME, V., IVENS, B., GUSMAO, J., LIU, B. H., HOSTYN, V., VAN SLYCKEN, J., VAN MONTAGU, M. & BOERJAN, W. (2001) Dense Genetic Linkage Maps of Three *Populus* Species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) Based on AFLP and Microsatellite Markers. *Genetics*, 158, 787-809.
- CHAGNÉ, D., BROWN, G., LALANNE, C., MADUR, D., POT, D., NEALE, D. & PLOMION, C. (2003) Comparative genome and QTL mapping between maritime and loblolly pines. *Molecular Breeding*, 12, 185-195.
- CHAGNÉ, D., CHAUMEIL, P., RAMBOER, A., COLLADA, C., GUEVARA, A., CERVERA, M. T., VENDRAMIN, G. G., GARCIA, V., FRIGERIO, J. M. & ECHT, C. (2004) Cross-species transferability and mapping of genomic and cDNA SSRs in pines. *TAG Theoretical and Applied Genetics*, 109, 1204-1214.

- CHAUMEIL, P. (2007). Comunicación Personal.
- CORNELIUS, J. (1994) Heritabilities and additive genetic coefficients of variation in forest trees. *Can. J. For. Res./Rev. can. rech. for.*, 24, 372-379.
- COSTA, P., BAHRMAN, N., FRIGERIO, J. M., KREMER, A. & PLOMION, C. (1998) Water-deficit-responsive proteins in maritime pine. *Plant Molecular Biology*, 38, 587-596.
- COSTA, P., POT, D., DUBOS, C., FRIGERIO, J. M., PIONNEAU, C., BODENES, C., BERTOCCHI, E., CERVERA, M. T., REMINGTON, D. L. & PLOMION, C. (2000) A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *TAG Theoretical and Applied Genetics*, 100, 39-48
- CATO, S., MCMILLAN, L., DONALSON, L., RICHARDSON, T., ECHT, C. & GARDNER, R. (2006) Wood formation from the base to the crown in *Pinus radiata*: gradients of tracheid wall thickness, wood density, radial growth rate and gene expression. *Plant Molecular Biology*, 60, 565-581.

## D

- DANTEC, L. L., CHAGNÉ, D., POT, D., CANTIN, O., GARNIER-GÉRE, P., BEDON, F., FRIGERIO, J. M., CHAUMEIL, P., LÉGER, P. & GARCIA, V. (2004) Automated SNP Detection in Expressed Sequence Tags: Statistical Considerations and Application to Maritime Pine Sequences. *Plant Molecular Biology*, 54, 461-470.
- DODD, R. S. & FOX, P. (1990) Kinetics of Tracheid Differentiation in Douglas-fir. *Annals of Botany*, 65, 649.
- DUMAIL, J. F., CASTERA, P. & MORLIER, P. (1998) Hardness and basic density variation in the juvenile wood of maritime pine. *Annales des sciences forestières*, 55, 911-923
- DURE, L., CROUCH, M., HARADA, J., HO, T. H. D., MUNDY, J., QUATRANO, R., THOMAS, T. & SUNG, Z. R. (1989) Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology*, 12, 475-486.
- DOLFERUS, R., DE BRUXELLES, G., DENNIS, E. S. & PEACOCK, W. J. (1994)

Regulation of the Arabidopsis Adh Gene by Anaerobic and Other Environmental Stresses. *Annals of Botany*, 74, 301-308.

DOXEY, A. C., YAISH, M. W. F., MOFFATT, B. A., GRIFFITH, M. & MCCONKEY, B. J. (2007) Functional Divergence in the Arabidopsis {beta}-1, 3-Glucanase Gene Family Inferred by Phylogenetic Reconstruction of Expression States. *Molecular Biology and Evolution*, 24, 1045.

## **E**

EGERTSDOTTER, U., VAN ZYL, L. M., MACKAY, J., PETER, G., KIRST, M., CLARK, C., WHETTEN, R. & SEDEROFF, R. (2004) Gene Expression during Formation of Earlywood and latewood in Loblolly Pine: Expression Profiles of 350 Genes. *Plant Biology*, 6, 654-663.

EKRAMODDOULLAH, A. K. M. & TAN, Y. (1998) Differential accumulation of proteins in resistant and susceptible sugar pine(*pinus lambertiana*) seedlings inoculated with the white pine blister rust fungus(*cronartium ribicola*). *Canadian journal of plant pathology*, 20, 308-318.

ERIKSSON, M. E., ISRAELSSON, M., OLSSON, O. & MORITZ, T. (2000) Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nature biotechnology*, 18, 784-788.

EVENO, E., COLLADA, C., GUEVARA, M., LEGER, V., SOTO, A., DIAZ, L., LEGER, P., GONZALEZ-MARTINEZ, S. C., CERVERA, M. T. & PLOMION, C. (2007) " Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses". *Molecular Biology and Evolution*

## **F**

FERRY-DUMAZET, H., HOUEL, G., MONTALENT, P., MOREAU, L., LANGELLA, O., NEGRONI, L., VINCENT, D., LALANNE, C., DE DARUVAR, A. & PLOMION, C. (2005) PROTICdb: a web-based application to store, track, query, and compare plant proteome data. *Proteomics*, 5, 2069-81.

- FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. (1990) Electrospray ionization-principles and practice. *Mass Spectrom. Rev*, 9, 37.
- FAO (2004) Preliminary review of biotechnology in forestry, including genetic modification. IN FGR/59E, F. G. R. W. P. (Ed. Rome, Italy., Forest Resources Development Service, Forest Resources Division.
- FIORANI CELEDON, P. A., DE ANDRADE, A., XAVIER MEIRELES, K. G., GALLO DE GAVARDO, M. C. D. C. G., GOMES CALDAS, D. G., HENRY MOON, D., TOZELLI CARNEIRO, R., FRANCESCHINI, L. M., ODA, S. & ALBERTO LABATE, C. (2007) Proteomic analysis of the cambial region in juvenile *Eucalyptus grandis* at three ages. *Proteomics*, 7, 2258-2274.
- FUKUDA, H. (1996) XYLOGENESIS: INITIATION, PROGRESSION, AND CELL DEATH. *Annual Review of Plant Physiology and Plant Molecular Biology*, 47, 299-325.
- FLORACK, D. E. A. & STIEKEMA, W. J. (1994) Thionins: properties, possible biological roles and mechanisms of action. *Plant Molecular Biology*, 26, 25-37.

## G

- GARAY-ARROYO, A., COLMENERO-FLORES, J. M., GARCIARRUBIO, A. & COVARRUBIAS, A. A. (2000) Highly Hydrophilic Proteins in Prokaryotes and Eukaryotes Are Common during Conditions of Water Deficit. *Journal of Biological Chemistry*, 275, 5668-5674.
- GION, J.-M., LALANNE, C., LE PROVOST, G., FERRY-DUMAZET, H., PAIVA, J., CHAUMEIL, P., FRIGERIO, J.-M., BRACH, J., AURÉLIEN, B., DE DARUVAR, A., CLAVEROL, S., MARC, B., SOMMERER, N., NEGRONI, L., PLOMION, C. & (2005) The proteome of maritime pine wood forming tissue. *Proteomics*, 5, 3731-3751.
- GREGORY, R. A. (1971) Cambial Activity in Alaskan White Spruce. *American Journal of Botany*, 58, 160-171.
- GROOME, M. C., AXLER, S. R. & GIFFORD, D. J. (1991) Hydrolysis of lipid and protein

reserves in loblolly pine seeds in relation to protein electrophoretic patterns following imbibition. *Physiologia Plantarum*, 83, 99-106.

GROOVER, A. T. (2005) What genes make a tree a tree? *trends in plant science*, 10, 210-214.

GYGI, S. P., ROCHON, Y., FRANZA, B. R. & AEBERSOLD, R. (1999) Correlation between Protein and mRNA Abundance in Yeast. *Molecular and Cellular Biology*, 19, 1720.

## H

HAN, K. H. (2001) Molecular Biology of Secondary Growth. *Journal of Plant Biotechnology*, 3, 45-57.

HARADA, H. & CÔTÉ, W. (1985) Structure of wood. IN HIGUCHI, T. (Ed.) *Biosynthesis and Biodegradation of Wood Components*. Orlando, FL., Academic Press.

HELLGREN, J. M. (2003) *Ethylene and auxin in the control of wood formation*, Dept. of Forest Genetics and Plant Physiology, Swedish Univ. of Agricultural Sciences.

HIGUCHI, T. (1996) *Biochemistry and Molecular Biology of Wood*, Springer-Verlag

## J

JONES, A. M. (2001) Programmed cell death in development and defense. *Plant Physiol*, 125, 94-97.

JORRIN, J. V., MALDONADO, A. M. & CASTILLEJO, M. A. (2007) Plant proteome analysis: a 2006 update. *Proteomics*, 7, 2947-62

JOSEPH, G. & KELSEY, R. G. (1997) Ethanol Synthesis and Water Relations of Flooded *Pseudotsuga menziesii* (Mirb.) Franco (Douglas-Fir) Seedlings under Controlled Conditions. *International Journal of Plant Sciences*, 158, 844.

JOSEPH, G. & KELSEY, R. G. (2004) Ethanol synthesis and aerobic respiration in the laboratory by leader segments of Douglas-fir seedlings from winter and spring. *Journal of Experimental Botany*, 55, 1095-1103.

## K

KIRCH, H. H., BARTELS, D., WEI, Y., SCHNABLE, P. S. & WOOD, A. J. (2004) The ALDH gene superfamily of Arabidopsis. *trends in plant science*, 9, 371-377.

KLECZKOWSKI, L. A., GEISLER, M., CIERESZKO, I. & JOHANSSON, H. (2004) UDP-Glucose Pyrophosphorylase. An Old Protein with New Tricks<sup>1</sup>. *Plant Physiology*, 134, 912–918

## L

LARSON, P. R., KRETSCHMANN, D. E., CLARCK, A. & ISEBRANDS, J. G. (2001) Formation and properties of juvenile wood in southern pines: a synopsis. *Gen. Tech. Rep.* Madison, WI:, U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.

LE PROVOST, G., PAIVA, J., POT, D., BRACH, J. & PLOMION, C. (2003) Seasonal variation in transcript accumulation in wood-forming tissues of maritime pine (*Pinus pinaster* Ait.) with emphasis on a cell wall glycine-rich protein. *Planta*, 217, 820-830.

LEE, R. C., HRMOVA, M., BURTON, R. A., LAHNSTEIN, J. & FINCHER, G. B. (2003) Bifunctional Family 3 Glycoside Hydrolases from Barley with  $\alpha$ -l-Arabinofuranosidase and  $\beta$ -d-Xylosidase Activity CHARACTERIZATION, PRIMARY STRUCTURES, AND COOH-TERMINAL PROCESSING. *Journal of Biological Chemistry*, 278, 5377-5387.

LOEWUS, V. A., KELLY, S. & NEUFELD, E. V. (1962) Metabolism of Myo-inositol in Plants: Conversion to Pectin, Hemicellulose, D-xylose, and Sugar Acids. *Proceedings of the National Academy of Sciences of the United States of America*, 48, 421-425.

LOEWUS, V. A. & MURTHY, P. P. N. (2000) myo-Inositol metabolism in plants. *Plant science(Limerick)*, 150, 1-19.

## M

MAEGLIN, R. (1987) JUVENILE WOOD, TENSION WOOD, AND GROWTH STRESS EFFECTS ON PROCESSING HARDWOODS. IN COUNCIL;, H. R. (Ed.) *Applying the latest research to hardwood problems*:. Memphis, TN.

## P

- PAIVA, J. (2006) PHENOTYPIC and MOLECULAR PLASTICITY of WOOD FORMING TISSUES in MARITIME PINE (*Pinus pinaster*Ait.). Bordeaux, Lisbon, Bordeaux 1 University New University of Lisbon.
- PAIVA, J., GARCES, M., ALVES, A., GARNIER-GÉRÉ, P., RODRIGUES, J.-C., LALANNE, C., PORCON, S., LE PROVOST, G., DA SILVA, D., BRACH, J., FRIGERIO, J. M., CLAVEROL, S., BARRE, A., FEVEREIRO, P. & PLOMION, C. (2008) Molecular and phenotypic profiling from the base to the crown in maritime pine wood forming tissue. *New Phytologist*.
- PECK, S. C. (2005) Update on proteomics in Arabidopsis. Where do we go from here? *Plant Physiol*, 138, 591-9.
- PERRY, D. J. & FURNIER, G. R. (1996) *Pinus banksiana* has at least seven expressed alcohol dehydrogenase genes in two linked groups. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 13020.
- PFLIEGER, S., LEFEBVRE, V. & CAUSSE, M. (2001) The candidate gene approach in plant genetics: a review. *Molecular Breeding*, 7, 275-291.
- PLOMION, C., BAHRMAN, N., DUREL, C. E. & O'MALLEY, D. M. (1995) Genomic mapping in *Pinus pinaster*(maritime pine) using RAPD and protein markers. *Heredity*, 74, 661-668.
- PLOMION, C., LALANNE, C., CLAVEROL, S., MEDDOUR, H., KOHLER, A., BOGEAT-TRIBOULOT, M. B., BARRE, A., LE PROVOST, G., DUMAZET, H. & JACOB, D. (2006) Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics*, 6, 6509-27.
- PLOMION, C., LEPROVOST, G. & STOKES, A. (2001) Wood formation in trees. *Plant Physiol*, 127, 1513-1523.
- PLOMION, C., PIONNEAU, C., BRACH, J., COSTA, P. & BAILLÈRES, H. (2000) Compression Wood-Responsive Proteins in Developing Xylem of Maritime Pine (*Pinus pinaster* Ait.). *Plant Physiology*, 123, 959.
- PLOMION, C., RICHARDSON, T. & MACKAY, J. (2005) Advances in forest tree genomics. *New Phytologist*, 166, 713-717.

- POT, D. (2004) Déterminisme génétique de la qualité du bois chez le pin maritime : du phénotype aux gènes. *Ecole Doctorale : Vie - Agronomie - Santé*. Rennes ENSA de Rennes
- POT, D., CHANTRE, G., ROZENBERG, P., RODRIGUES, J.-C., JONES, G. L., PEREIRA, H., HANNRUP, B., CAHALAN, C. & PLOMION, C. (2002) Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). *Ann. For. Sci.*, 59, 563-575.
- POT, D., MCMILLAN, L., ETCH, C., LE PROVOST, G., GARNIER-GÉRÉ, P., CATO, S. & PLOMION, C. (2005) Nucleotide variation in genes involved in wood formation in two pine species. *New Phytologist*, 167, 101-167.
- PLA, M., HUGUET, G., VERDAGUER, D., PUIGDERRAJOLS, P., LLOMPART, B., NADAL, A. & MOLINAS, M. (1998) Stress proteins co-expressed in suberized and lignified cells and in apical meristems. *Plant Science*, 139, 49-57.
- PUIGDERRAJOLS, P., FERNANDEZ-GUIJARRO, B., TORIBIO, M. & MOLINAS, M. (1996) Origin and Early Development of Secondary Embryos in *Quercus suber* L. *International Journal of Plant Sciences*, 157, 674.

## R

- RAMANJULU, S. & BARTELS, D. (2002) Drought- and desiccation-induced modulation of gene expression in plants. *Plant, Cell and Environment* 25, 141–151.

## S

- SAMUELS, A. L. (1995) Cytokinesis in tobacco BY-2 and root tip cells: a new model of cell plate formation in higher plants. *The Journal of Cell Biology*, 130, 1345-1357.
- SEITZ, B., KLOS, C., WURM, M. & TENHAKEN, R. (2000) Matrix polysaccharide precursors in Arabidopsis cell walls are synthesized by alternate pathways with organ-specific expression patterns. *The Plant Journal*, 21, 537-546.
- SCHIRCH, V., AND D. M. SZEBENYI. 2005. Serine hydroxymethyltransferase 399 revisited. *Curr. Opin. Chem. Biol.* 9: 482-487.

SCHERP, P., GROTHA, R. & KUTSCHERA, U. (2001) Occurrence and phylogenetic significance of cytokinesis-related callose in green algae, bryophytes, ferns and seed plants. *Plant cell reports(Print)*, 20, 143-149.

## T

TADEGE, M. & KUHLEMEIER, C. (1997) Aerobic fermentation during tobacco pollen development. *Plant Molecular Biology*, 35, 343-354.

TSUJI, H., MEGURO, N., SUZUKI, Y., TSUTSUMI, N., HIRAI, A. & NAKAZONO, M. (2003) Induction of mitochondrial aldehyde dehydrogenase by submergence facilitates oxidation of acetaldehyde during re-aeration in rice. *FEBS Letters*, 546, 369-373.

## U

UGGLA, C., ELISABETH MAGEL, THOMAS MORITZ & SUNDBERG, B. (2001) Function and Dynamics of Auxin and Carbohydrates during Earlywood/Latewood Transition in Scots Pine. *Plant Physiology*, 125, 2029–2039.

## Z

ZOBEL, B. J. & SPRAGUE, J. R. (1998) *Juvenile wood in forest trees*, Berlin.

ZOBEL, B. J. & VAN BUIJTENEN, J. P. (1989) *Wood Variation: Its Causes and control*, Berlin, Springer-verlag.

## ANEXOS

Anexo 1: Extracción de proteínas totales desde xilema en formación	i
Anexo 2: Cuantificación proteínas por el método de Bradford modificado	iv
Anexo 3: Electro-Focalización con el sistema IPGphor: migración según el pI	vii
Anexo 4: Electroforesis denaturante en gel de poliacrilamida	viii
Anexo 5: Coloración de geles con azul coloidal	ix
Anexo 6: Modelos Estadísticos	x
Anexo 7: Programas R utilizados en esta tesis	xiv
Anexo 8 : poster « Effect of cambial age in protein accumulation in maritime pine wood forming tissue »	

## Anexo 1: Extracción de proteínas totales

### 1/Principio del método

Extracción de proteínas totales en condiciones denaturantes

### 2/ Materiales necesarios

Balanza de precisión

Morteros

Centrifuga refrigerada

Vortex

Pipeta plástica estéril de 10ml.

Campana para secar pellets al vacío

Tubos “oak-ridge” de 10ml. con tapa

Tubos Eppendorf de 1,5 o 2ml.

Tubos Falcon de 15 o 50ml.

Botellas de vidrio de 200ml.

Bombona con nitrógeno líquido

Juego de pipetas (P2, P20, P200, P1000) con sus correspondientes puntas.

### 3/Reactivos

TCA	WR Prolabo RP Normapur 20 742.293
2-mercaptoetanol	Sigma M6250
Acetona	VWR Prolabo Rectapur 20 065.362
DTT	Euromedex EU0006-D
Triton X100	Sigma T8787
CHAPS	Sigma C30023
Tiourea	Sigma T7875
Urea	Euromedex EU0014-B
IPG Buffer pH 4-7	Amersham Biosciences 17-6000-86

#### 4/Modo de operación

Antes de comenzar, verificar la presencia de Nitrógeno líquido y encender la centrifuga refrigerada, para que esta alcance una temperatura ligeramente inferior a 0°C

- 1) Preparar las soluciones de precipitación y de lavado y guardar a -20°C
- 2) Peso de tubos “oak-ridge” para la extracción (sin las tapas)
- 3) Retiro del tejido desde congelador a -80°C
- 4) Pulverizar el material vegetal en un mortero, con nitrógeno líquido: 500 a 700mg
- 5) Adjuntar a la muestra pulverizada, 10ml. de solución de precipitación guardada a -20°C (8 ml para recuperación del material y 2ml. para enjuague del mortero)
- 6) Agregar el líquido en un tubo oak-ridge de 10 ml. y dejar reposar una hora a -20°C
- 7) Centrifugar 15min a -20°C a 16000 g
- 8) Descartar el sobrenadante
- 9) Lavar los pellets con 10ml. de la solución de lavado
- 10) Dejar reposar una hora a -20°C y luego centrifugar 10min a 16000 g
- 11) Eliminar el sobrenadante
- 12) Secar el pellet en desecador por 2 a 3 horas
- 13) Pesar los tubos con los pellets (sin las tapas)
- 14) Pulverizar los pellets con la ayuda de una bagueta de vidrio
- 15) Resuspender los pellets en un volumen de 15 a 30ml. de tampón de resolubilización por mg. de pellet.
- 16) Centrifugar los tubos a 400g por 2min. a 20°C. Extraer el sobrenadante y guardar en un Tubo Eppendorf. Reagrupar los sobrenadantes si se han hecho varias repeticiones. Centrifugar los tubos nuevamente a 400g por 2 minutos a 20°C.
- 17) Recuperar los sobrenadantes y guardarlos a -80°C

#### 5/Preparación de Soluciones:

<u>Solución de precipitación</u>	<u>10 muestras</u>
T.C.A.	10g
2-mercaptoetanol	70µl
Acetona	c.s. 100ml

### Solucion de lavado

2-mercaptoetanol		70µl
Acetona	c.s.	100ml

### Tampón de resolubilizacion de proteínas

Urea	21g
Triton X100	1ml
Chaps	2g
Tiourea	7,61g
DTT	77,12mg
Buffer IPG	0,5ml
Agua miliQ	c.s.50ml

Nota: Comenzar temprano la preparación del tampón de resolubilización, este requiere bastante tiempo de agitación, por la alta concentración de urea, guardar esta solución a -20°C.

## **6/Referencias**

- Damerval C, de Vienne D, Zivy M, Thiellement H (1986) Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *Electrophoresis* 7: 52-54
- 2-D Electrophoresis using immobilized pH gradient: principles and Methods . Amersham Pharmacia Biotech (1998)
- Mode opératoire : Extraction des protéines totales, Rédacteur Céline Lalanne Assistant Ingénieur. INRA UMR Biogeco (2004)

## Anexo 2: Cuantificación proteínas por el método de Bradford modificado

### 1/Principio del método

Cuantificación de proteínas a continuación de la extracción descrita en “Extracción de proteínas totales”

### 2/ Materiales necesarios

Cubetas estándar desechables de 4,5ml

Espectrofotómetro (SPECTRA max Plus de Molecular Devices)

Balanza de precisión

Botellas Shott de vidrio

Papel Wattman n°1

Embudo de vidrio

Pipeta plástica estéril de 10 ml.

Juego de pipetas (P2, P20,P200 ,P1000) con sus correspondientes puntas

### 3/Modo de operación

Encender el Espectrofotómetro 15 minutos antes de su utilización

Junto antes de la cuantificación, descongelar las muestras, mezclar suavemente y centrifugar los tubos a 400g, por 4 minutos a 20°C. Si hay pellet luego de la centrifugación, recuperar el sobrenadante en un nuevo tubo y recentrifugar.

### Preparación de la curva de calibración

Diluir el reactivo colorante concentrado “Dye reagent concentrate”(BIORAD): 1 volumen de reactivo con 3 volúmenes de agua mili Q. Filtrar la solución en una botella, con papel Wattman n°1

Preparar una solución de ovoalbumina a 5mg/ml. Para esto pesar 10mg de ovoalbumina y agregar 2ml. de tampón de resolubilización.

Preparación de la curva de calibración directamente en las cubetas

<b>Cantidad de proteínas (µg)</b>	<b>0</b>	<b>5</b>	<b>10</b>	<b>20</b>	<b>30</b>	<b>40</b>	<b>50</b>
Solución ovoalbumina (µl)	-	1	2	4	6	8	10
Tampón de resuspensión (µl)	10	9	8	6	4	2	-
HCl 0,1N (µl)	10	10	10	10	10	10	10
H <sub>2</sub> O (µl)	80	80	80	80	80	80	80
Reactivo tinción diluido	3,5	3,5	3,5	3,5	3,5	3,5	3,5

No olvidar mezclar las cubetas luego de agregar el reactivo de tinción. Esperar 5 minutos antes de la cuantificación

Preparación de las muestras en las cubetas

Volumen de muestra (µl)	10
HCl 0,1N (µl)	10
H <sub>2</sub> O (µl)	80
Reactivo tinción diluido	3,5

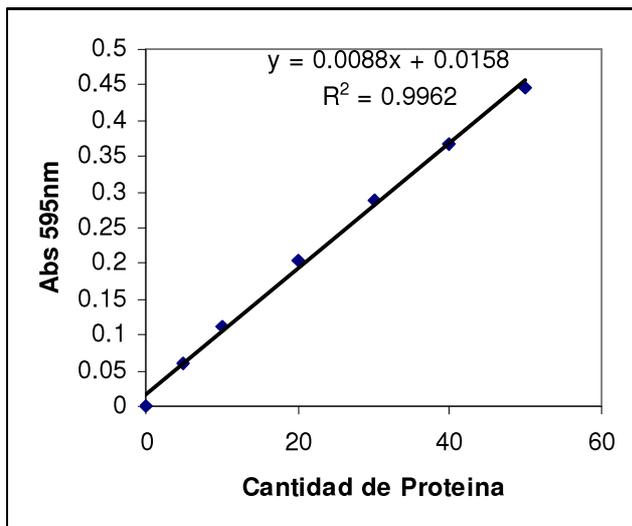
No olvidar de mezclar las cubetas luego de agregar el reactivo de tinción. Esperar 5 minutos antes de la cuantificación. Realizar 6 repeticiones por muestra.

Medir la absorbancia a 595nm, para la curva de concentración y las muestras.

### **Cuantificación utilizando una curva de calibración promedio.**

A partir de muchas curvas de calibración establecidas en el laboratorio, se calculo una curva de calibración promedio.

Cantidad de proteínas	0	5	10	20	30	40	50
Abs a 595nm	0	0,059	0,112	0,204	0,287	0,366	0,446



La curva de regresión lineal es  $Y=0,008X+0,016$

### Anexo 3: Electro-Focalización con el sistema IPGphor: migración según el pI

#### Isoelectroenfoque

##### Tampon de rehidratación de cintas.

Se utiliza la solución de resuspensión a la cual se agrega una pequeña punta de espátula de bromofenol. Esta solución se alícuota en tubos de 2ml y se guarda a -20°C

##### Tampón de equilibración para las cintas

Tris-HCl 1,5M, pH 8,8	16,75ml, Filtrar con membrana 0,4 $\mu$
Urea	180,17g
Glicerol(87% stock)	172,5ml
SDS	10g
Azul de Bromofenol	punta de espátula
Agua milliQ	c.s. 500ml

El DTT (100mg/10ml) o Acetamida (250mg/10ml) son agregadas al momento de uso de la solución.

## Anexo 4: Electroforesis denaturante en gel de poliacrilamida

### Preparación de genes de resolución: Soluciones para electroforesis Bidimensional

#### **Tampón de electrodo 2da dimensión**

Tris		102,3g
Glicina		510g
SDS		34g
Agua destilada	c.s.	34 lt

#### **Solución madre de acrilamida al 30%: (36,5:1)**

Archilamida		584g
Bis-acrilamida		16g
Agua destilada	c.s.	2 lt

#### **Gel de resolución**

Acrilamida 30% sol. Stock		989,3ml
Tris HCl pH 8,8 (2M solución stock)		675ml
SDS (10% sol. Stock)		40,5ml
Sacarosa		27g
Agua destilada	c.s.	2700ml

#### **Preparación de agarosa bajo punto de fusión 1%**

Agarosa Low melting point		1g
SDS (10% stock)		2ml
1,2M Bis-Tris, 0,8M HCl		12,5ml
agua miliQ	c.s.	100ml

## Anexo 5: Coloración de geles con azul coloidal

La coloración se lleva a cabo en cajas de poliestireno transparentes con dos geles por caja en agitación. Se ocupan 400 ml. De solución por caja.

### Fijación (2horas)

Acido fosfórico (85% sol. Stock)	47 ml
Etanol(95% Sol. Stock)	1052 ml
Agua	c.s. 2lt

### Lavado (3 veces por 30min)

Agua

### Incubacion (1h)

Metanol (100% sol. Stock)	680ml
Sulfato de Amonio	340g
Acido Fosfórico(85% sol. Stock)	47ml
agua	c.s. 2lt

Nota: solubilizar el sulfato de amonio con 700ml de agua, luego juntar con la mezcla metanol (680ml) con agua (500ml) y finalmente agregar el acido.

### Coloración (5 días)

Agregar 1g de azul coloidal de Coomasie BG250 para 2lt de solución de incubación

Nota: disolver el azul en la solución metanol agua al menos un hora

### Conservación

5% acido acético en bolsas plásticas selladas a 4°C

## Anexo 5: Modelos estadísticos

### La regresión lineal

Hay tres tipos de modelos:

- a) Modelo Mecanista. Hay un solo resultado posible  
Ej:  $pV=nRT$
- b) Modelo Lógico. Toma decisión respecto a algo (usado en informática)
- c) Modelo Estadístico  
Busca explicar el valor de una variable respuesta en base a una o más variables explicativas y un término residual.

Variable respuesta = Componente sistemática + componente residual

### El análisis de varianza

Es un test de estructuración de los datos, se ayuda de variables explicativas de tipo cualitativo, denominadas factores.

$$Y_{ir} = \mu + \alpha_i + \varepsilon_{ir}$$

$Y_{ir}$  : Variable respuesta

$\mu$  : media general

$\alpha_i$  :Efecto de un factor sobre la variable respuesta

$\varepsilon_{ir}$ : componente residual

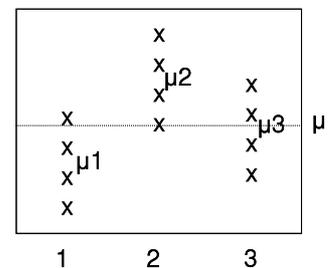
Usando esta ecuación general, consideremos por ejemplo un set de datos de la altura de los árboles de tres bosques,  $\alpha_i$  sera el efecto del bosque  $i$  sobre la altura,  $\mu$  el promedio general de la altura

$$\alpha_1 = \mu - \mu_1$$

$$\alpha_2 = \mu - \mu_2$$

$$\alpha_3 = \mu - \mu_3 \rightarrow 0, \text{ si tiende a cero, no hay efecto bosque 3}$$

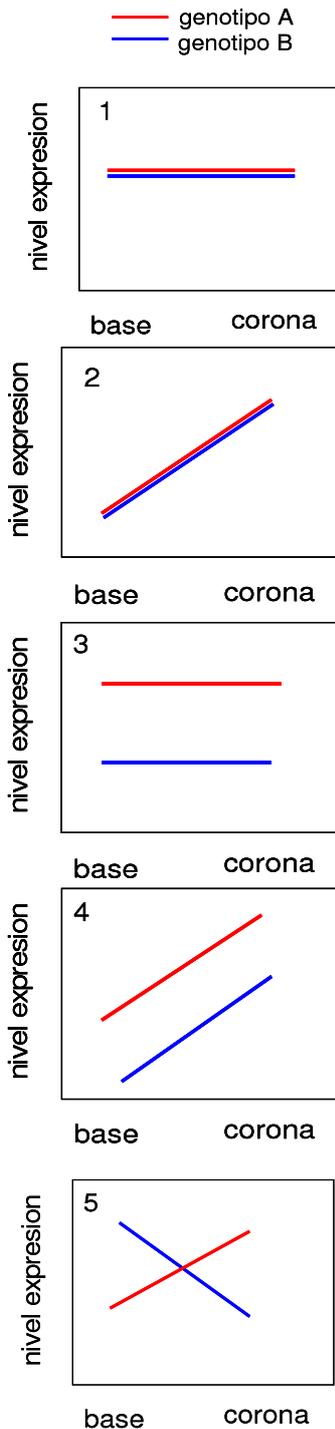
En este caso habría una diferencia en los promedios para el bosque 1 y 2



## Anova de dos factores, (Two-way ANOVA)

Supongamos tenemos dos factores,  $\alpha$  en  $i$  niveles y  $\beta$  en  $j$  niveles. El modelo más general sería

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$



El efecto interacción  $(\alpha\beta)_{ij}$  es interpretado como parte de una respuesta principal no atribuible al efecto aditivo de  $\alpha_i$  y  $\beta_j$ . Por ejemplo usted puede disfrutar de frutillas( $\alpha$ ) y crema( $\beta$ ) individualmente, pero la combinación es superior. En contraste puede gustar del pescado y helado pero no su combinación. A continuación ejemplos gráficos para un ANOVA de dos factores tomando en cuenta como variable respuesta la expresión de proteínas y como factores el factor genotipo ( $\alpha$ ), con un genotipo A y un genotipo B (niveles  $i$ ) y el factor altura del árbol ( $\beta$ ), con los niveles ( $j$ ) base y corona.

En las siguientes figuras se representa en el eje Y el nivel de expresión de una proteína en particular y en el eje X dos niveles de altura del árbol. La línea de color rojo representa el genotipo A y la línea azul representa el genotipo B. En la figura 1 no hay efectos significativos, En la figura 2 el efecto altura es significativo, En la figura 3 solo efecto genotipo es significativo, En la figura 4 hay efecto de el factor genotipo y efecto altura. En la figura 5 solo el efecto interacción es significativo.

## **ANALISIS CLUSTERING**

El análisis de clustering es una técnica de agrupamiento de individuos u objetos en grupos desconocidos. Se diferencia de otros tipos de agrupamiento, como análisis discriminativo, porque en el análisis clustering, el número y las características de los grupos derivados de los datos no son siempre conocidos previamente al análisis.

El análisis de clustering ha sido usado por décadas en el área de la taxonomía, la clasificación se realiza desde un plano general a uno más específico. También el clustering ha sido usado en medicina, clasificando pacientes en determinadas patologías, en antropología clasificando herramientas de piedra y en marketing, clasificando los consumidores en base a sus gustos. En resumen, es posible encontrar aplicaciones para el clustering en casi todo campo de investigación. Es importante destacar que el clustering es altamente empírico, diferentes métodos pueden dar origen a diferentes agrupaciones, además siendo que los grupos no son conocidos con anticipación, es comúnmente difícil decidir si los resultados tienen significado, en el contexto del problema estudiado.

### **Conceptos básicos**

#### **Medición de distancias**

Todos los métodos de clustering requieren definir alguna medida de cercanía o similaridad entre dos observaciones. Esto se puede definir igualmente como distancia.

La distancia más comúnmente usada es la **distancia Euclidiana**. En dos dimensiones, supongamos dos puntos de coordenadas  $(X_{11}, X_{21})$  y  $(X_{12}, X_{22})$ , respectivamente. Entonces la distancia Euclidiana entre dos puntos estaría definida por

$$\text{Distancia} = [(X_{11} - X_{12})^2 + (X_{21} - X_{22})^2]^{1/2}$$

#### **Técnicas Analíticas de Clustering**

Las técnicas comúnmente usadas de clustering son de dos tipos: jerárquicos y no jerárquicos.

#### **Clustering jerárquico**

Métodos jerárquicos pueden ser aglomerativos o divisivos. En los métodos aglomerativos comenzamos con N clusters, por ej. Cada observación constituye su cluster propio. En pasos sucesivos se combinan los dos clusters más próximos, reduciendo el número de clusters de uno

en uno, por cada ciclo, al final del análisis, todas las observaciones están agrupadas en un solo cluster. En métodos divisivos se comienza con un solo cluster conteniendo todas las observaciones. En pasos sucesivos se van separando los casos más diferentes. La mayoría de los programas de clustering son de tipo aglomerativo.

El método del centroide es un método ampliamente usado del tipo aglomerativo. En el método del centroide, la distancia entre dos grupos es la distancia entre sus centroides (el centroide es el punto que sus coordenadas son los promedios de todas las observaciones del cluster). Si un cluster tiene una sola observación, entonces el centroide es la misma observación. El proceso prosigue combinando los grupos de acuerdo a su distancia entre sus centroides, los grupos con menor distancia son combinados en primer lugar.

Un importante problema es como seleccionar el número de clusters. No existe un procedimiento objetivo estándar para hacer esta selección. Las distancias entre clusters en los sucesivos pasos pueden servir como guía. El investigador puede elegir detenerse, cuando la distancia alcance un valor específico o cuando se evidencie un salto en los valores de las distancias. Además el diseño del experimento puede sugerir un número natural de clusters. Si este número es conocido, una técnica apropiada es la técnica de clustering “k-means”.

### **Clustering K-means**

El Clustering K-means es una popular técnica de clustering no jerárquico. Para un número específico de k clusters el algoritmo básico procede en los siguientes pasos:

1. División de los datos en K clusters iniciales. Los miembros de estos clusters pueden ser seleccionados por el programa o por el usuario de acuerdo a procedimientos arbitrarios.
2. Calculo de los promedios o centroides para los k clusters.
3. Para cada valor, calcular su distancia a cada centroide, asignar al grupo cuyo centroide sea el más cercano.
4. Se repite el punto 3 para cada caso.
5. Se repiten los puntos 2, 3 y 4 hasta que no sean reasignados nuevos casos

El clustering es muy sensible a “outliers” o datos alejados de la media general. Los datos deben ser normalizados antes de ser analizados por clustering.

## Anexo 7: Programas R

```
## statistical analisys of spot expression##
##writen by marcelo garces and phillipe chaumeil##

# read data coming from Image Scan in format text#
#the name Group ID has to be manually edited to GroupID, need to erase the
space#
rm(list=ls())
setwd("D:/")
data<-read.table("D:/cambialage2006.txt",header=T,dec=".",sep=" ",as.is=TRUE )
dim (data)

## transform into long format
long <-
reshape(data,v.names="vol",idvar="GroupID",varying=c(list(names(data)[2:13])),
direction="long")

##add row with the number of treatments in replicated#
##and add the column treatment to the data frame#
## gl(number of treatments,replicates,total lenght)#
treatment<-gl(2,nrow(data)*3,nrow(data)*12)
long2<-( cbind(long, treatment))
dim(long2)

write.table(long2,file="long2.txt",append=FALSE,quote=TRUE,sep="\t",eol="\n",n
a="0",dec=".",row.names=FALSE,col.names=TRUE)
  #add supplementary raw in excel and save as text

  #start here with data completed from excel
setwd("D:/")
data<-read.table("D:/cambialage2006ok.txt",header=T,dec=".",sep=" ",as.is=TRUE
)
#perform aov in a bucle and summaryse the results in a table
ref.fact<-as.factor(data$GroupID)
refspotlist<-levels(ref.fact)
refdone<-" "
matresu<-" "
for(i in 1:length(refspotlist)){
tabletemp<-data[data$GroupID==refspotlist[i],]
ind.aov<-
aov(vol~treatment+clone+treatment*clone+gel,data=tabletemp,na.action=na.omit)
tabanova<-summary(ind.aov)
Somcar<-data.frame((tabanova[[1]])$"Sum
Sq",row.names=rownames(summary(ind.aov)[[1]]))
Pvaltab<-
data.frame((tabanova[[1]])$"Pr(>F)",row.names=rownames(summary(ind.aov)[[1]]))
#old r2treat<-(Somcar["treatment",1]/sum(Somcar[,1]))*100
r2treat<-(Somcar[1,1]/sum(Somcar[,1]))*100
r2clone<-(Somcar[2,1]/sum(Somcar[,1]))*100
r2gel<-(Somcar[3,1]/sum(Somcar[,1]))*100
r2treatinclone<-(Somcar[4,1]/sum(Somcar[,1]))*100
```

```

refdone<-c(refdone,refspotlist[i])
#old matresu<-rbind(matresu,c(refspotlist[i],Pvaltab["treatment", 1],r2treat))

matresu<-
rbind(matresu,c(refspotlist[i],Pvaltab[1,1],r2treat,Pvaltab[2,1],r2clone,
Pvaltab[4,1],r2treatinclone,Pvaltab[3,1],r2gel))

}

##write a table with the spot number,p and r2#
rownames(matresu)<-refdone
colnames(matresu)<-c("refspot", "Pvaltreatmet",
"r2treatment", "pvalclone", "r2clone", "pvatreat*clone", "r2treat*clone", "pvalgel"
, "r2gel")
matresu2<-matresu[-1,]
write.table(matresu2,file="aovca2006.txt",append=FALSE,quote=TRUE,sep="\t",eol
="\n",na="0",dec=".",row.names=FALSE,col.names=TRUE)

##write a table with the spot number, p and r2 of spots with a p<0.001
setwd("D:/")
results<-
read.table("D:/ca2006/aovca2006.txt",header=T,dec=".",sep="",na.strings=NULL )
dim (results)
essay<-results[results$pvalclone<0.001,]
essay3<-results[results$Pvaltreatmet<0.001| results$pvatreat.clone<0.001 ,]
dim(essay3)#gives the number of significant spots
# results$pvalclone<0.001
write.table(essay,file="clone.txt",append=FALSE,quote=TRUE,sep="\t",eol="\n",n
a="",dec=".",row.names=FALSE,col.names=TRUE)

#to select the expression data of only the spots with p<0.001
original<-read.table("D:/ca2006/cambialage2006.txt",header=T,dec=".",sep="" )
essay2<-read.table("D:/pselec solo
cambial.txt",header=T,dec=".",sep="",na.strings=NULL )
selclust<-original[match(essay$refspot,original$GroupID),]
#write a table used for clustering
setwd("D:/")
write.table(selclust,file="cloneexp.txt",append=FALSE,quote=TRUE,sep="\t",eol=
"\n",na="",dec=".",row.names=FALSE,col.names=TRUE)

#to generate jpeg objects with individual bloxplots for each representative
spot
datac<-
read.table("L:/Genetique/Marcelo/clustca2005.txt",header=T,dec=".",sep="",na="
")
refspotlistc<-as.factor(datac$GroupID)

for(i in 1:length(refspotlistc)){
tabletempc<-long2[long2$GroupID==refspotlistc[i],]
ind.aov<-aov(vol~treatment,data=tabletempc,na.action=na.omit)
spot<-(tabletempc$GroupID[1])
tabanova<-summary(ind.aov)
Pvaltab<-
data.frame((tabanova[[1]])$"Pr(>F)",row.names=row.names(summary(ind.aov)[[1]]))
jpeg(filename=paste(spot,".jpg"))
#change boxplot to modifie colors #

```

```
#boxplot(vol~treatment,data=tabletemp,ylab="vol",xlab="treatment",main=paste(
"spot no", (tabletemp$GroupID[1]), ", Pvalue:", Pvaltab["treatment", 1]))
boxplot(vol~treatment,data=tabletemp,ylab="vol",xlab="treatment")
dev.off()
}
#
```

## Referencias

- Practical Regression and Anova using R . Julian J. Faraway. July 2002  
<http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf>
- Cluster Analysis in “Computer-Aided Multivariate Análisis”, A.A.Afifi, V.Clark,  
Chapman and Hall
- Formation a la statistique le Modele Lineaire. Frédérique Pellerin.  
<http://www.inra.fr/mia/T/FPstat/FP2.html>
- Electrophorese Bidimensionelle des Proteines. Protocole Utilise a l’INRA Pierroton  
[http://www.pierroton.inra.fr/genetics/2D/Proteomevert/Protocoles/proto\\_bidinew.pdf](http://www.pierroton.inra.fr/genetics/2D/Proteomevert/Protocoles/proto_bidinew.pdf)