UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE SANTE, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT

INRA Centre Val de Loire
Unité Amélioration, Génétique et Physiologie Forestières

THÈSE présentée par :

Alexandre MOREL

Soutenance publique le **20 mars 2014**

pour obtenir le grade de : **Docteur de l'université d'Orléans**
Discipline/ Spécialité : Biologie Végétale et Forestière

Physiologie moléculaire du développement des embryons somatiques de pin maritime (*Pinus pinaster* Ait.) : approches transcriptomique et protéomique

THÈSE dirigée par :
Marie-Anne LELU-WALTER Directrice de Recherches INRA Orléans

RAPPORTEURS :
Pierre COUTOS-THEVENOT Professeur Université Poitier
Hervé ETIENNE Chercheur (HDR) CIRAD Montpellier

JURY :
Françoise CORBINEAU Professeur Université Pierre et Marie Curie Paris VI
Eric LAINE Professeur Université Orléans
Caroline TEYSSIER Chargée de Recherches INRA Orléans

Remerciements

Je tiens à remercier en premier lieu tous les membres de mon jury d’avoir accepté d’évaluer mon travail.

J’exprime toute ma gratitude à Marie-Anne LELU-WALTER pour m’avoir permis de réaliser mes travaux au sein de son équipe. Pour les responsabilités qui m’ont été confiées durant mon séjour au laboratoire, témoignage de sa confiance, pour sa passion envers les "bébés plantes" qu'elle a su me transmettre avec générosité, ainsi que pour sa grande expérience de chercheur qui m'a permis de tenir le cap et d'achever ce projet.

Je remercie Caroline TEYSSIER d’avoir accepté de co-encadrer mes travaux de recherche, pour son implication de tous les instants, ses critiques constructives, sa patience et sa gentillesse sans faille apportées tout au long du projet.

J’exprime toute ma reconnaissance à Philippe Label qui a toujours répondu présent et été de très bon conseil, de par ses réflexions pertinentes et sa rigueur.

Je tiens à adresser tous mes remerciements à Gilles Pilate de m'avoir accueilli au sein de l'Unité AGPF et toutes les personnes qui ont partagé ma présence dans l'unité, et en particulier à Claire Le Metté, Kevin Ader et Nathalie Boizot pour leur aide précieuse dans la réalisation de mes différentes manips. Un grand merci à Patricia Montes, Brigitte Viguier, Franck Rogeon et Jean-Léandre Haton pour leur grande disponibilité, parce qu'il n'y a pas que la paillasse dans la vie.

Merci aussi à toutes les équipes de recherche avec qui j'ai eu le privilège de travailler tout au long de ce projet :

Jean-François Trontin du FCBA de Bordeaux, pour sa participation active et son aide précieuse dans mon projet, ainsi que Isabelle Reymond.

Françoise Corbineau, de part le grand intérêt qu'elle a porté à mon travail et pour sa disponibilité, et son équipe du LCMP de Jussieu, en particuliers Dominique Vinel et Alain d'Harlingue.

Martin Vágner et son équipe de l'Institut expérimental de botanique à Prague, et en particulier Kateřina Eliášová et Bedřich Pešekc.

Martin Cadene et Martine Beaufour du CBM à Orléans, Stéphane Claverol et Anne-Marie Lomenech de l'équipe Protéome au CGFB de Bordeaux.

Sans oublier une pensée particulière à ma famille et mes proches pour leur soutien.

Table des matières

ABREVIATIONS **1**

INTRODUCTION GENERALE **6**

CHAPITRE I - Etat de l'art **8**

Partie I - Le pin maritime ***8***

I.1. Habitat du pin maritime 8
I.2. Importances agronomique et industrielle du pin maritime 10
I.3. Caractéristiques botaniques 12
I.4. Reproduction sexuée du pin maritime 14
I.4.1. La proembryogenèse 16
I.4.2. Embryogenèse précoce 19
I.4.3. Embryogenèse tardive 19
I.4.4. Maturation et déshydratation 20
I.5. Programme d'amélioration et variétés améliorées du pin maritime 22
I.6. Effets post-tempêtes 26

Partie II - Multiplication végétative du pin maritime ***28***

II.1. Bouturage horticole 29
II.2. Greffage horticole 29
II.3. Embryogenèse somatique 30
II.3.1. Initiation des masses embryogènes 32
II.3.2. Multiplication des masses embryogènes 32
II.3.3. Cryoconservation des masses embryogènes 33
II.3.4. Développement et maturation de l'embryon somatique cotylédonaire 34
II.3.5. Germination de l'embryon somatique cotylédonaire et développement en plant 35
II.3.6. Essais au champ de plants de pin maritime 35
II.3.7. Place de l'embryogenèse somatique dans les programmes d'amélioration 37
II.3.7.1. Appui des biotechnologies pour maximiser le gain génétique 37
II.3.7.2. Association de l'embryogenèse somatique et de la cryoconservation 38

Partie III - Marqueurs moléculaires du développement des embryons somatiques et zygotiques ***40***

III.1. Embryogenèse précoce 49
III.1.1. Expression des gènes dans le suspenseur 49
III.1.2. Evénements précoces de mort cellulaire programmée 49
III.1.3. Gènes régulant le cycle cellulaire et la paroi cellulaire 50
III.1.4. Les gènes de régulation du développement embryonnaire 51
III.1.4.1. Gènes Leafy Cotylédon (LEC) 51
III.1.4.2. Gènes de type WUSCHEL et WUSCHEL-related homeobox (WOX) 52
III.1.4.3. Récepteur kinase de l'embryogenèse somatique (SERK) 52
III.1.5. Protéines extracellulaires 53
III.1.5.1. Arabinogalactane protéines, chitinases et lipochitooligosaccharides 53
III.1.5.2. Protéines transporteur de lipides non-specifiques (LTPs) 53

III.1.5.3. Germines et Germin-Like Proteins (GLP) 54
III.2. Embryogenèse tardive 55
III.2.1. Protéines de réserve 55
III.2.1.1. 11S globuline ou legumine-like 56
III.2.1.2. 7S globuline ou viciline-like 57
III.2.1.3. 2S albumine 57
III.2.2. Glucides et amidon 57
III.2.2.1. Métabolisme des glucides lors de l'embryogenèse 58
III.2.2.2. Oligosaccharides non-réducteurs 59
III.2.3. Acide abscissique (ABA) 60
III.2.4. Protéines abondantes de l'embryogenèse tardive (LEA) 61
III.2.5. Les protéines de réponse au choc thermique (HSP) 62

Partie IV - Méthodologies de la physiologie moléculaire 63
IV.1. Approche de transcriptomique globale 65
IV.1.1. Séquençage par terminaison cyclique réversible d'Illumina/Solexa 65
IV.1.2. Difficultés de l'assemblage 67
IV.2. Approche de protéomique globale 70
IV.2.1. L'électrophorèse bidimensionnelle (2DE) 70
IV.2.2. Limites de la protéomique bidimensionnelle 74
IV.3. Annotations des séquences 74
IV.4. Physiologie moléculaire et intégrative 75

Partie V - L'embryogenèse somatique du pin maritime aujourd'hui : problématique et objectifs 78

Références 82

CHAPITRE II - Événements moléculaires précoces impliqués dans le développement des embryons somatiques de *Pinus pinaster* Ait. en réponse à la disponibilité réduite de l'eau : analyses transcriptomique et protéomique 97

II.1. Présentation de la publication 97

II.2. Publication 100
Résumé 100
Summary 101
Abbreviations 101
Introduction 102
Material and Methods 103
Pinus pinaster somatic embryogenesis 103
Fresh Weight (FW), Dry Weight (DW), and water content of EMs 103
Histology and microscopy 103
ABA and abscisic acid glucose ester (ABA-GE) measurements 104
Sample preparation and determination of ABA and ABA-GE 104
ABA determination by GC-MS 104
ABA-GE determination by LC-MS 104
Determination of carbohydrate and total protein content 104
Carbohydrate and starch assay 104
Total protein assay 104
Transcriptomic analysis 105
Sequencing 105
Functional characterization and gene ontology (GO) annotation 105

Functional Classification by KEGG 105
2-D gel proteomic analysis 105
2-D PAGE 105
Gel spot processing 105
NanoLC-MS/MS analysis 105
Protein identification by database sequence assignment 106
Statistical analysis 106
Results 106
Histological and biological characteristics of maritime pine EMs 106
ABA content 106
Carbohydrate and total protein contents 107
Transcriptomic analysis 107
Data annotation by enrichment analysis and functional classification 107
Transcript expression levels 108
2-D gel proteomic analysis 109
Discussion 109
Does gellan gum concentration control maturation of somatic embryos? 111
What happens in EMs when somatic embryo maturation is hindered? 112
Which components appear during maturation of somatic embryos? 112
Conclusions 114
Author contributions 114
References 114
Supporting information 118

CHAPITRE III – Les embryons somatiques cotylédonaires de *Pinus pinaster* [Ait.] se développent de façon similaire à des embryons zygotiques cotylédonaires frais : approches physiologique, biochimique et protéomique 156

III.1 Présentation de la publication 156

III.2 Publication 159
Abstract 163
Introduction 163
Material and Methods 165
Plant material 165
Cotyledonary somatic embryos (SEs) 165
Cotyledonary zygotic embryos (ZEs) 165
Dry weight, water content of SEs and ZEs 166
Determination of carbohydrate and total protein content 166
Carbohydrates 166
Total protein assay 166
1D gel electrophoresis 166
2-D gel proteomic analysis 166
Proteomic study design 166
2-D PAGE 166
Mass spectrometry and data analysis 167
Statistical analysis 167
Results 168
Dry weight and water content of cotyledonary SEs and maturing cotyledonary ZEs 168
Carbohydrate content in cotyledonary SEs and maturing cotyledonary ZEs 168

Quantitative and qualitative changes in protein contents in cotyledonary SEs and maturing cotyledonary ZEs 168
Cotyledonary SEs *vs.* maturing cotyledonary ZEs: a synthesis of quantitative biological and biochemical data 169
2-D gel proteomic analysis of cotyledonary SEs and fresh maturing cotyledonary ZEs 169
Identification of candidate protein markers of the fresh, cotyledonary stage of embryo 169
Discussion 170
Different embryogenic lines produced similar quality standard of cotyledonary SEs..... 170
Quality of cotyledonary SEs did not vary as a function of the duration of maturation 170
Cotyledonary SEs resemble fresh, maturing cotyledonary ZEs 171
Proteomics revealed similar protein profiles in somatic *vs* fresh zygotic embryos ... 171
Protein markers of the fresh cotyledonary embryos stage during somatic and zygotic embryogenesis 172
Conclusions 173
References 175
Figure legends 179
Supplementary data legends 180
Figures 181
Supplementary Figures 196

III.3 Résultats complémentaires 210
III.3.1 Evolutions physiologique et biochimique de l'EZ de pin maritime au cours de sa maturation 210
III.3.2 Evolutions physiologique et biochimiques du mégagamétophyte de pin maritime au cours de la maturation de la graine 211

III.3.2.1 Mesures physiologiques 211
III.3.2.2 Contenu en sucres solubles 213
III.3.2.3 Contenu en protéines totales et en protéines de réserve 214

CHAPITRE IV – Discussion générale 216

IV.1 La diminution de la disponibilité en eau n'est pas perçue comme une contrainte par les ES indifférenciés 216
IV.2 La diminution de la teneur en eau impacte-elle le métabolisme des sucres ? 217
IV.3 Les processus cellulaires et moléculaires clés impliqués dans le développement précoce des embryons somatiques 219
IV.4 Comparaison du développement de l'EZ et de son mégagamétophyte au cours de la maturation de la graine de pin maritime 238

CHAPITRE V – Conclusions et Perspectives 240

V.1 Conclusions 240
V.2 Perspectives 243

BIBLIOGRAPHIE 247

ANNEXES 269

1. Somatic embryo maturation in maritime pine (*Pinus pinaster*): contribution of a 2-DE proteomic analysis for a better understanding..... 270

2. *De novo* assembly of maritime pine transcriptome: implications for forest breeding and biotechnology 282

3. L'embryogenèse somatique : une méthode de multiplication végétative du pin maritime pour demain ?..... 302

4. Composition des milieux de culture..... 311

Abréviations

µM ─ Micromolaire

2,4-D ─ Acide 2,4-dichlorophénoxyacétique (auxine)

2DE ─ Electrophorèse bidimensionnelle

4G ─ 4 g L^{-1}, 4 grammes de gélifiant par litre de milieu de maturation

9G ─ 9 g L^{-1}, 9 grammes de gélifiant par litre de milieu de maturation

ABA ─ Acide abscissique

ABI3 ─ ABA insensitive 3

ACP ─ Analyse en composantes principales

ADH ─ Alcohol dehydrogenase

ADN ─ Acide désoxyribonucléique

ADNc ─ Acide désoxyribonucléique complémentaire

AFLPs ─ Amplified fragment length polymorphism

AFOCEL ─ Laboratoire de Biotechnologie, Association Forêt-Cellulose, ancien FCBA

AGP ─ Arabinogalactan protein

AIB ─ Acide Indole Butyrique

AK ─ Adenosine kinase

AOS ─ Active oxygen species

APRT ─ Adenine phosphoribosyltransferase

APX ─ Ascorbate peroxydase

ARN ─ Acide ribonucléique

ARNi ─ ARN interférent

ARNm ─ ARN messager

BA ─ Benzyl adénine (6-benzylaminopurine)

Chi-1 ─ Chitinase 1

CHIA4 ─ Class IV chitinase

COV ─ Composés d'origine volatile

CPFA ─ Centre de Productivité et d'Action Forestière Aquitaine

CPPU ─ Cytokinine-1-(2-chloro-4-pyridyl)-3-phenylurée

DIGE ─ Differential in gel-electrophoresis

ES ─ Embryon(s) somatique(s)

EST ─ Expressed Sequence Tag

EZ ─ Embyon(s) zygotique(s)

F-actin capping protein ─ Filament-actin capping protein

FAO ─ Food and Agriculture Organization of the United Nations

FCBA ─ Institut technologique chargé de la Forêt, de la Cellulose, du Bois-construction et de l'Ameublement.

FPKM ─ Fragments per kilobase per million reads sequenced

fru ─ Fructose

FT ─ FLOWERING LOCUS

FUS3 ─ FUSCA3

GA2OX ─ Gibberelline 2-oxidase

Gb ─ Giga paires de bases, 1.10^9 paires de bases

GC/MS ─ Gas chromatography/Mass spectrometry

GC/TOF-MS ─ Gas chromatography/Time-of-flight mass spectrometry

GER1 ─ Germin protein 1

GIS ─ Groupement d'Intérêt Scientifique

GLP ─ Germin-like protein

glu ─ Glucose

Glu-1 ─ β-1,3-glucanase 1

GO ─ Gene Ontology

GS1 ─ Glutamine synthetase 1

GTPase ─ Guanosine triphosphate hydrolase

HD-Zip I ─ Homeodomain-leucine zipper I

HR ─ Humidité Relative

HSP ─ Heat Shock Protein, Protéine de réponse au stress thermique

IAA11 ─ Indole-3-acetic acid inducible 11

IEF ─ Isoélectrofocalisation

IFN ─ Inventaire Forestier National

IK ─ Inosine kinase

INRA ─ Institut National de la Recherche Agronomique

IRSTEA ─ Institut de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture

kDa ─ Kilodalton

L1L ─ LEAFY COTYLEDON 1-like

LBD12/ASL5 ─ LOB domain-containing protein 12/ASYMMETRIC LEAVES 2-like protein 5

LC-MS/MS ─ Liquid chromatography coupled to tandem mass spectrometry

LCOs ─ Lipochitooligosaccharides

LEA ─ Late embryogenesis abundant protein

LEC1 ─ LEAFY COTYLEDON 1

LMW HSPs ─ Low molecular weight HSP

LRR ─ Leucine-rich-repeat protein

LTP ─ Lipid transfer protein

m/z ─ Ratio masse/charge

MAAPRAT ─ Ministère de l'Agriculture, de l'Alimentation, de la Pêche, de la Ruralité et de l'Aménagement du Territoire

MC9 ─ Metacaspase de type 9

mcII-Pa ─ Metacaspase de type 2 de *Picea abies*

Méga. ─ Mégagamétophyte

MEs ─ Masses embryogènes

MF ─ Masse fraîche

MM ─ Masse Moléculaire

MS ─ Masse sèche

MYB77 ─ Myeloblastosis transcription factor type 77

NARS2 ─ NAC-regulated seed morphology 2

NCBI ─ National Center for Biotechnology Information

NIP ─ Nodulin-like intrinsic protein

NOD factor ─ Nodulation factor

NPA ─ N-1-naphthylphthalamic acid

ONF ─ Office National des Forêts

p34cdc2 ─ Cell division cycle protein 2 homolog

Pa1-18 ─ Gènes de *Picea abies*

PaHB1 ─ Homeobox1de *Pice abies*

Pavp 1 ─ Transcription factor viviparous 1 de *Picea abies*

pb ─ Paire de bases

PCD ─ Programmed cell death (mort cellulaire programmée)

PCR ─ Polymerase chain reaction

PDC ─ Pyruvate decarboxylase

PDF2 ─ Protodermal factor 2

PEG ─ Polyéthylène glycol

PgAGO ─ Protéine ARGONAUTE de *Picea abies*

PgEMB ─ Embryogenesis-associated protein de *Picea glauca*

PGR ─ Plant Growth Regulator, régulateur de la croissance des plantes

pH ─ Potentiel Hydrogène

pHi ─ Point iso-électrique

PIB ─ Produit intérieur brut

PIN ─ Polar auxin inhibitor

PPC2 ─ Protein phosphatase 2C

Protéomique 2D ─ Protéomique bidimensionnelle

QrEm ─ Early methionine-labelled protein de *Quercus robus*

qRT-PCR ─ PCR quantitative en temps réel

Rab1 ─ Response sensitive ABA 1

raf ─ Teneur en raffinose soluble, exprimée en µg de raffinose par mg de MS (µg mg^{-1} MS)

ratio hex─sac ─ Hexoses/saccharose = (glucose + fructose)/saccharose

ratio RFO─sac ─ RFO/saccharose = (raffinose + stachyose)/saccharose

RFO ─ Raffinose

RGP ─ Reversibly glycosylated polypeptides

RISC ─ RNA-induced silencing complex

RLK ─ Receptor-like kinase

Roxy1 ─ Glutaredoxin 1

RPN1 ─ Regulatory proteasome non-ATPase subunit 1

RT-PCR ─ Reverse transcriptase PCR

sac ─ Saccharose

SAM ─ S-adénosylméthionine synthase

SAUR ─ Small auxin-up RNA

SDS ─ Dodécylsulfate de sodium

SDS-PAGE ─ Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SERK ─ Somatic embryogenesis receptor-like kinase

smHSP ─ Small Heat Shock Protein

SNP ─ Single nucleotide polymorphism

SPY ─ SPINDLY

sqRT-PCR ─ Semi-quantitative RT-PCR

SRK2 ─ Serine/threonine-protein kinase

SSR ─ Simple sequence repeat

sta ─ Stachyose

TE ─ Teneur en eau

TF 1 ─ TERMINAL FLOWER 1

TIR1 ─ Transport inhibitor response 1

TMP-C ─ Transmembranaire protein C

UDP ─ Uridine diphosphate

WOX ─ WUSCHEL-related homeobox

WUS ─ WUSCHEL

Introduction générale

L'augmentation rapide de la population mondiale s'est traduite par une forte augmentation de la demande en bois (combustible, pâte à papier, construction…). Au cours des prochaines décennies, les besoins toujours croissants en bois et dérivés vont entraîner une pression considérable sur la production forestière. La forêt cultivée aura donc un rôle économique de plus en plus important à jouer avec une exigence croissante de la qualité des produits. Disposer d'une ressource ligneuse abondante et de qualité nécessitera d'optimiser les pratiques sylvicoles ainsi que le développement de variétés génétiquement améliorées dans le cadre d'un développement durable. L'amélioration génétique des ligneux qui se fait par sélection classique, est un processus lent en raison de la durée des cycles biologiques des arbres, tandis que la multiplication du matériel forestier amélioré fait face à des difficultés de propagation en masse par graines et/ou boutures. Seule l'association de nouvelles méthodologies de sélection et des biotechnologies représente une véritable perspective pour améliorer et multiplier en grand nombre les espèces forestières. Les génotypes élites sélectionnés devront nécessairement être produits en quantités considérables sur la base de méthodes de dissémination végétative rapides et bien maîtrisées. Le maintien de la compétitivité de la forêt française dépendra en partie de la prise en compte de ces technologies dans les travaux d'amélioration génétique. Leur mise au point est particulièrement justifiée pour les espèces faisant l'objet d'une exploitation intensive comme le pin maritime.

En Europe, les pins constituent des espèces d'intérêt majeur. Le pin maritime (*Pinus pinaster* Ait.) est l'espèce forestière industrielle majeure en Europe du Sud-Ouest dont le massif aquitain, en France. La multiplication végétative du pin maritime est à ce jour assurée par greffage ce qui limite l'obtention d'un grand nombre de plants. Cela limite donc la mise en place de tests clonaux permettant d'étudier la valeur génétique totale de l'individu. Ces problèmes ne se rencontrent pas avec des plants issus d'embryons somatiques (ES), obtenus par culture *in vitro.* L'embryogenèse somatique est reconnue comme la méthode la plus performante pour la propagation clonale d'arbres forestiers. C'est pourquoi des recherches dans ce domaine ont été entreprises chez le pin maritime à l'INRA d'Orléans (Lelu et al.,

1999; Lelu-Walter et al., 2006) et au FCBA (Ramarosandratana et al., 2001a, b; Breton et al., 2005, 2006).

Les recherches se concentrent actuellement sur le développement d'ES cotylédonaires à partir des cultures embryogènes afin d'optimiser le processus d'embryogenèse somatique pour garantir une régénération de plants de qualité. Cependant, chez le pin maritime l'embryogenèse somatique n'est pas optimisée, la limite étant le contrôle du développement des ES.

L'objectif de ce travail de thèse est d'étudier le développement de l'ES au travers de 2 principales questions:

i.) Dans le milieu de maturation des ES, la disponibilité en eau est impactée par le gélifiant du milieu de culture. L'augmentation de sa concentration, favorable à la différenciation des embryons, réduit la disponibilité en eau du milieu vers les ES. Quels sont les mécanismes physiologiques, cellulaires et moléculaires qui contrôlent la réponse des ES à leur environnement et favorisent leur différenciation ?

ii.) La régénération de plants s'obtient en prélevant des ES cotylédonaires et en les transférant sur milieu de germination. A ce jour, cette sélection est basée uniquement sur des critères de ressemblance morphologique avec l'embryon zygotique (EZ). Or ces critères sont insuffisants pour décrire le véritable stade de maturité atteint par l'ES. Existe-t-il, comme pour l'EZ, une relation entre la composition en composés de réserve (protéines et sucres) et la maturité physiologique ?

La compréhension des mécanismes ontogénétiques nécessite la prise en compte des différents niveaux de complexité biologique et de leurs relations intra- et inter-cellulaires. Ainsi plusieurs techniques complémentaires ont été mises en œuvre pour mieux décrire ces relations et obtenir une vision multi-échelle et donc plus globale du processus d'embryogenèse somatique. Les ES ont ainsi été décrits selon des observations macro et microscopiques, au niveau de leur teneur en eau, de leur contenu en sucres solubles, en phytohormone endogène, en protéines totales et au niveau de leur transcriptome, de leur protéome. Chez les conifères cette approche multi-échelle, la comparaison du transcripome et du protéome sont très innovantes dans le domaine de l'étude de l'embryogenèse somatique, car jamais entreprises à ce jour.

L'objectif *in fine* est de mettre à disposition un outil puissant de multiplication permettant la propagation d'arbres élites.

CHAPITRE I - Etat de l'art

Partie I - Le pin maritime

I.1. Habitat du pin maritime

Le pin maritime est une espèce méditerranéenne bien adaptée aux climats marins très tempérés, à température douce et régulière. En effet, il croît dans les régions à climats océaniques sur les sols sablonneux des zones littorales, en revanche, il exige une légère humidité de l'air mais il supporte bien la sécheresse estivale. Le pin maritime pousse d'une façon optimale en terres acides et profondes, il est considéré comme espèce frugale car il s'adapte à tous les types de sols même les plus pauvres (sur la roche, dans les sables arides et dans les terres crayeuses qui ne conviennent à aucun autre végétal) et les plus acides (sables dunaires), voire les sols hydromorphes, mais il est totalement intolérant au calcaire (espèce calcifuge). Pour sa culture, il faut absolument éviter de le planter sur des sols limono-sableux riches ou sur des terres agricoles, autrement il présentera des défauts de forme et se montrera très sensible aux attaques de la pyrale du tronc (*Dioryctria sylvestrella*).

L'espèce se localise dans les zones non loin de la mer proches des côtes sud atlantiques et méditerranéennes où elle y pousse spontanément (zone entre 0 et 800 m. d'altitude). Elle est assez commune en Espagne, Italie, Portugal et Maroc mais aussi en Afrique du Sud où elle a été introduite pour être cultivée à grande échelle. Elle reste l'espèce la plus répandue dans l'ouest du bassin méditerranéen (**Figure I.1**), actuellement elle couvre plus de 10% de la surface boisée française (Alia et Martin, 2003).

En France l'aire de répartition naturelle du pin maritime se situe dans le massif forestier des Landes de Gascogne, la Corse, les Maures, l'Estérel et les Corbières, dans la forêt de la Double (massifs forestiers du sud-ouest, **Figure I.2**). Cette espèce a également été introduite dans d'autres régions comme sur la côte méditerranéenne principalement en Cévennes méridionales, dans les massifs des Maures ou de l'Estérel, mais aussi dans la région Charente, Aquitaine et dans la vallée de la Loire. Le pin maritime constitue le plus souvent des forêts monospécifiques.



Figure I.1. Aire de répartition (bleu) du pin maritime en Europe et Afrique du Nord **(Alia et Martin, 2003)**.

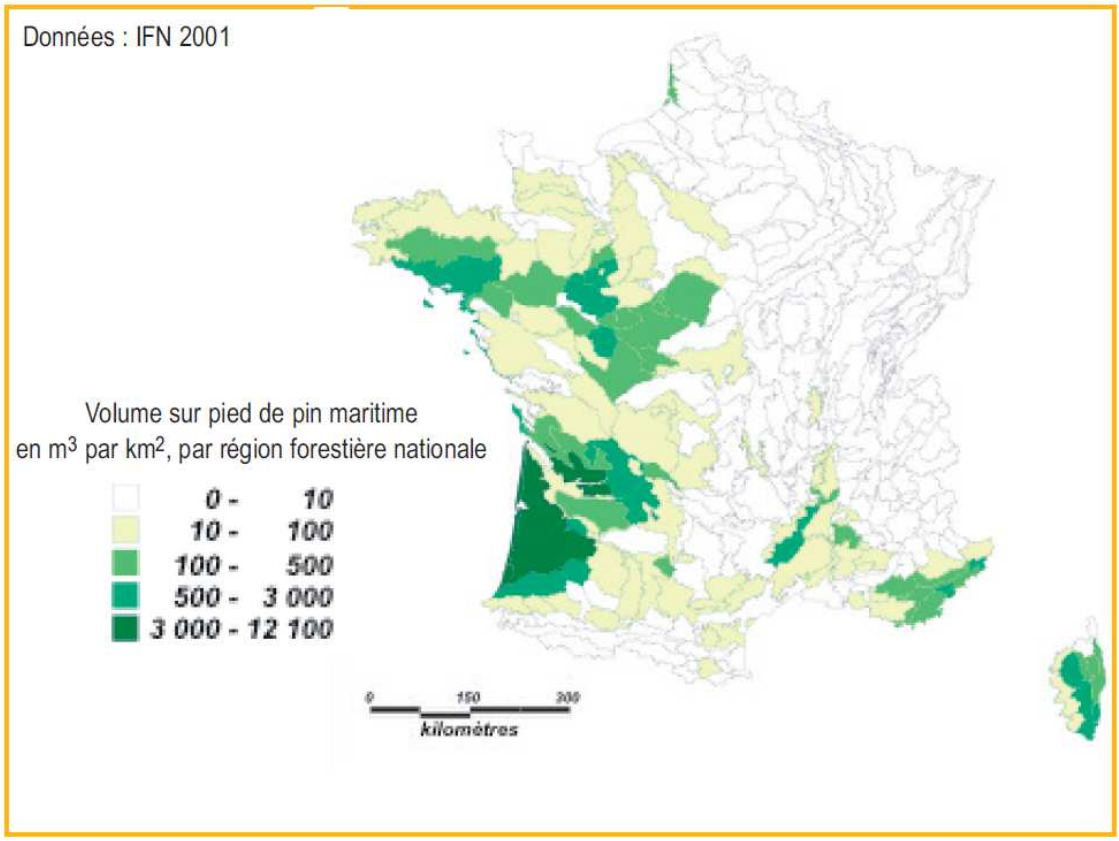


Figure I.2. Aire de répartition du pin maritime en France (**Pignard et al., 2001**).

Néanmoins, des associations végétales existent mais varient en fonction du sol : sur sols pauvres, il est associé à des plantes acidophiles des Landes ; sur sols riches, il côtoie l'arbousier, le chêne pédonculé, le chêne vert, voire le chêne liège dans le sud des Landes. Son couvert, clair et léger, contribue au développement d'une strate arbustive et herbacée épaisse particulièrement combustible.

Pour une régénération naturelle par semis, il exige la pleine lumière, donc aucun couvert, la chaleur et une certaine humidité atmosphérique. L'enracinement s'effectue à deux niveaux ; pivotant d'abord, puis traçant particulièrement adapté pour le reboisement sur des sols difficiles. Ordinairement, le pin maritime a une croissance végétative rapide sur les sols profonds. En général, il produit annuellement des pousses de 30 à 40 cm de long durant son jeune âge. Sensible aux fortes gelées, il craint les hivers très rigoureux avec des expositions prolongées à des températures inférieures à -15°C.

I.2. Importances agronomique et industrielle du pin maritime

Les espèces de pins jouent un rôle particulièrement important dans les plantations forestières dans le monde moderne et constituent à la fois une grande partie de la récolte annuelle de bois et des forêts de jeunes plants qui fourniront du bois dans l'avenir (Boyer et al., 1993; FAO, 2006). En effet, le grand nombre d'espèces de la famille des pins leur permet d'occuper des environnements très variés, certaines espèces présentant des rendements élevés même dans des conditions défavorables. Les pins sont également bien adaptés dans le cadre du reboisement et de la monoculture. Leur bois est facilement traité et utilisé dans une grande variété de produits finaux tels que le bois de charpente, les pâtes à papiers ou encore le contre-plaqué (Plomion et al., 2007).

En raison de son très bon développement en zones littorales, le pin maritime est utilisé à la fois pour la fixation des dunes littorales (Favennec, 1999) ainsi que pour la mise en valeur des zones marécageuses. En France l'exemple le plus spectaculaire est celui des dunes fixées au niveau des Landes de Gascogne où des milliers d'hectares ont été reboisés durant le siècle dernier afin de protéger les régions avoisinantes (habitations, parcelles agricoles...) des dépôts de sable provenant de l'océan atlantique (Palmisano et al., 1994).

Le pin maritime est l'espèce forestière industrielle majeure du sud-ouest de l'Europe (France, Espagne et Portugal). Plus d'un quart de la ressource (1,1 Mha) est localisée en France où il représente la première espèce résineuse en surface (**Table I.1**). Depuis 1860 dans

les Landes de Gascogne, les peuplements de pin ont toujours été exploités, dans un premier temps maintenant révolu, pour la récolte de sa résine, puis pour l'exploitation de son bois. Cette espèce fournit à elle seule chaque année 24,7 % du bois d'œuvre (5,6 Mm^3) et 27,1 % du bois d'industrie (3,3 Mm^3, Thivolle-Cazat et al., 2013). La quasi-totalité de cette production (près de 8,5 Mm^3/an) est concentrée en région Aquitaine, première forêt cultivée d'Europe, où le pin maritime est l'espèce prépondérante sur près de 830 000 ha (Colin et al., 2010).

Table I.1. Classement des 10 premières espèces forestières les plus vendues en nombre de plants. Extrait des résultats de l'enquête statistique annuelle MAAPRAT/IRSTEA sur les ventes de plants forestiers pour la campagne de plantation 2010-2011.

N° en 2011 2012	Espèces les plus vendues* (hors matériels destinés aux TCR)	Ventes en France 2011-2012	*Rappel Ventes en France 2010-2011*	*Variation France*	Export 2011-2012	*Rappel Export 2010-2011*	*Variation Export*	TOTAL ventes (France+Export) 2011-2012
1	Pinus pinaster	29 095 032	*18 472 430*	+58%	250	*2 000*	-88%	29 095 282
2	Pseudotsuga menziesii	7 221 432	*5 479 286*	+32%	1 038 860	*519 900*	+100%	8 260 292
3	Quercus petraea	4 019 902	*3 033 218*	+33%	25 490	*80 700*	-68%	4 045 392
4	Picea abies	1 804 549	*1 781 572*	+1%	237 628	*174 500*	+36%	2 042 177
5	Abies nordmanniana	1 773 892	*1 196 573*	+48%	457 500	*360 350*	+27%	2 231 392
6	Pinus nigra corsicana	1 120 070	*1 140 143*	-2%	3 825	*12 940*	-70%	1 123 895
7	Peupliers-cultivars (plançons, cf. ANN. B)	1 038 765	*794 613*	+31%	22 391	*59 942*	-63%	1 061 156
8	Robinia pseudoacacia	938 289	*779 215*	+20%	5 900	*3 800*	+55%	944 189
9	Larix decidua	894 162	*894 009*	+0%	11 900	*10 300*	+16%	906 062
10	Pinus taeda	805 931	*647 938*	+24%	0	*0*		805 931

Le pin maritime est un poids lourd économique : en France, c'est la 1ère espèce en nombre de plants vendus pour la reforestation (plus de 18 millions en 2010-2011).

Le pin maritime est donc un pilier de l'économie de la filière forêt-bois en Aquitaine qui pèse environ 34 000 emplois (16,5 % des effectifs nationaux) et 2,5 milliards d'euros de chiffre d'affaire, dont 0,9 milliards à l'export (Pédezert, 2009) (**Table I.1**). « *Le nombre de ses salariés, son implication dans le maintien de la cohésion territoriale, son poids dans la constitution du PIB aquitain confèrent à cette filière un rôle décisif tant économique que social* » (Conseil Régional d'Aquitaine).

Avec une productivité moyenne de 11,8 m^3/ha/an (Colin et al., 2010), contre seulement 4m^3/ha/an dans les années 50, le pin maritime est très supérieur aux autres pins présents en France (5,9 m^3/ha/an) et reste très compétitif face aux meilleurs résineux (sapin, épicéa ou douglas) pour lesquels les forêts de production occupent des surfaces plus faibles et/ou plus morcelées et souvent en zone de montagne ne favorisant pas le bûcheronnage mécanisé. Cette très nette augmentation a pu être réalisée par l'apparition de pratiques sylvicoles mécanisées et par la mise en place d'un programme d'amélioration du pin maritime.

I.3. Caractéristiques botaniques

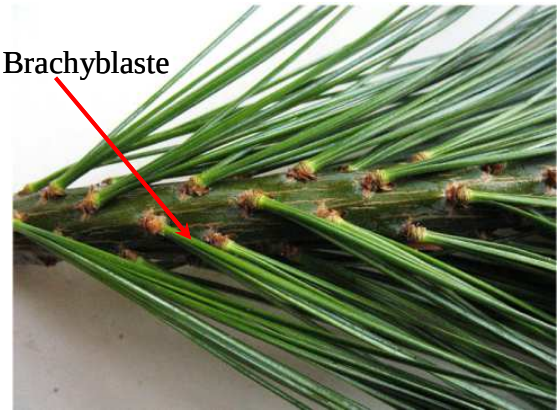
Le pin maritime, *Pinus pinaster* Ait., est un conifère appartenant à la division des gymnospermes et à la famille des Pinacées. C'est une espèce post-pionnière nomade et sempervirente (toujours verte), à longévité moyenne (200 ans) qui atteint le stade de maturité sexuelle vers l'âge de 15-20 ans (**Figure I.3a**). L'arbre se caractérise par une croissance initiale rapide, un port droit, le fût étant souvent flexueux surtout à la base. A son jeune âge, la cime est de forme conique devenant, chez le sujet plus âgé, irrégulière plutôt plate et étalée.



A. Pin maritime mature de 15 ans



B. Ecorce de pin maritime



C. Jeune rameau de pin maritime



D. Cône femelle mature de pin maritime



E. Graine de pin maritime

- Nombre de cônes par hectolitre: 240
- Poids de graines ailées par hectolitre de cône : 2,2 kg
- % d'ailes dans les graines : 13%
- Rendement en graines par cône : 100 graines par cône
- Poids moyen de 100 graines : 5.6 grammes
- Nombre de graines au kilo : 18000
- % de graines vaines : 4 à 5 %
- Faculté germinative moyenne : 90 à 95 %

Figure I.3. Caractéristiques botaniques du pin maritime.

Les défauts de la forme du tronc sont très caractéristiques de l'espèce. L'arbre est alors dégarni à la base avec un houppier de couleur vert sombre plus dispersé. Il atteint généralement 20 à 30 mètres de haut. L'écorce est de couleur brun violet, rouge sombre en profondeur, profondément crevassée et très épaisse en vieillissant (**Figure I.3b**). Le bois est rougeâtre foncé avec un aubier nettement plus clair. Les canaux résinifères sont nombreux, gros, et visibles à l'œil nu, la sécrétion de résine est abondante. Les aiguilles ou pseudophylles persistantes, sont regroupées par deux au niveau d'un court rameau, le brachyblaste (bourgeon dormant situé à l'insertion des aiguilles) très caractéristique de l'espèce (**Figure I.3c**). Elles sont de couleur vert foncé, épaisses, rigides et luisantes et peuvent atteindre 15 cm ou plus de long.

I.4. Reproduction sexuée du pin maritime

Le pin maritime est une espèce monoïque et allogame quasi-stricte. La fructification commence vers 6 ou 7 ans mais la mise à fleur ne devient régulière et abondante que vers une quinzaine d'années. D'une manière générale, les cônes femelles (**Figure I.4**) sont positionnés dans la partie supérieure de l'arbre sur les rameaux principaux des 3 ou 4 derniers verticilles. Les cônes mâles (**Figure I.4**) sont au contraire situés dans la partie inférieure du houppier. Dans la plupart des cas, un rameau ne porte qu'une catégorie de cône.

Le cycle de reproduction sexuée du pin maritime s'étale sur 3 années (**Figure I.4**). En été de l'année n, se produit l'initiation florale qui est en grande partie conditionnée par les conditions environnementales (site, stress hydrique..). Les ébauches florales femelles sont décelables dès la fin de l'hiver. Ensuite le cône se transforme, les écailles ovulifères se développent et s'écartent pour devenir réceptives au pollen vers la mi-avril. Les conditions météorologiques et notamment la température (somme des températures depuis le mois de février) définissent la période de réceptivité de la fleur (Renauld, 1995) : ainsi la période de réceptivité des cônes femelles peut varier de plus de 1 mois d'une année à l'autre (fin mars à début mai). Après la période de pollinisation du cône (année n+1), celui-ci se transforme alors en cônelet au cours de l'été. La fécondation proprement dite ne se produit que 12 mois après la pollinisation (année n+2). Le cônelet se développe alors très rapidement pour atteindre sa taille définitive vers le mois de juillet. La maturation du cône se produit en été : dès le mois d'août le cône commence à changer de couleur. Fin octobre, le cône est mûr (couleur marron clair). L'ouverture naturelle du cône intervient au printemps de l'année suivante (année n+3),

sous l'effet de fortes chaleurs, qui entraînent mécaniquement l'ouverture des écailles et la libération des graines ailées.

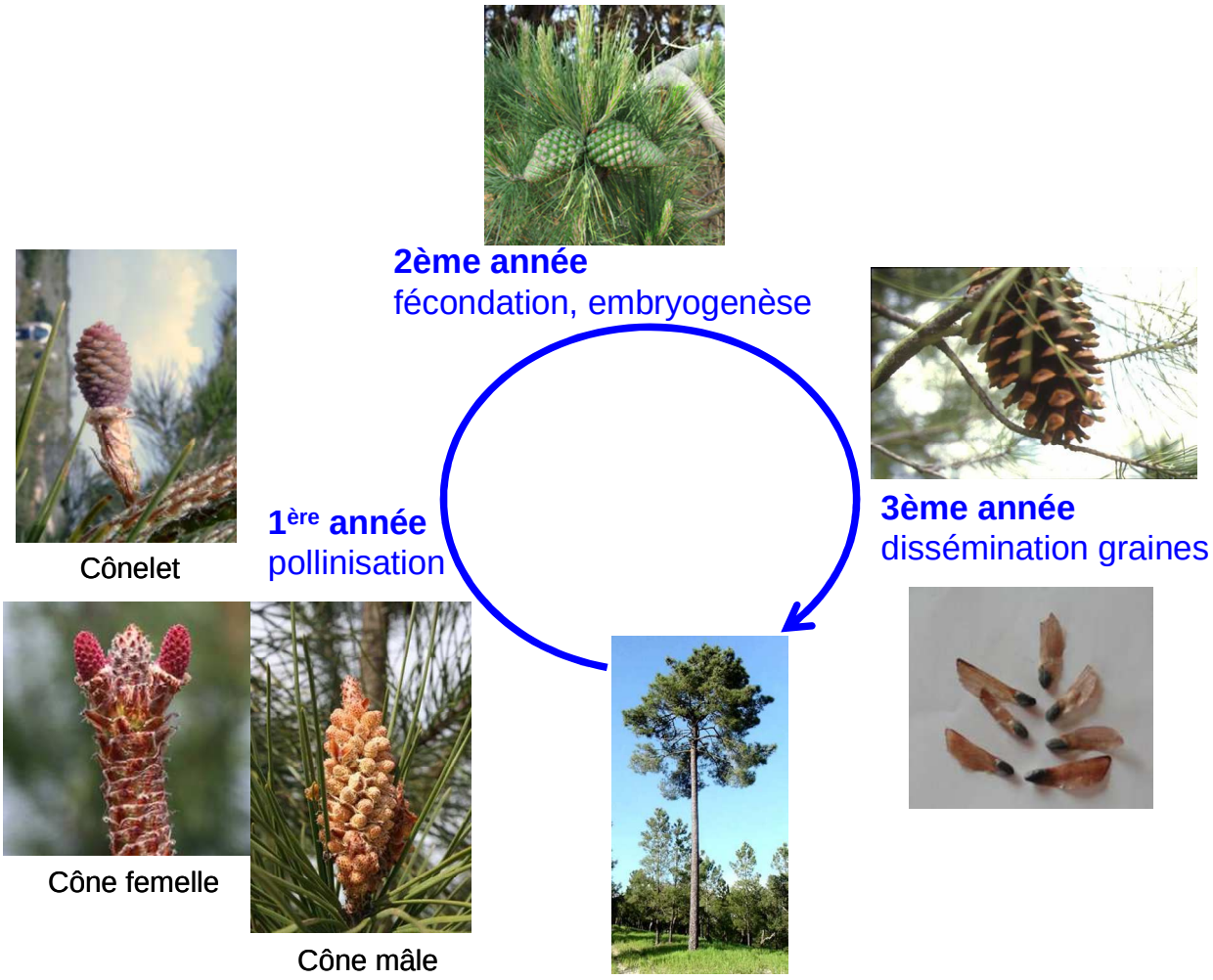


Figure I.4. Cycle de reproduction sexuée du pin maritime.

Le cône peut être récolté fin octobre **(Figure I.3d)**. Après la récolte et dans l'attente d'une ouverture artificielle, le cône doit être mis à ressuyer afin que son taux d'humidité soit suffisamment bas pour permettre l'ouverture en étuve. La graine de pin maritime est composée d'une coque dure, de forme ovoïde, de couleur marron noire et d'une aile membraneuse qui permet la dispersion naturelle des graines par le vent, la forme, la couleur et la striation des ailes étant spécifiques à chaque individu **(Figure I.3e)**. Le rendement en graines par cône peut être très variable suivant plusieurs facteurs (**Figure I.3**) : le site, l'âge des arbres (les jeunes arbres ont des cônes de petits calibres avec de faibles rendements en graines), l'origine génétique de l'arbre et enfin la qualité de la pollinisation (Alazard et al., 2005).

Parmi les gymnospermes, les conifères sont les plus représentés. Leur plan de développement embryonnaire, souvent considéré comme le modèle pour les gymnospermes, est très différent de celui des angiospermes. Le processus peut être divisé en quatre phases (Singh, 1978) :

(1) proembryogenèse - tous les stades depuis la fécondation et avant l'allongement du suspenseur

(2) embryogenèse précoce - tous les stades pendant et après l'allongement du suspenseur et avant l'établissement du méristème racinaire

(3) embryogenèse tardive - mise en place des méristèmes racinaire et caulinaire, suivie du développement de l'embryon

(4) maturation - accumulation des réserves, réduction du métabolisme et déshydratation

I.4.1. La proembryogenèse

Les graines de gymnospermes possèdent de un à quatre archégones, chacun contenant un ovule. La fécondation pouvant se produire dans chaque archégone, entraîne initialement la présence potentielle de plusieurs zygotes, donc de différents génotypes dans une graine. Ce processus appelé polyembryogenèse « simple » est caractéristique de la reproduction des gymnospermes **(Figure I.5**).

Contrairement aux angiospermes où il y a une double fécondation (formation du zygote et de son tissu nutritif l'albumen, tissu triploïde) chez les gymnospermes il n'y a qu'un seul évènement de fécondation ; le zygote se développe dans un tissu ayant subi la méïose donc haploïde nommé endosperme ou megagamétophyte (Dogra, 1967; Nagmani et al., 1995, **Figure I.5**).

Après la fécondation, l'EZ subit deux évènements de duplication nucléaire sans cytodiérèse pour entrer dans une phase à quatre « noyaux libres », suivie par une phase de cellularisation formant deux couches de cellules et huit pro-embryons unicellulaires. Ces couches se divisent par la suite pour former quatre niveaux cellulaires. Les cellules du premier et deuxième niveaux se multiplient pour former l'embryon proprement dit, tandis que les cellules des troisième et quatrième niveaux vont s'allonger subissant une division cellulaire limitée pour former le suspenseur embryonnaire primaire (Dogra, 1967; Singh, 1978; Nagmani et al., 1995, **Figure I.6**).

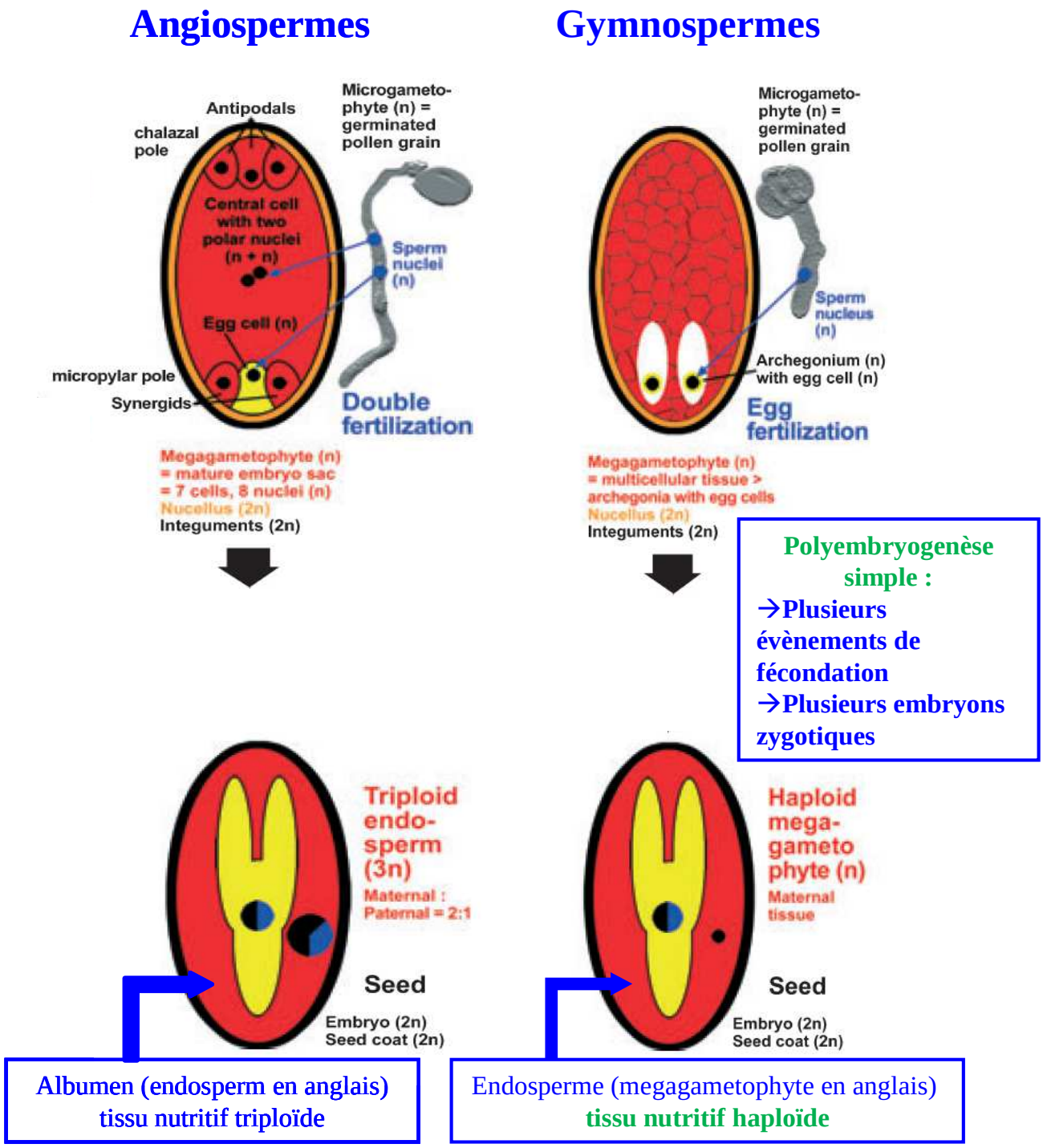


Figure I.5. Différence entre les évènements de fécondation chez les angiospermes et les gymnospermes **(Linkies et al., 2010)**.

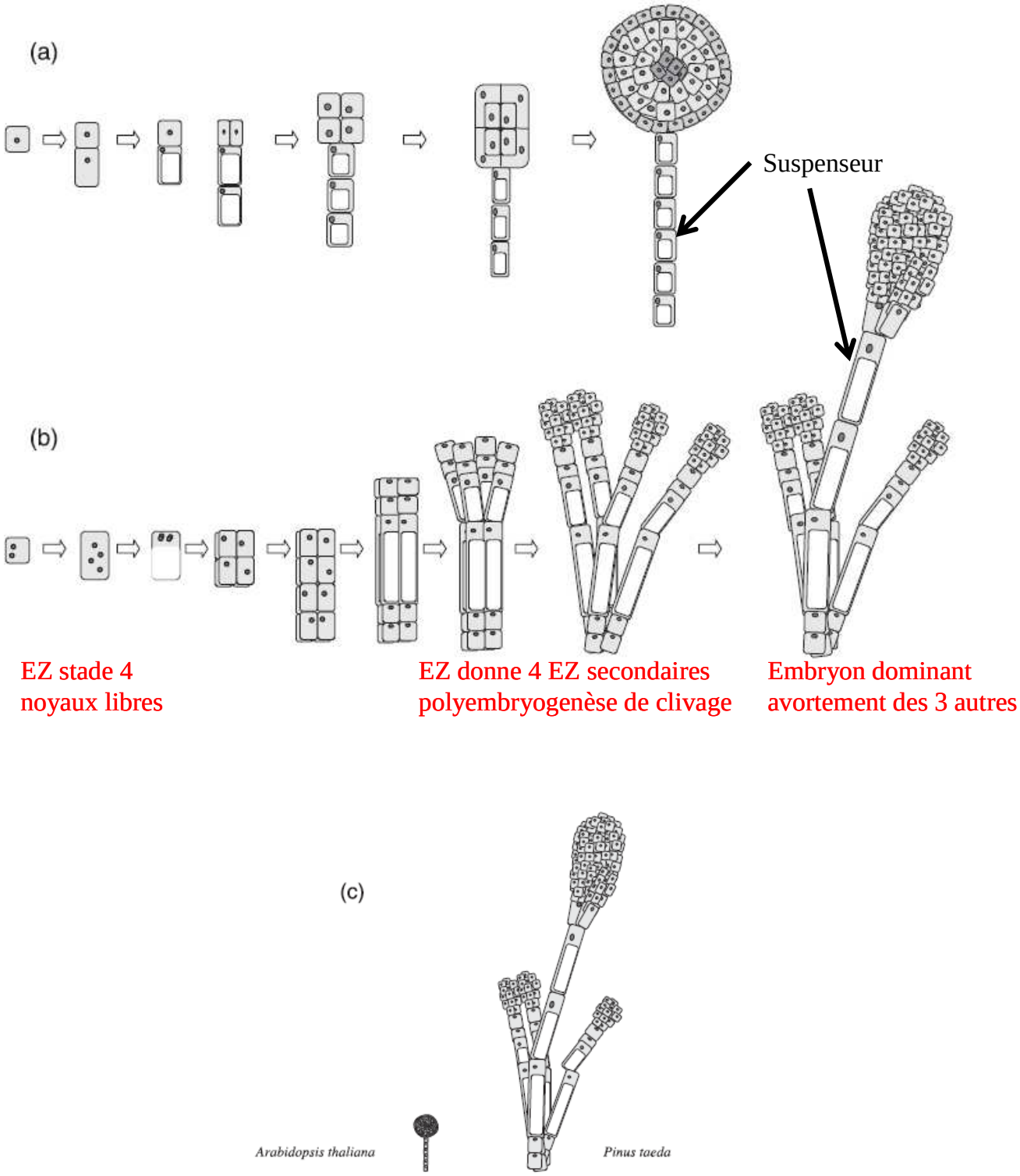


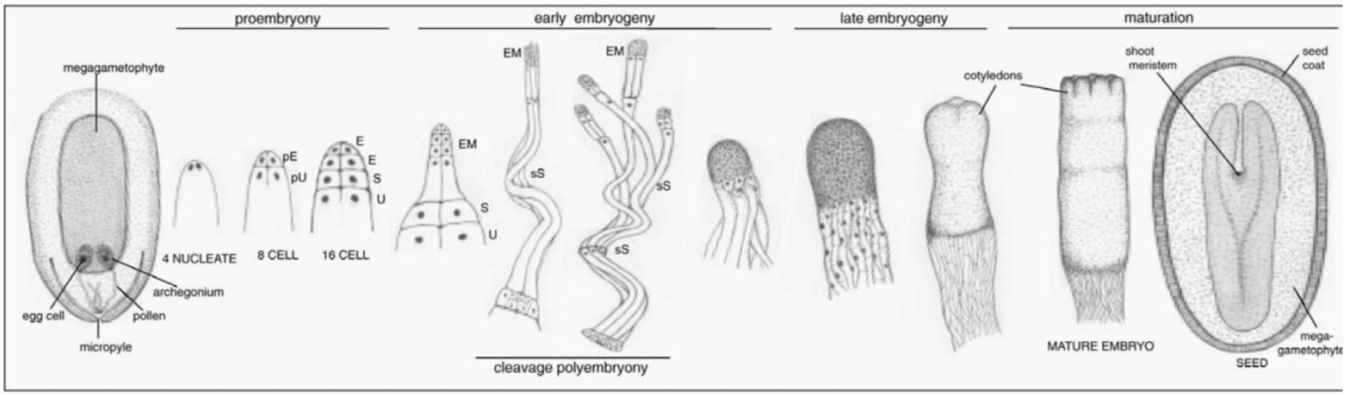
Figure I.6. Représentation schématique des événements précoces de l'embryogenèse chez **(a)** *Arabidopsis thaliana* et **(b)** *Pinus taeda*. **(c)** la taille relative de l'embryon globulaire d'*Arabidopsis* et de l'embryon dominant de pin **(Cairney et Pullman, 2007)**.

I.4.2. Embryogenèse précoce

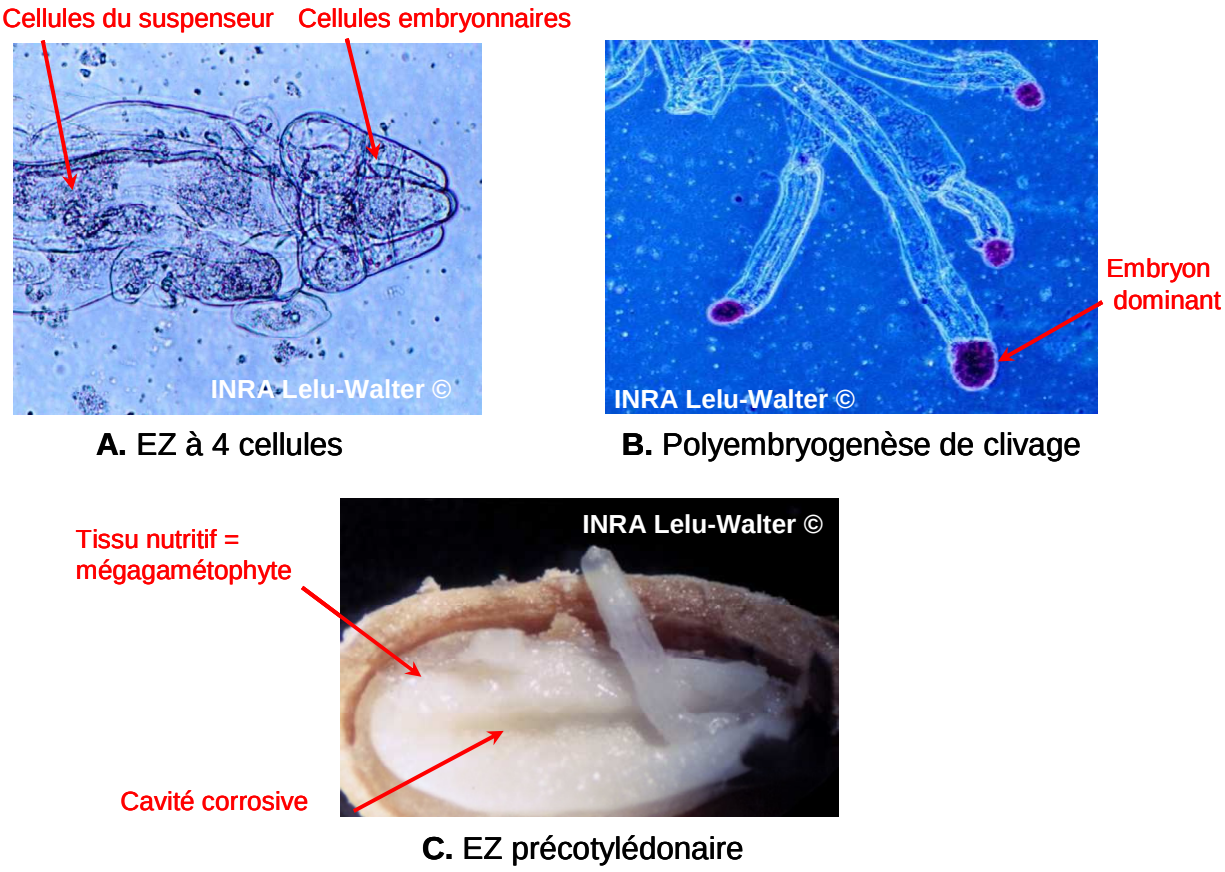
Au cours de l'embryogenèse précoce le suspenseur s'allonge. Contrairement à la division en une seule file de cellules du suspenseur d'*Arabidopsis*, le suspenseur de conifère est constitué de plusieurs files de cellules qui ne se divisent pas et qui sont originaires des cellules proximales de la base de l'embryon (**Figure I.7**). A mesure de son allongement, le suspenseur pousse l'embryon en développement vers la partie centrale du mégagamétophyte. Le suspenseur joue plusieurs rôles, dans le soutien et le positionnement de l'embryon dans la graine, la synthèse de régulateurs de croissance, et en particulier de gibbérellines et la conduction des nutriments à partir du mégagamétophyte vers les cellules de l'embryon (Larkins et al., 1997). Survient alors un processus spécifique des gymnospermes du genre *Pinus* appelé polyembryogenèse de « clivage » : les quatre cellules initiales de l'EZ se mettent sur un plant équatorial, chacune d'entre elles se divisant pour donner quatre EZ secondaires (Dogra, 1967; von Aderkas et al., 1991) (**Figure I.6, I.7**). Cela consiste donc à un processus naturel de clonage. Cependant par la suite l'un des EZ secondaires devient dominant, aboutissant à l'avortement des trois autres embryons, et colonise la cavité corrosive de l'endosperme où il se développe (**Figure I.7**). Sur les dix genres de la famille des *Pinacees*, seuls *Cedrus*, *Keteleeria*, *Pinus* et *Tsuga* présentent une polyembryonie de clivage (Dogra, 1967).

I.4.3. Embryogenèse tardive

L'embryogenèse tardive est caractérisée par une phase d'histogénèse et d'organogenèse le long de l'axe de l'embryon (von Arnold et al., 2002). La partie embryonnaire se divise en deux régions, une région (proximale au suspenseur) dans laquelle les cellules subissent des divisions transversales, et une région (distale par rapport au suspenseur), dans laquelle les cellules se divisent dans tous les plans. Le méristème racinaire se différencie dans la région proximale et bientôt des ébauches racinaires pourront être visibles. La région distale est ensuite délimitée en hypocotyle et méristème caulinaire apical, puis les cotylédons sont formés. Contrairement à *Arabidopsis*, les embryons de conifères développent une couronne de plusieurs cotylédons entourant le méristème caulinaire apical (**Figure I.7**). A ce stade, les tissus provasculaires et corticaux sont différenciés (Vestman et al., 2011).



A **B** **C**



A. EZ à 4 cellules

B. Polyembryogenèse de clivage

C. EZ précotylédonaire

Figure I.7. Représentation schématique du développement de l'embryon zygotique chez les conifères du genre *pinus* **(von Arnold et al., 2002)**. PU - primary upper tier, PE - primary embryonal tier, E - embryonal tier, S - suspensor tier, U - upper tier, EM - embryonal mass, sS - secondary suspensor.

I.4.4. Maturation et déshydratation

Au cours de la maturation, la graine met en place des tissus de réserve, l'objectif étant de préparer l'embryon à la déshydratation et à la dormance, mais aussi d'assurer l'autonomie du jeune plant dans les premières phases de la germination. Parmi les composés de réserve les plus importants se trouvent les protéines, les lipides, les sucres et l'amidon, les rôles multiples

joués par ces différentes molécules étant détaillés dans la partie III. La phase de maturation est considérée comme terminée lorsque l'embryon a accumulé toutes ses réserves. La maturation physiologique est ainsi définie par le moment où la graine atteint son maximum de poids sec, c'est-à-dire le moment où l'accumulation des réserves est optimale pour permettre une bonne germination. En effet la qualité, en termes qualitatif et quantitatif, des composés de réserve accumulés va directement impacter le succès de la germination de la graine et la croissance ultérieure du semis. De façon concomitante à l'accumulation des réserves, une augmentation de la teneur en acide abscissique (ABA) endogène et de la déshydratation de la graine s'effectue progressivement. La majorité des graines angiospermes et de conifères, dont celles de pin maritime, sont de type orthodoxe. Elles subissent en fin de maturation, une intense déshydratation qui peut abaisser le taux interne d'humidité relative de la graine jusqu'à 5%. La viabilité des graines déshydratées peut ensuite être maintenue à basses températures et à teneur en eau la plus faible possible (Engelmann, 1991). La phase de déshydratation de l'embryon est cruciale pour permettre la transition entre la phase de maturation et la phase de germination (Bewley et Black, 1985; Misra et al., 1985; Kermode, 1990, **Figure I.8).**

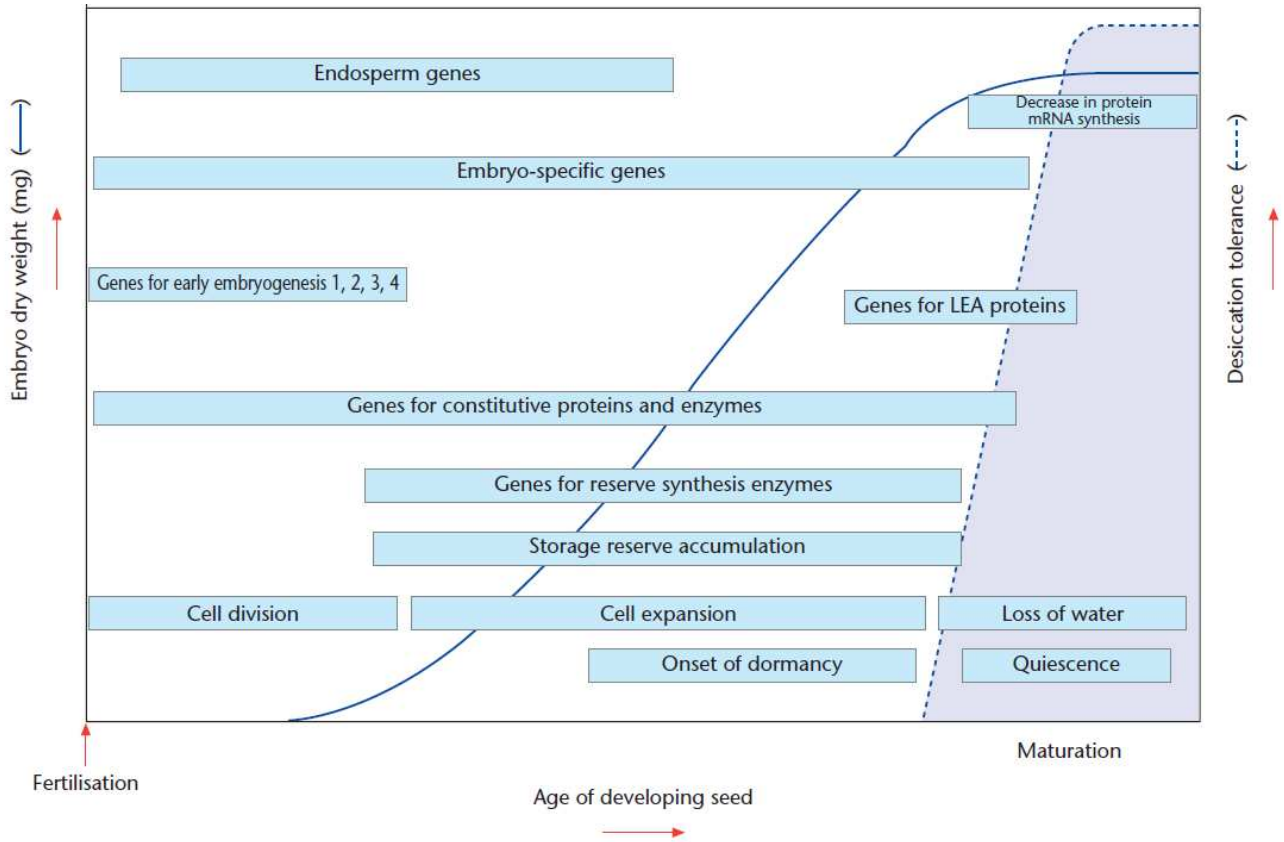


Figure I.8. Résumé des évènements biochimiques et moléculaires survenant lors du développement de la graine chez les angiospermes **(Bewley et al., 2013)**.

I.5. Programme d'amélioration et variétés améliorées du pin maritime

Dans les années soixante, la valeur génétique du matériel reproductif utilisé dans les plantations a été reconnue comme un enjeu important pour la productivité et la qualité du bois sur le long terme. Des essais comparant différentes provenances ont révélé la supériorité de croissance et la bonne adaptation aux sols sableux podzolisés de la région landaise, ayant des caractéristiques variables allant de très sec (dunes le long de la côte), à sec (avec *Caluna vulgaris* et *Erica cinerea*), semi-humide (avec *Pteridium aquilinum*) et humides (avec *Molinia caerulea*) des essences landaises. Les ressources locales de la forêt d'Aquitaine ont ainsi fourni la meilleure base génétique pour construire une population reproductrice. Le programme d'amélioration a commencé dans le but d'améliorer à la fois la croissance et la rectitude du tronc (**Figure I.9**), conjointement à l'INRA (Institut National de la Recherche Agronomique) à au FCBA (ex AFOCEL, Association Forêt-Cellulose). Les deux instituts ont uni leurs forces en 1995, dans le cadre du GIS Pin Maritime du Futur (Groupement d'Intérêt Scientifique) qui comprenait également le CPFA (Centre de Productivité et d'Action Forestière Aquitaine) et l'ONF (Office National des Forêts).



Figure I.9. Troncs flexueux de pins maritimes non-améliorés (**A**) ; Amélioration du pin maritime sur le caractère rectitude **(B).**

Aujourd'hui, 100% des plantations sont faites avec des semences améliorées de la deuxième génération de vergers à graines, et le gain génétique des variétés améliorées est d'environ 30% pour le volume et la rectitude (Alazard et al., 2005, **Figure I.10**). La 3ème génération de vergers à graines a été plantée en 2010. Elle devrait commencer à produire des semences au cours des prochaines années, avec des gains génétiques attendus de 40-45% pour la croissance et la rectitude du tronc par rapport au matériel d'origine (Merzeau et al., 2002, **Figure I.10**). D'autres critères de sélection sont maintenant pris en compte comme la qualité du bois, autant pour le marché du bois que celui des pâtes à papier. Plusieurs observations récentes corroborent cette demande : d'abord le gain de productivité a été suivi d'un abaissement de l'âge de la récolte (de 60 à 40 ans), ce qui a conduit à une récolte en plus grande proportion de bois juvénile provoquant une baisse intrinsèque de la qualité globale du bois (Pot et al., 2002; Bouffier et al., 2009). De plus, les enquêtes préliminaires ont rapporté des corrélations négatives entre la productivité et la qualité (Pot et al., 2002).

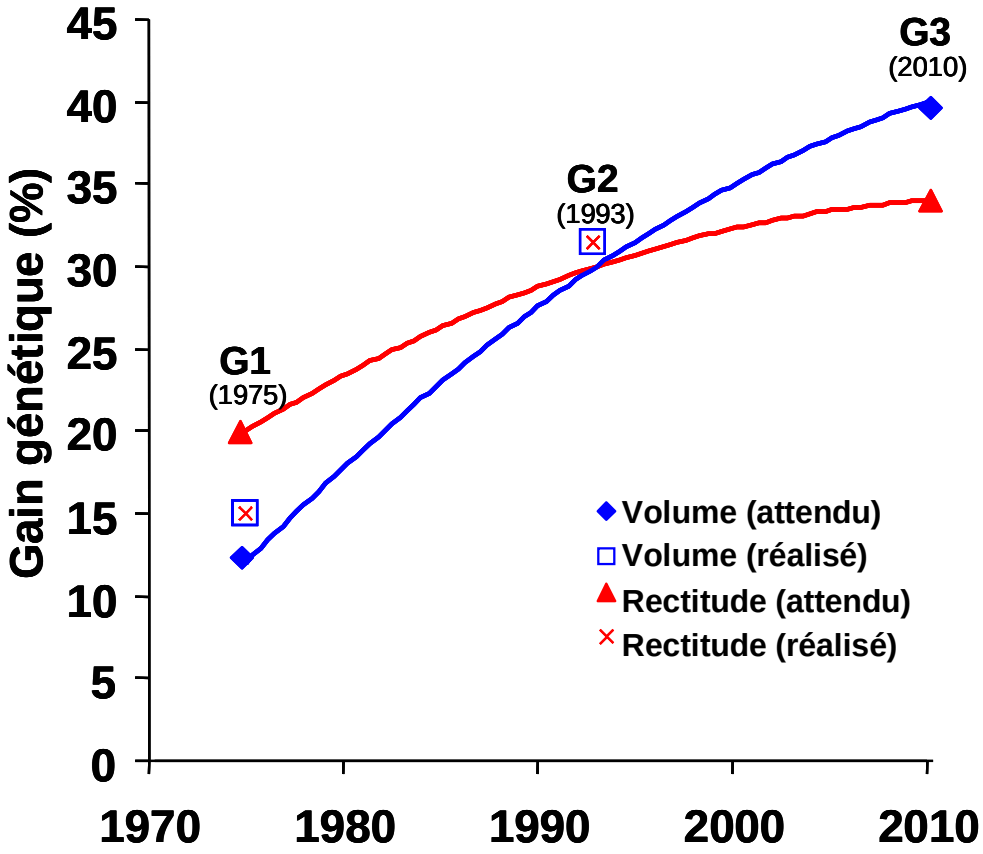


Figure I.10. Gain génétique (%) attendu et réalisé à 15 ans pour la vigueur (volume) et la forme (rectitude basale) de variétés de pin maritime améliorées de 1ère (**G1**), 2ème (**G2**) ou 3ème génération (**G3**). La date d'entrée en production des vergers à graines correspondante est indiquée **(Merzeau et al., 2002)**.

Dans le cas de l'amélioration d'espèces forestières, il est très important d'intégrer la notion de temps. A titre d'exemple, la première étape du cycle de sélection récurrente chez le pin maritime nécessite environ 15 ans. Ces quinze années se répartissent comme ceci : deux années pour la mise en place des tests de descendance et des croisements contrôlés et une dizaine d'années pour pouvoir évaluer correctement les descendants. L'évaluation précoce des descendants peut être effectuée sans attendre l'âge adulte (les pins maritimes sont récoltés entre 35 et 40 ans dans les Landes de Gascogne). Mais en raison d'une mauvaise corrélation juvénile-adulte (Kremer, 1992), il est nécessaire d'attendre un âge minimum de 12 ans pour évaluer la rectitude et le volume. Au final, en comptant une quinzaine d'années supplémentaires pour la deuxième étape du cycle de sélection, une période d'environ 30 ans est nécessaire pour réaliser un cycle complet (**Figure I.11**, Bouffier et al., 2009). Le programme d'amélioration du pin maritime est actuellement bien avancé: la 3ème génération de variétés améliorées est en cours d'évaluation, ce qui est remarquable pour une espèce ligneuse comme le pin maritime qui fructifie tardivement.

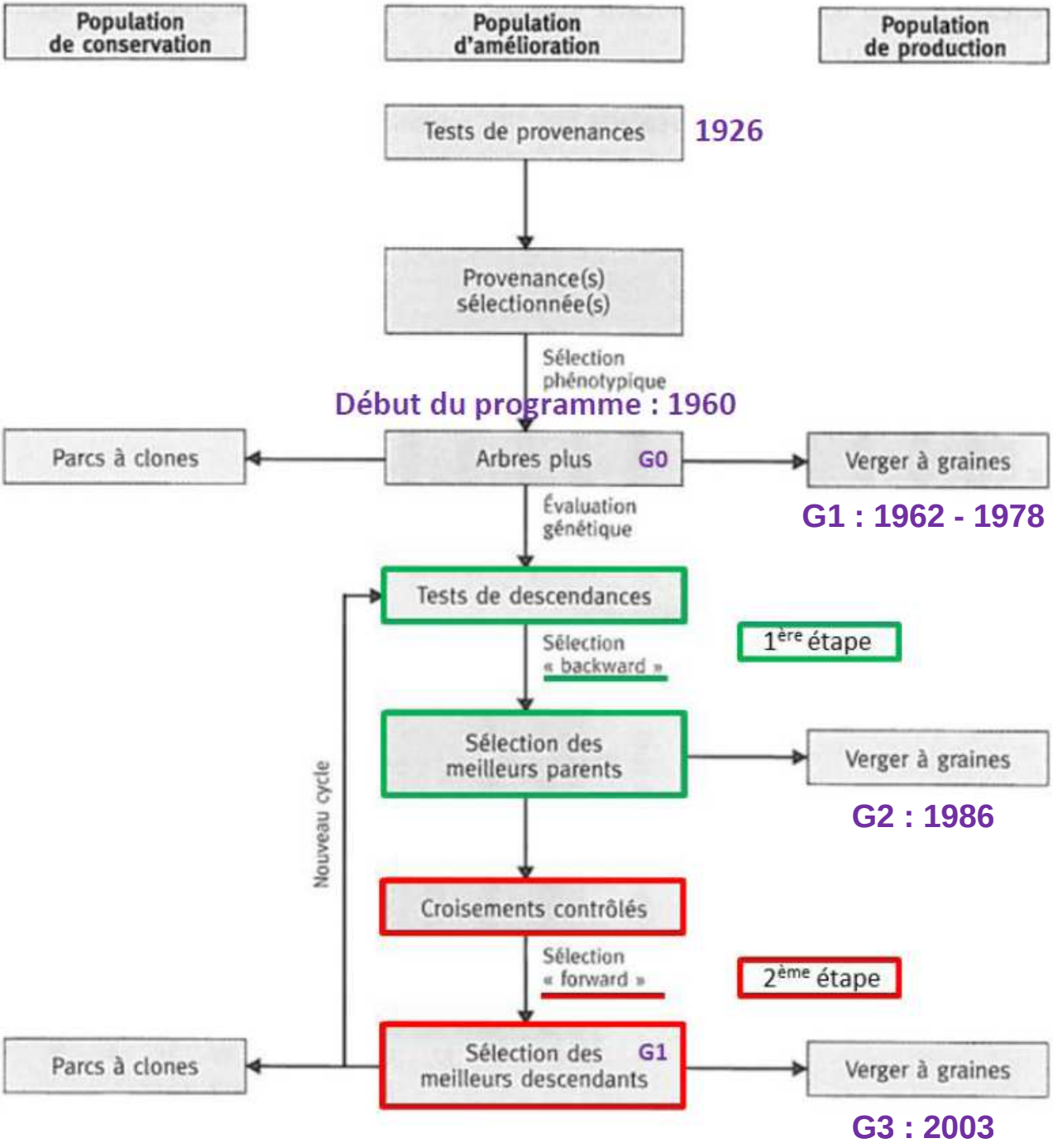


Figure I.11. Représentation schématique du programme d'amélioration du pin maritime **(Bouffier et al., 2009)**.

I.6. Effets post-tempêtes

Dans un contexte post-tempêtes (1999 et 2009) qui se traduit par la destruction d'une partie du capital de production, un déficit sévère de la ressource disponible en Aquitaine est à craindre à l'échéance de 2020. Il a ainsi été estimé que la surface à reboiser impérativement atteint 95100 ha après la tempête Martin de 1999 (plus de 40 % de dégâts ; 24 Mm3 abattus = 3 ans de récolte) et 202 000 ha après la tempête Klaus de 2009 (37 Mm^3 abattus = 5 ans de récolte, Thivolle-Cazat et al., 2013, **Figure I.12**). Ce capital perdu doit dorénavant être reconstitué le plus vite possible. L'intérêt, pour les variétés améliorées, se trouve donc renforcé pour reconstituer ce capital. Aujourd'hui de nouveaux traits de sélection sont envisagés pour augmenter la valeur génétique des variétés améliorées comme la résistance aux stress biotique (ravageurs) et abiotiques (sécheresse, salinité) ou encore la résistance aux épisodes de tempete par la sélection de systèmes racinaires mieux développés (Alazard et Raffin, 2002).

Le pin maritime constitue ainsi une ressource économique importante en Europe, dont le renouvellement est assuré à ce jour par la reproduction sexuée, c'est à dire la production de graines puis semis. Cependant, il faut 2 ans au pin maritime pour réaliser son cycle de reproduction, ce qui constitue une période longue et incompressible pour obtenir des graines. De plus, ce cycle est soumis aux aléas climatiques, ce qui constitue un frein certain à la propagation de l'espèce. Afin de répondre aux réalités économiques qui imposent une demande en bois toujours croissante, il apparait crucial de pouvoir disposer d'outils de propagation alternatifs des variétés améliorés. En effet, il existe d'autres types de dissémination pour le pin maritime, basés sur la multiplication végétative propre aux espèces végétales.



Figure I.12. Effets des tempêtes Martin (1999) et Klauss (2009) sur la forêt des landes.

Partie II - Multiplication végétative du pin maritime

L'utilisation de la multiplication végétative dans le secteur forestier est le moyen le plus rapide, le plus souple et le plus efficace de produire suffisamment de matériel génétiquement amélioré pour répondre aux besoins en bois. La plupart des plantes, y compris les espèces forestières, sont propagées par les graines qui sont faciles à collecter et qui peuvent être stockées pendant de longues périodes. La multiplication végétative possède un certain nombre d'avantages qui en font un outil utile dans les programmes d'amélioration :

- Elle offre une production plus rapide et plus stable de matériel amélioré, car il n'est pas affecté par les saisons de floraison et la production de graines et permet une plus grande flexibilité de déploiement de matériels forestiers
- Elle reproduit exactement une combinaison génétique unique (à la fois en termes additifs et non additifs), à la différence de la reproduction sexuée où seuls les traits additifs sont transmissibles. Ainsi les sources non additives de variation sont facilement accessibles et peuvent être exploitées pour renforcer encore l'intensité de sélection et donc le gain génétique par rapport à la reproduction sexuelle.
- Elle est bien insérée au sein de schémas de sélection précoce, en particulier ceux basés sur l'utilisation de marqueurs moléculaires seulement aux premiers stades de développement, permettant un raccourcissement de l'intervalle de génération entre deux cycles de sélection consécutifs.
- Elle fournit une méthode de tests puissants pour l'évaluation génétique avec une estimation plus simple du microenvironnement et des interactions G x E (génotype par environnement) que dans les tests de descendance. Il a été montré que le test clonal est plus efficace que le test conventionnel de descendance des semis (Isik et al., 2003).
- Elle permet d'utiliser des clones pour mieux évaluer les génotypes, la plasticité phénotypique, ou pour raccourcir la durée du cycle de reproduction.
- Elle peut produire du matériel clonal qui croît de manière uniforme, ce qui rend la gestion plus prévisible, ainsi que la production d'un produit de meilleure qualité (bois ou fibre de bois) impactant directement les coûts de récolte et de transformation. L'uniformité du produit pourrait avoir un impact plus fort sur les coûts que l'augmentation du taux de croissance (Shelbourne, 1997).

Afin de répondre plus efficacement aux demandes de reboisement, un certain nombre de techniques basées sur la capacité de multiplication végétative naturelle des plantes, a été développé pour le pin maritime. Le principe est de pouvoir régénérer un individu entier à

partir d'organes d'un parent donneur, il s'agit donc de clonage. Le pin maritime peut être multiplié indifféremment par bouturage, greffage et par les biotechnologies de propagation *in vitro* (embryogenèse somatique).

II.1. Bouturage horticole

La multiplication peut se réaliser soit à partir de rameaux vigoureux prélevés sur l'arbre soit à partir de pousses interfasciculaires développées après une taille. La taille consiste à éliminer des pousses âgées de 2 à 3 ans de manière à provoquer le développement des brachyblastes. Afin d'initier la rhizogenèse, la base des boutures est plongée dans une solution hormonale d'Acide Indole Butyrique (AIB). Quarante jours après la mise en place des boutures, les racines apparaissent et les plants peuvent alors être élevés. La réussite du bouturage dépend de nombreux facteurs (qualité et origine génétique du matériel végétal, contrôle de l'ambiance de bouturage...) mais c'est l'âge des arbres ou des pied-mères qui constitue le principal facteur de réussite du bouturage : d'une manière générale pour les ligneux, le pourcentage d'enracinement et sa qualité diminue fortement avec l'âge. En effet le problème majeur de la multiplication végétative d'arbres âgés est le vieillissement des plants, se traduisant par une baisse du potentiel de rhizogenèse et par une croissance faible et souvent plagiotrope des plants racinés (phénomène de maturation). Ce phénomène récurrent est en lien avec le vieillissement ontogénétique (ou changement de phase depuis le stade juvénile jusqu'aux stades adultes végétatif et reproducteur) qui entraine une perte des capacités organogénétiques. Les arbres de pin maritime âgés sont, comme chez beaucoup d'espèces forestières et la plupart des conifères, récalcitrants au bouturage. Pour ces raisons, la multiplication de clones de pin maritime à des fins de reboisement n'a jamais abouti.

II.2. Greffage horticole

Le greffage est une technique de multiplication végétative qui est utilisée depuis longtemps chez le pin maritime. En effet, les premières applications à grande échelle de cette technique remontent à l'initiation du programme d'amélioration pour la sauvegarde et le transfert en parcs à clones des « arbres + » sélectionnés en forêt. Ainsi, toute la population d'amélioration a pu être transférée et sauvegardée grâce à la technique du greffage. A l'heure actuelle, le greffage est encore utilisé soit pour la création de vergers à graines, soit pour le transfert et la sauvegarde du matériel sélectionné.

Son principe est de récolter un greffon sur un arbre (dont on veut conserver le patrimoine) et de l'apposer sur un porte greffe constitué par un jeune plant de deux ans. Les meilleurs greffons sont les rameaux végétatifs ou ceux qui porteront les cônes femelles. Le greffage est réalisé au printemps lorsque le porte greffe est en pleine sève après débourrement végétatif très net. Un mois environ après le greffage, le greffon commence à se développer. Lorsque les conditions favorables sont réunies (matériel adéquat, bon état du porte greffe, conditions climatiques satisfaisantes et greffeur expérimenté), le taux de reprise avoisine les 90%. Le greffage est une opération minutieuse, longue et coûteuse à ce jour réservée à des opérations spécifiques (création de banque de clones, de vergers à graines de clones).

Ainsi le greffage et le bouturage du pin maritime sont des techniques limitées par leur faible rendement de production et leurs coûts parfois prohibitifs. Il convient aujourd'hui de se tourner vers d'autres méthodes de multiplication végétative plus performantes.

II.3. Embryogenèse somatique

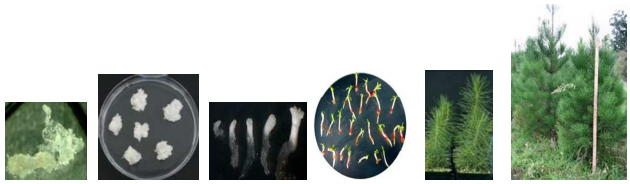
L'embryogenèse somatique consiste en la production d'embryons à partir de cellules somatiques (cellules du soma non reproductrices) sans passer par la multiplication sexuée (cellules germinales). On parle alors d'embryons « somatiques ».

L'embryogenèse somatique a été réalisée pour la 1ère fois chez la carotte (Steward et al., 1958). Depuis, l'avancée des connaissances en matière d'embryogenèse somatique n'a eu de cesse de progresser chez les angiospermes (Yang et Zhang, 2010; Helal, 2011; Rai et al., 2011). Chez les gymnospermes, plus récalcitrants à l'embryogenèse somatique (Bonga et al., 2010), il aura fallu attendre près de 30 ans pour obtenir un résultat similaire chez l'épicéa commun, *Picea abies* (Hakman et al., 1985). Depuis plusieurs études très complètes portant sur développement des ES ont apporté des progrès dans la maîtrise de l'embryogenèse somatique des conifères, et ont permis de combler une certaine forme de retard vis-à-vis des autres espèces forestières (Tautorus et al., 1991; Attree et Fowke, 1993; von Aderkas et Bonga, 2000; von Arnold et al., 2002; Stasolla et Yeung, 2003; Cairney et Pullman, 2007; Klimaszewska et al., 2007).

Chez les conifères, le matériel végétal utilisé pour l'initiation de l'embryogenèse somatique est l'EZ à un stade immature (Attree et Fowke, 1993). Concernant le genre *Pinus*, les espèces les plus étudiées sont *Pinus taeda* aux USA (Pullman et al., 2003), *Pinus radiatata* en Nouvelle-Zélande (Bishop-Hurley et al., 2003) et *Pinus pinaster* et *Pinus sylvestris* en Europe (Klimaszewska et al., 2007; Lelu-Walter et al., 2008).

Chez le pin maritime, l'embryogenèse somatique a été réalisée pour la première fois au cours d'un travail de thèse par Jarlet-Hugues (Jarlet-Hugues, 1989). La régénération de plants n'a été effective que 10 ans plus tard (Lelu et al., 1999). Depuis, des recherches approfondies ont été entreprises en France (Ramarosandratana et al., 2001a; Jordy et Favre, 2003; Breton et al., 2005; Lelu-Walter et al., 2006; Park et al., 2006; Klimaszewska et al., 2009; Trontin et al., 2013), en Espagne (Perez Rodriguez et al., 2006 ; Humanez et al., 2012) et au Portugal (Miguel et al., 2004) (**Table I.2**). L'embryogenèse somatique du pin maritime est un processus complexe qui se déroule au travers de plusieurs étapes détaillées ci-après (**Figure I.13**).

Table I.2. Historiques des avancées sur la maitrise des différentes étapes de l'embryogenèse somatique du pin maritime. **1.** Initiation ; **2.** Prolifération ; **3.** Maturation ; **4.** Germination ; **5.** Plants ; **6.** Mise en champs.
+ Étape maitrisée ; **(q)** Etude de la qualité de l'embryon somatique ; **(2)** Etude de l'embryogenèse somatique secondaire.



Années	Références	1	2	3	4	5	6
1989	(Jarlet-Hugues, 1989)	+					
1995	(Bercetche et Pâques, 1995)	+	+				
1999	(Lelu et al., 1999)	+	+	+	+	+	
2001	(Ramarosandratana et al., 2001a, b)		+	+	+		
2003	(Jordy et Favre, 2003)			+ **(q)**	+		
2004	(Miguel et al., 2004)	+	+	+			
2005	(Breton et al., 2005)		+				
2006	(Breton et al., 2006)			+ **(q)**			
2006	(Perez Rodriguez et al., 2006)	+	+	+	+	+	
2006	(Lelu-Walter et al., 2006)	+	+	+	+	+	
2006	(Park et al., 2006)	+	+	+ **(q)**			
2009	(Klimaszewska et al., 2009)		+ **(q)**	+ **(2)**			
2012	(Humanez et al., 2012)	+	+	+			
2013	(Trontin et al., 2013)	+	+	+	+	+	+

II.3.1. Initiation des masses embryogènes

Pour le pin maritime l'embryogenèse somatique est initiée à partir d'EZ immatures à un stade de polyembryogenèse de clivage **(Figure I.13.1**). Ce stade de développement, optimal pour obtenir l'embryogenèse somatique, est limité à quelques jours (7-10 jours, début juillet en France). En général les cultures embryogènes de conifères, ne sont constituées que d'ES très immatures, d'où le terme de masses embryogènes (MEs). Les MEs provenant d'un même explant et partageant donc le même génotype, forment une « lignée » constituée d'embryons immatures indifférenciés. Les milieux d'initiation nécessitent la présence d'une auxine (l'acide 2,4-dichlorophenoxyacetique, 2,4-D) et d'une cytokinine en général la benzyl adénine (BA) à faibles concentrations (0.5 mg L^{-1}, Lelu et al 2006). La Cytokinine-1-(2-chloro-4-pyridyl)-3-phenylurée (CPPU, Park et al 2006) a permis une amélioration de 77% du taux d'initiation moyen (Park et al., 2006). Les MEs peuvent aussi être obtenues en l'absence totale de régulateur de croissance (Lelu et al 1999). Cependant cette absence conduit à des cultures asynchrones rendant le contrôle de la régénération d'embryons cotylédonaires plus difficile.

II.3.2. Multiplication des masses embryogènes

Après obtention, les MEs sont repiquées sur un milieu favorisant la multiplication des ES (**Figure I.13.2**). La multiplication des MEs est réalisée sur milieu supplémenté en auxine (2,4-D, 0.5 mg L^{-1}) et cytokinine (BAP, 0.5 mg L^{-1}). Cette étape de multiplication permet l'amplification du matériel embryogène. Les accroissements relatifs moyens en matière fraîche estimés pour 6 lignées embryogènes de 3 familles cultivées sont supérieurs à 1000 % sur 2 semaines, il est ainsi possible d'obtenir 1kg de MEs à partir d'un gramme initial de matériel en 6 semaines (Trontin et al., 2013). C'est donc cette étape de multiplication qui fait la force de l'embryogenèse somatique. Néanmoins le vieillissement des MEs en multiplication, se traduit rapidement (en 6 mois de temps) à une chute progressive du potentiel embryogène (aptitude des MEs à régénérer des ES cotylédonaires, Klimaszewska et al., 2009). Ce phénomène semble bien spécifique aux pins car il apparaît beaucoup plus limité chez d'autres conifères (épicéa, mélèze). Des phénomènes épigénétiques (qui ne modifient pas la structure de base de l'ADN) en seraient responsables, comme la méthylation de l'ADN, la modification de la structure de la chromatine ou la régulation par l'expression d'ARNs interférants. Ce problème est techniquement contourné *via* la bonne gestion à long terme des stocks de MEs cryoconservées dans un état juvénile.



Figure I.13. Etapes de l'embryogenèse somatique du pin maritime.

II.3.3. Cryoconservation des masses embryogènes

Les MEs peuvent être facilement cryoconservées dans l'azote liquide (-196°C, Lelu-Walter et al 2006). Les très basses températures permettent de figer non seulement l'eau libre dans les cellules mais également l'eau liée aux molécules. Toute réaction métabolique est ainsi stoppée, en particulier les phénomènes de vieillissement. La décongélation des MEs ne pose pas de problème de reprise de croissance (100 % de réactivation). Il est donc possible

d'établir des collections statiques de milliers d'individus dans un état juvénile au sein de cryo-conservateurs qui occupent quelques m² seulement de surface utile. Une telle collection peut être maintenue à très long terme dans l'azote liquide. La bonne gestion des échantillons permet de garantir le niveau du stock cryoconservé pour chaque lignée embryogène. Depuis 1993, une importante collection de lignées embryogènes (2000) de génotypes élites a déjà été cryoconservée (Trontin et al. 2013).

II.3.4. Développement et maturation de l'embryon somatique cotylédonaire

Lors de la maturation, les embryons immatures des MEs se différencient en embryons précotylédonaires dans un premier temps, puis en embryons cotylédonaires (de couleur blanche avec des cotylédons plus ou moins développés) comportant chacun un méristème racinaire et un méristème caulinaire.

Chez le pin maritime, le développement des ES cotylédonaires et leur maturation se déroulent en 12 semaines (**Figure I.13.3**). De nombreux paramètres contrôlent le développement des ES cotylédonaires (Lelu et al., 1999; Ramarosandratana et al., 2001a) :

- L'acide absicissique (ABA) est en général utilisé à de fortes concentrations allant de 80 µM (Lelu-Walter et al., 2006; Ramarosandratana et al., 2001b) à 120 µM (Miguel et al., 2004).
- Le potentiel osmotique du milieu est le plus souvent élevé par augmentation de la concentration de saccharose (0.2 M) (Ramarosandratana et al., 2001a; Lelu-Walter et al., 2006).
- La concentration du gélifiant du milieu est élevée variant de 9 à 12 g L^{-1} (Phytagel™, Sigma, voir Annexes) (Lelu et al., 1999; Ramarosandratana et al., 2001). Chez les pins, ce paramètre est indispensable au bon développement d'ES cotylédonaires. La forte teneur en gélifiant augmente la dureté du milieu et diminue la disponibilité en eau (potentiel matriciel élevé, Klimaszewska et al., 2000).

En dépit des fortes concentrations en ABA et en gélifiant, plusieurs phénomènes viennent entraver la régénération des ES cotylédonaires. Tout d'abord, le potentiel embryogène, varie selon les lignées embryogènes (Ramarosandratana et al., 2001b). Il existe un effet génétique quant à l'aptitude des lignées à l'embryogenèse somatique (Lelu-Walter et al., 2006). Ensuite, le vieillissement des lignées embryogènes en multiplication, entraîne une chute progressive mais inexorable de leur potentiel embryogène qui peut être perdu après 6 mois de multiplication des MEs. Cela se traduit par une incapacité des lignées à régénérer des ES. Ce vieillissement des MEs affecte non seulement la quantité mais aussi la qualité des ES

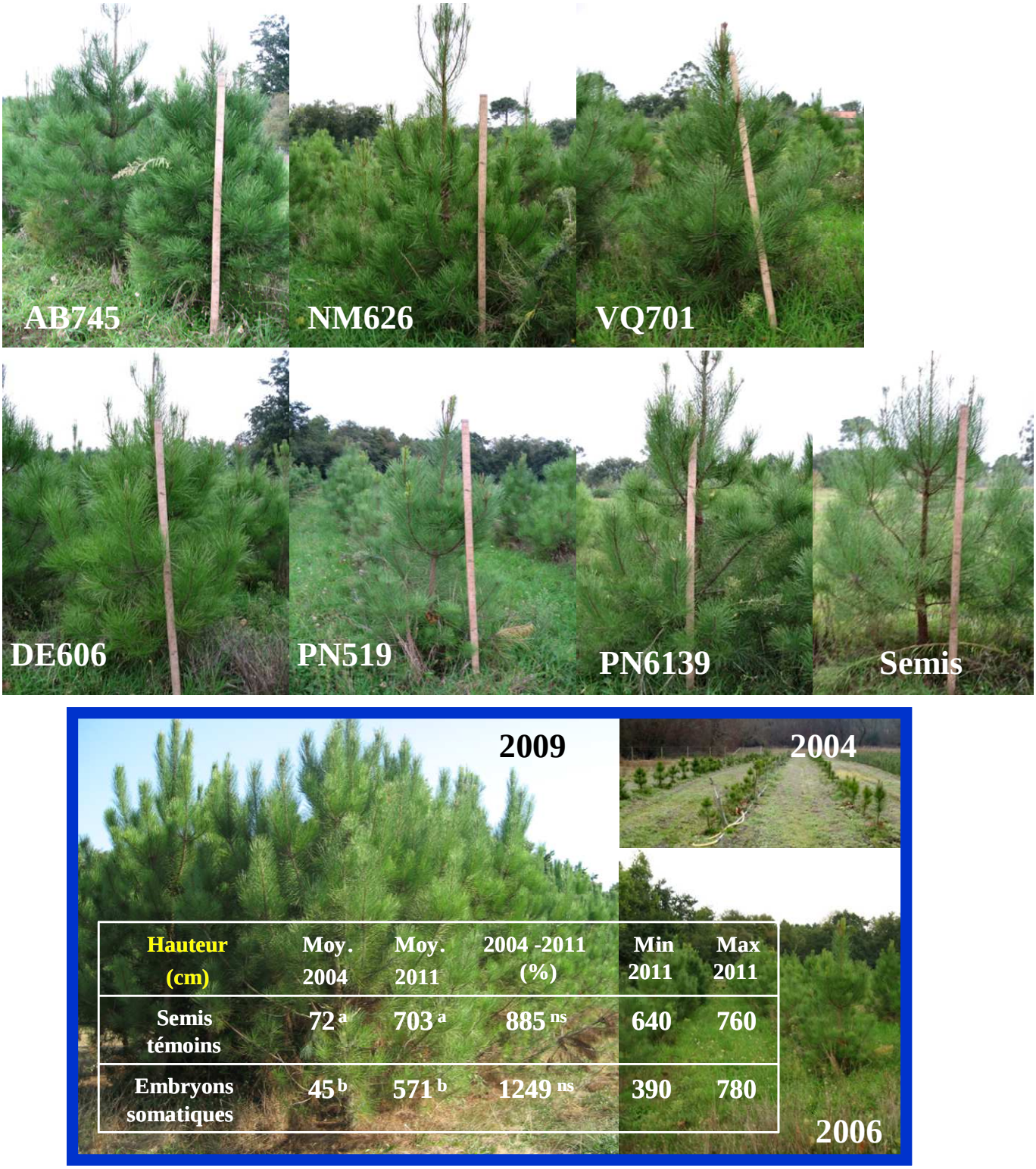
régénérés qui voit leur taille diminuer (Breton et al., 2005). La sauvegarde des cultures embryogènes par cryoconservation revêt donc ici toute son importance. Enfin lors des premières semaines de maturation, les MEs en culture prolifèrent encore très fortement processus antagoniste à la régénération d'ES. Par conséquent, l'étape de maturation du pin maritime n'est pas à ce jour maîtrisée. Plus de recherches sont nécessaires pour optimiser la capacité de différenciation des MEs.

II.3.5. Germination de l'embryon somatique cotylédonaire et développement en plant

Un nombre plus restreint d'études a été réalisé sur la germination et la conversion en plantes des ES chez *P. pinaster*. Ceci peut être expliqué par la difficulté à régénérer de grandes quantités d'ES. À l'heure actuelle, les ES sont récoltés sur milieu de maturation à partir de 12 semaines, et sont choisis en fonction de leur morphologie. La germination est réalisée sur un milieu solide sans régulateur de croissance (**Figure I.13.4**). Il a été montré que l'obscurité ou une diminution de la luminosité est bénéfique pour les cultures d'ES dans les 7 à 10 premiers jours avant exposition à la lumière chez *P. strobus* et *P. pinaster* (Klimaszewska et al., 2001; Lelu-Walter et al., 2006). La fréquence moyenne de germination est assez élevée, (78% comme cela a été déterminé à partir de plus de 2000 ES cotylédonaires). Cependant, il existe des variations considérables du taux de germination entre les ES des différentes lignées (allant de 61% à 96%). L'allongement de l'hypocotyle est visible après 6 à 7 semaines sur milieu de germination. Les plants cultivés durant 14 et 16 semaines sur milieu de germination sont transférés depuis les boîtes de Pétri en conditions contrôlées sous serre (**Figure I.13.5**, Lelu-Walter et al., 2006). La survie moyenne des plantes acclimatées après 4 mois est de 65%, variant de 36% à 100% selon le génotype.

II.3.6. Essais au champ de plants de pin maritime

Dès 1999 les équipes de FCBA et de l'INRA ont été capables de produire conjointement des ES cotylédonaires (**Table I.2**, Lelu et al. 1999). Depuis, un objectif constant est d'estimer, au travers d'essais au champ, si les plants régénérés se développent normalement, fructifient et produisent des graines. Le plus ancien essai, d'ampleur limitée (12 clones FCBA, 59 plants en pépinière) date de 1999. Lors de la dernière mesure à 12 ans (2011) les arbres somatiques mesuraient de 6,50 à 12,10 m pour un diamètre de 9 à 29 cm (**Figure I.14**).



Hauteur (cm)	Moy. 2004	Moy. 2011	2004 -2011 (%)	Min 2011	Max 2011
Semis témoins	72 [a]	703 [a]	885 [ns]	640	760
Embryons somatiques	45 [b]	571 [b]	1249 [ns]	390	780

Figure I.14. Phénotype d'un semis témoin de pin maritime et de plants obtenus par embryogenèse somatique (6 clones) après 2,5 ans de plantation (essai 2004, Sivaillan). Le

tableau donne les hauteurs moyennes, minimales et maximales observées ainsi que l'accroissement relatif entre 2004 et 2011. [a] et [b] : Différences significatives. [ns] : non significatif **(Trontin et al., 2013)**.

Les dimensions de certains arbres somatiques étaient comparables aux semis témoins de la parcelle (hauteur : 9,60 à 11,20 m ; diamètre : 19 à 33 cm). Les premières fructifications ont été notées dès l'âge de 5 ans pour les plants issus d'embryons somatiques et les semis témoin. Cet essai démontre donc que les plants somatiques présentent des phases juvénile et adulte végétatives normales et entrent dans la phase adulte reproductive (**Figure I.14**). Les plants somatiques ont néanmoins une vigueur initiale plus faible affectant les paramètres de productivité. Cependant, les arbres issus d'ES présentant les meilleurs taux de croissance pourraient être capables de rattraper leur retard.

II.3.7. Place de l'embryogenèse somatique dans les programmes d'amélioration

L'embryogenèse somatique est devenue la méthode de choix pour la propagation végétative de conifères (Sutton, 2002) en raison de son taux de multiplication élevé et du maintien quasi infini de la juvénilité du materiel embryogène par cryoconservation : les masses embryogènes à un stade très juvénile sont conservables à long terme par congélation à très basse température sous azote (-196°C). L'embryogenèse somatique peut être intégrée au programme d'amélioration offrant un outil pour les tests clonaux et pour la propagation rapide de variétés améliorées.

II.3.7.1. Appui des biotechnologies pour maximiser le gain génétique

La forte inertie du programme d'amélioration génétique, inhérente au temps de génération relativement long du pin maritime (15-20 ans), constitue une contrainte forte et spécifique des arbres qui implique de maximiser le gain génétique obtenu à chaque cycle de sélection. Actuellement, la cadence d'entrée en production des vergers à graines améliorées est de 17-18 ans (**Figure I.11**). Même s'il est envisagé que cette cadence augmente en lien avec la croissance améliorée des géniteurs (la 4ème génération de verger est prévue pour 2020), une érosion des prévisions de gains génétiques dès la 3ème génération est attendue (Merzeau et al., 2002). Il est donc nécessaire de dépasser les limitations des méthodes d'amélioration conventionnelles, en les couplant à des biotechnologies afin d'accélérer et optimiser les processus d'amélioration. Dans le cas du pin maritime, plusieurs biotechnologies très prometteuses sont en développement depuis 10 à 20 ans au FCBA et à l'INRA. Il s'agit des technologies d'analyse à haut débit de la structure et de l'expression du génome (génomique,

transcriptomique, protéomique), et de propagation conforme des génomes améliorés (clonage par embryogenèse somatique, Trontin et al., 2013).

II.3.7.2. Association de l'embryogenèse somatique et de la cryoconservation

L'embryogenèse somatique est donc initiée à partir de graines immatures. Par conséquent les performances potentielles des plants régénérés en champs sont inconnues. Il est nécessaire d'établir des tests clonaux pour évaluer les performances des lignées initiées, mais ces tests sont longs, d'autant plus si les caractères évalués s'expriment tardivement (comme la qualité du bois). De plus le matériel clonal testé une fois arrivé à l'âge adulte, où les traits analysés sont exprimés, a déjà subi un vieillissement ontogénétique empêchant toute régénération par multiplication végétative. Alors comment valoriser les résultats de ces tests clonaux ? L'intérêt de disposer d'une méthode de conservation performante et sure de la juvénilité tissulaire prend ici tout son sens. La multiplication d'un embryon, dont nous ne connaissons pas à priori les performances au champ (contrairement aux parents), n'aurait qu'un intérêt limité si les MEs ne pouvaient être conservées à long très terme par cryoconservationon. Dans un programme d'amélioration bien engagé, comme celui du pin maritime, la possibilité de sécuriser les ressources génétiques en cryocollections dans un état très juvénile pendant toute la phase de test clonal sur le terrain (10 ans environ) permet de bénéficier rétroactivement d'un stock de tissu embryogène pour chacun des génotypes finalement sélectionnés pour leurs performances (**Figure I.15,** Trontin et al., 2013**).** Il est alors possible d'envisager la multiplication de ces arbres élites par ES pour contribuer aux sorties variétales. C'est en ce sens que la multiplication végétative est actuellement un objectif crédible chez le pin maritime comme chez d'autres arbres forestiers en Europe (Lelu-Walter et al., 2013).

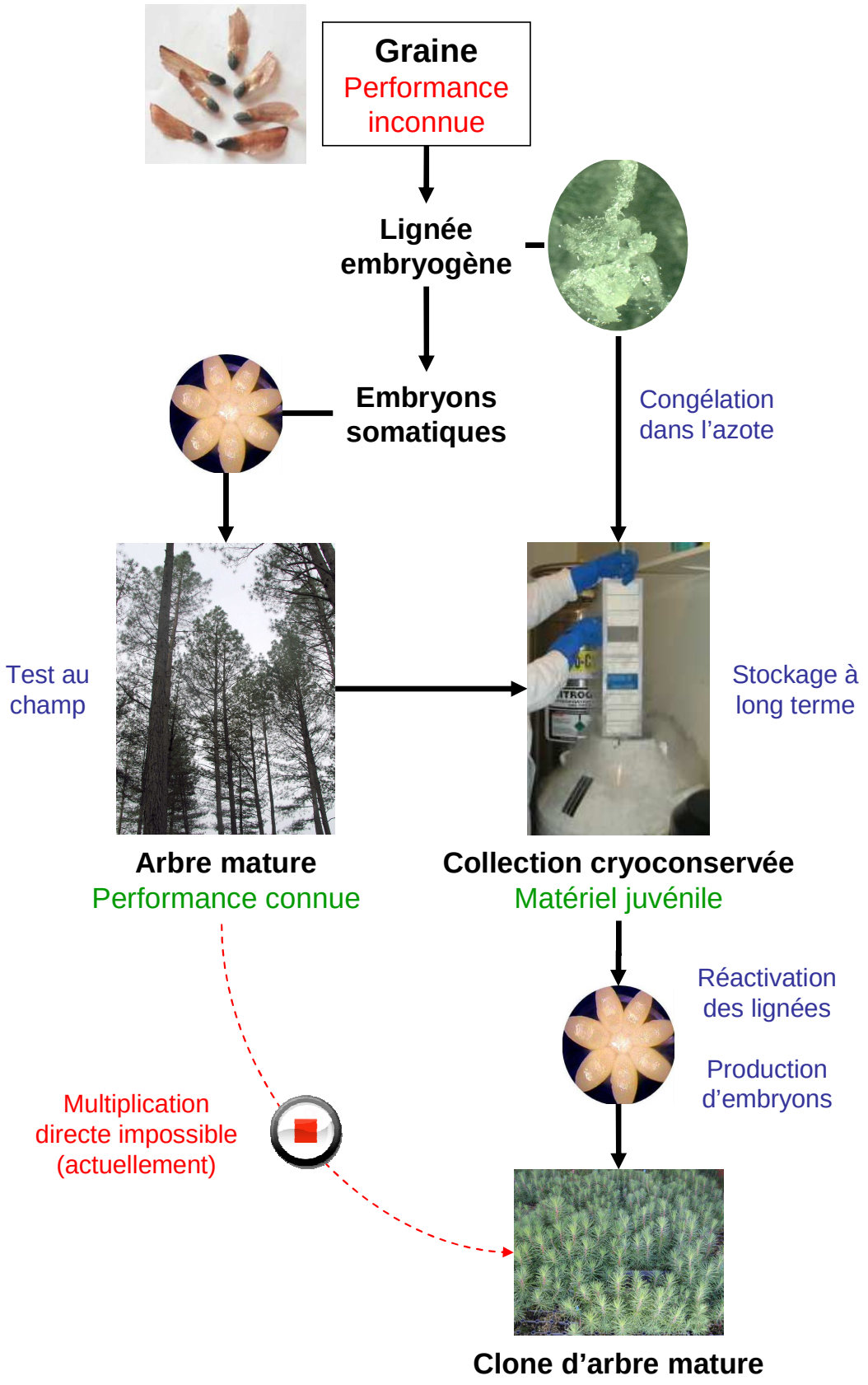


Figure I.15. Appui de l'embryogenèse somatique et de la cryoconservation pour le clonage rétroactif d'arbres matures de pin maritime **(Trontin et al., 2013)**.

Partie III - Marqueurs moléculaires du développement des embryons somatiques et zygotiques

L'EZ constitue la référence en matière de développement embryonnaire, référence à approcher afin d'optimiser le développement de l'ES. Une meilleure compréhension des différences physiologiques existant entre le développement de l'EZ et celui de l'ES pourrait être obtenue grâce à l'analyse des profils d'expression génique. Cela pourrait contribuer à l'amélioration des protocoles *in vitro* (Bonga et al., 2010). Les analyses d'expression génique peuvent aussi être un outil utile pour le suivi du développement de l'embryon, depuis l'acquisition de la compétence embryonnaire jusqu'à la maturation (Feher et al., 2003). Ainsi, l'identification, dans un premier temps, de marqueurs pouvant discriminer la qualité des embryons de différentes lignées cellulaires, en termes de potentiel embryogène et en qualité de maturation, permettrait la sélection de lignées embryogènes plus proches des EZ. Dans l'idéal, ces gènes « marqueurs » devraient être universels, sensibles, détectables dans de petites quantités de tissu, et révélateurs de processus spécifiques qui caractérisent chacune des étapes ou des transitions au cours du développement (Lara-Chavez et al., 2012).

De nombreux gènes nécessaires au développement de l'embryon ne sont pas spécifiques de l'embryogenèse et leurs fonctions basales sont nécessaires tout au long du cycle de vie de la plante, ce sont les gènes appelés « de ménage » (Tzafrir et al., 2004). Sur 16 000 gènes actifs tout au long du développement des graines d'Arabidopsis, seulement 289 sont spécifiques de la graine (Le et al., 2010).

L'expression des gènes peut être visualisée par la présence des transcrits et des protéines qui en découlent. La **Table I.3** recense un nombre non-exhaustif de transcrits ou de protéines caractérisés au cours des différentes étapes de développement des ES et/ou des EZ chez les conifères.

Table I.3. Marqueurs moléculaires caractéristiques des embryogenèses (somatique et zygotique) des conifères dont *Pinus pinaster*. Etudes identifiant des marqueurs protéiques par protéomique 2D (en bleu) ; études portant sur le pin maritime (en rouge).

Stade	fonction	gène/protéine	nature	technique de détection	espèce	Expression	références
Proembryogenèse	contrôle de la division cellulaire	STIMPY/WOX9	facteur de transcription à homéodomaine	criblage de banque de clones ADNc	*Pinus taeda*	spécifique des cultures embryogènes	(Cairney et al., 2006)
	métabolisme de l'azote	GS1b	Glutamine synthetase	northern blot assay + hybridation in situ	*Pinus pinaster et Pinus sylvestris*	ES précoce	(Perez Rodriguez et al., 2006)
	modification de la paroi cellulaire	β-expansine	expansine	crible de banque d'ADNc	*Pinus radiata*	masses proembryogènes de stade I	(Bishop-Hurley et al., 2003)
	modification de la paroi cellulaire	cellulase	hydrolase	crible de banque d'ADNc	*Pinus radiata*	masses proembryogènes de stade I	(Bishop-Hurley et al., 2003)
	modifications de la paroi cellulaire	germine-like protéine	germine	crible de banque d'ADNc	*Pinus radiata*	masses proembryogènes de stade I	(Bishop-Hurley et al., 2003)
	modifications de la paroi cellulaire	PcGER1	germine	criblage de banque de clones ADNc	*Pinus caribaea*	spécifique des lignées cellulaires embryogènes	(Domon et al., 1995)
	réponse aux stress	cytochrome P450	monooxygénase/oxydoréductase	crible de banque d'ADNc	*Pinus radiata*	masses proembryogènes de stade I	(Bishop-Hurley et al., 2003)
	signalisation cellulaire	arabinogalactan protein (AGP)	polypeptides glycosylés	réctif β-glucosyl Yariv + anticorps monoclonaux	*Picea abies*	spécifique des cultures embryogènes	(Egertsdotter et Arnold, 1995)
	signalisation cellulaire	*PgChi-1*	basic class IV chitinase	crible de banque d'ADNc	*Picea glauca*	spécifique des cultures embryogènes	(Dong et Dunstan, 1997b)
	signalisation cellulaire	chitinase basique	chitinase	SDS-PAGE	*Pinus caribaea*	spécifique des cultures embryogènes	(Domon et al., 2000)
	signalisation cellulaire	lipochitooligosaccharides (LCOs)	facteur NOD	western blot	*Picea abies*	spécifique des cultures embryogènes	(Egertsdotter et von Arnold, 1998)
	signalisation cellulaire	*PgGlu-1*	β-1,3-glucanase	crible de banque d'ADNc	*Picea glauca*	spécifique des cultures embryogènes	(Dong et Dunstan, 1997b)
	transduction du signal	AaSERK1 (SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1)	leucine-rich repeat receptor-like kinase	RT-PCR	*Araucaria angustifolia*	cellules embryogènes	(Steiner et al., 2012)

Stade	fonction	gène/protéine	nature	technique de détection	espèce	expression	références
Embryogenèse précoce	voie de signalisation régulée par l'auxine	PIN protéine, polar auxin inhibitor N-1-naphthylphthalamic acid (NPA)	efflux-facilitating protein transport	immunolocalisation	*Picea abies*	embryogenèse somatique précoce	(Hakman et al., 2009)
	biosynthèse de l'auxine	small auxin-up RNA (SAUR)	ARN interférant	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique précoce	(Vestman et al., 2011)
	contrôle du développement des méristèmes	CLAVATA1-LIKE	leucine-rich-repeat protein receptor kinase (LRR RLK)	qRT-PCR	*Pinus pinaster* et *Pinus pinea*	premières phases de la caulogenèse adventive dans les cotylédons	(Alvarez et al., 2013)
	contrôle de la croissance cellulaire radiale	Scarecrow	facteur de transcription	semi-quantitative RT-PCR	*Araucaria angustifolia*	embryogenèse somatique précoce	(Schlögl et al., 2012)
	contrôle de la division cellulaire	WUSCHEL RELATED HOMEOBOX 2 (WOX 2)	facteur de transcription à homéodomaine	microarray + qRT-PCR	*Picea abies*	exprimé pendant la différenciation précoce des embryons	(Vestman et al., 2011)
	contrôle de la division cellulaire	PaWOX2 et PaWOX8/9	facteur de transcription à homéodomaine	hybridation in situ	*Picea abies*	embryogenèse précoce ES et EZ	(Palovaara et al., 2010)
	contrôle du développement des téguments	NAC-regulated seed morphology 2 (NARS2)	facteur de transcription	microarray + qRT-PCR	*Picea abies*	différenciellement exprimé pendant la différenciation précoce des embryons	(Vestman et al., 2011)
	défense contre le stress oxydatif, détoxification	thioredoxine	oxydoréductase	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	défense contre le stress oxydatif	thioredoxin H9-like	oxydoréductase	protéomique 2D	*Pinus massoniana*	embryogenèse zygotique précoce	(Zhen et al, 2012)
	défense contre le stress oxydatif, détoxification	ascorbate peroxidase	oxydoréductase, peroxydase	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	défense contre le stress oxydatif	ascorbate peroxidase	oxydoréductase, peroxydase	protéomique 2D	*Cunninghamia lanceolata*	embryogenèse zygotique précoce	(Shi et al, 2010)
	différenciation du protoderme	Protodermal factor2 (PDF2)	facteur de transcription	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique précoce	(Vestman et al., 2011)
	élongation cellulaire	TRANSMEMBRANE PROTEIN (TMP-C)	aquaglyceroporine	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique précoce	(Vestman et al., 2011)

Stade	fonction	gène/protéine	nature	technique de détection	espèce	expression	références
Embryogenèse précoce	élongation cellulaire	*PtNIP1;1* (NIP, nodulin-like intrinsic protein)	aquaglycéroporine	marquage histochimique GUS	*Pinus taeda*	stade précoce chez l'ES et l'EZ et en concordance avec le suspenseur	(Ciavatta et al., 2002)
	méthylation de l'ADN	adenosylhomocysteinase	adenosylhomocysteinase	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	méthylation de l'ADN	méthionine synthase	méthionine synthase	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	méthylation de l'ADN	S-adenosylmethionine synthase (SAM)	S-adenosylmethionine synthase (SAM)	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	méthylation de l'ADN	S-adenosylmethionine synthase (SAM)	S-adenosylmethionine synthase (SAM)	protéomique 2D	*Cunninghamia lanceolata*	embryogenèse zygotique précoce	(Shi et al, 2010)
	modification de la paroi cellulaire	SEPR43	myo-inositol oxygénase	ADNc-AFLPs (Amplified fragment length polymorphism) + RT-PCR	*Pinus radiata*	embryogenèse somatique précoce	(Aquea et Arce-Johnson, 2008)
	modification de la paroi cellulaire	SEPR91, SEPR110 et SEPR114	pectinestérase	ADNc-AFLPs (Amplified fragment length polymorphism) + RT-PCR	*Pinus radiata*	embryogenèse somatique précoce	(Aquea et Arce-Johnson, 2008)
	modification de la paroi cellulaire	SEPR1	α-D-galactosidase	ADNc-AFLPs (Amplified fragment length polymorphism) + RT-PCR	*Pinus radiata*	embryogenèse somatique précoce	(Aquea et Arce-Johnson, 2008)
	modifications de la paroi cellulaire	UDP-galactose 4-epimerase	Epimérase	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	modifications de la paroi cellulaire	UDP-glucose dehydrogenase	Oxidoreductase	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	modifications de la paroi cellulaire	GLP	germine	protéomique 2D	*Picea glauca*	abondante dès le 7ème jour de développement de l'embryon	(Lippert et al., 2005)
	modifications de la paroi cellulaire	LmGER1	germine	immunoserum anti-GLP + analyse d'activité sur gel	*Larix kaempferi* × *Larix decidua*	embryogenèse somatique précoce	(Mathieu et al., 2006)
	modifications de la paroi cellulaire	pcGER1	germine	crible de banque d'ADNc	*Pinus caribaea*	stade pré-globulaire	(Neutelings et al., 1998)

Stade	fonction	gène/protéine	nature	technique de détection	espèce	expression	références
Embryogenèse précoce	modifications de la paroi cellulaire	*Pa18*	lipid transfer protein (LTP)	crible de banque d'ADNc	*Picea abies*	embryogenèse somatique précoce	(Sabala et al., 2000; Hjortswang et al., 2002)
	mort cellulaire programmée	caspase-6-like (mcII-Pa)	protéase	crible de banque ADNc + hybridation in situ	*Picea abies*	activité forte dans les ES à un stade précoce et faible lorsque la morphogenèse de l'embryon est complète	(Suarez et al., 2004)
	mort cellulaire programmée	Cathepsin B-like cysteine protease + Metacaspase 9 (MC9)	protéase	microarray + qRT-PCR	*Picea abies*	différentiellement exprimé au cours de la différentiation précoce de l'embryon	(Vestman et al., 2011)
	régulateur majeur de l'embryogenèse	Pavp 1	facteur de transcription viviparous 1	northern blot hybridation	*Picea abies*	stade cotylédonaire précoce	(Fischerova et al., 2008)
	régulateur majeur de l'embryogenèse	LEC1-type	Leafy cotyledon 1-like (L1L), facteur de transcription	RT-PCR	*Picea abies*	embryogenèse somatique précoce	(Uddenberg et al., 2011)
	régulateur majeur de l'embryogenèse	*L1L*	Leafy cotyledon 1-like	qRT-PCR	*Pinus pinaster*	embryogenèse somatique précoce	(de Vega-Bartol et al., 2013b)
	régulateur négatif des gibbérellines (GA)	Roxy1 (ROXY1)	glutarédoxine	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique précoce	(Vestman et al., 2011)
	réorganisation du cytosquelette	*Pa1*, *Pa2*, *Pa3* and *Pa4*	Actin isoformes	semi-quantitative RT-PCR	*Picea abies*	embryogenèse somatique précoce	(Schwarzerova et al., 2010)
	réorganisation du cytosquelette	F-actin capping protein	actine	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique précoce	(Vestman et al., 2011)
	réorganisation du cytosquelette	tubuline	tubuline	microarrays	*Picea abies*	embryogenèse somatique précoce	(Stasolla et al., 2004)
	signalisation cellulaire	CHIA4-Pa chitinases	chitinase	immunolocalisation + western blot + hybridation in situ	*Picea abies*	embryogenèse somatique précoce	(Wiweger et al., 2003)
	signalisation cellulaire	chitinase	chitinase	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique précoce	(Vestman et al., 2011)
	signalisation cellulaire	reversibly glycosylated polypeptides (RGPs)	UDP-arabinopyranose mutase	semi-quantitative RT-PCR	*Araucaria angustifolia*	embryogenèse somatique précoce	(Schlögl et al., 2012)

Stade	fonction	gène/protéine	nature	technique de détection	espèce	expression	références
Embryogenèse précoce	transduction du signal	PmSERK1 (SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1)	leucine-rich repeat receptor-like kinase	crible de banque d'ADNc	*Pinus massoniana*	embryogenèse somatique précoce	(Yan et al., 2010)
	transduction du signal	SERK (SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE)	leucine-rich repeat receptor-like kinase	microarray + qRT-PCR	*Picea abies*	exprimé pendant la différenciation précoce des embryons	(Vestman et al., 2011)
	transduction du signal	Rab1	small GTPase superfamily	qRT-PCR	*Pinus pinaster*	ES précoce puis ES en germination	(de Vega-Bartol et al., 2013b)
Embryogenèse tardive	voie de signalisation régulée par l'auxine	Indole-3-acetic acid inducible 11 (IAA11)	facteur de transcription	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique tardive	(Vestman et al., 2011)
	voie de signalisation régulée par l'auxine	MYB77	facteur de transcription	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique tardive	(Vestman et al., 2011)
	voie de signalisation régulée par l'auxine	auxin receptor, TIR1	récepteur de l'auxine	microarray + qRT-PCR	*Picea abies*	embryogenèse somatique tardive	(Vestman et al., 2011)
	biosynthèse des gibbérellines	GA2OX2 et GA2OX6 (baisse de l'expression)	Gibberelline 2-oxidase	microarray + qRT-PCR	*Picea abies*	régulé négativement durant ES précoce et ES tardive de l'embryon	(Vestman et al., 2011)
	contrôle le développement des méristèmes	Clavata-like	leucine-rich-repeat protein receptor kinase (LRR RLK)	RT-PCR	*Pinus taeda* et *Pinus oocarpa*	maximum au stade globulaire puis constant jusqu'au satde cotylédonaire tardif	(Lara-Chavez et al., 2012)
	contrôle du cycle cellulaire	p34cdc2 protein kinase (cdc2Pa)	Serine/threonine-protein kinase	southern blot	*Picea abies*	embryogenèse somatique tardive	(Footitt et al., 2003)
	contrôle de la différenciation des méristèmes (apical et racinaire)	PgAGO, ARGONAUTE	fait partie du complexe RNA-induced silencing complex (RISC)	crible de banque d'ADNc	*Picea glauca*	différentiellement exprimé au cours de la différentiation tardive de l'embryon	(Tahir et al., 2006)
	contrôle la séparation des organes	LBD12/ASL5 et LBD40 (LOB domain-containing protein)	LOB domain-containing protein	microarray + qRT-PCR	*Picea abies*	différenciellement exprimé pendant la différenciation tardive des embryons	(Vestman et al., 2011)

Stade	fonction	gène/protéine	nature	technique de détection	espèce	expression	références
Embryogenèse tardive	contrôle du développement du protoderme	*PaHB1, Homeobox1*	facteur de transcription à homéodomaine	clone ADNc + RT-PCR +hybridation in situ	*Picea abies*	recuis pour la transition vers l'embryogenèse tardive	(Ingouff et al., 2001)
	régulateur majeur de l'embryogenèse	ABA insensitive 3 (ABI3)	facteur de transcription	microarray + qRT-PCR	*Picea abies*	augmente lors de la transition vers l'embryogenèse tardive	(Vestman et al., 2011)
	régulateur majeur de l'embryogenèse	Leafy cotyledon 1-like (L1L)	Leafy cotyledon 1-like	microarray + qRT-PCR	*Picea abies*	diminue lors de la transition vers l'embryogenèse tardive	(Vestman et al., 2011)
	régulateur majeur de l'embryogenèse	Pavp1	facteur de transcription viviparous 1	southern blot	*Picea abies*	ES mature	(Footitt et al., 2003)
	régulation de la structuration des tissus	HD-Zip I	Homeodomain-leucine zipper I, facteur de transcription	RT-PCR	*Pinus taeda* et *Pinus oocarpa*	maximum au stade globulaire puis diminue progressivement	(Lara-Chavez et al., 2012)
	régulation négative des voies de biosynthèse des gibbérellines	Spindly (SPY)	facteur de transcription	microarray + qRT-PCR	*Picea abies*	marque la transition vers l'embryogenèse somatique tardive	(Vestman et al., 2011)
	synthèse protéique	proteasome subunit alpha type-3	protéasome	protéomique 2D	*Pinus massoniana*	embryogenèse zygotique précoce	(Zhen et al, 2012)
	synthèse protéique	RPN1	26S proteasome regulatory subunit S2	RT-PCR	*Pinus taeda* et *Pinus oocarpa*	maximum au stade globulaire puis constant jusqu'au stade cotylédonaire tardif	(Lara-Chavez et al., 2012)
Maturation	régulation de la germination des graines	Protein MOTHER of FT and TF 1	phosphatidylethanolamine-binding protein family	qRT-PCR	*Pinus pinaster*	ES mature	(de Vega-Bartol et al., 2013b)
	métabolisme de l'azote	GS1a	Glutamine synthetase cytosolic isozyme	northern blot assay + hybridation in situ	*Pinus pinaster* et *Pinus sylvestris*	ES en germination, absent de l'EZ	(Perez Rodriguez et al., 2006)
	protection contre la dessiccation	LEA	Late Embryogenesis Abundant	RT-PCR	*Pinus taeda* et *Pinus oocarpa*	présent dès le stade MEs, augmente jusqu'au stade cotylédonaire tardif	(Lara-Chavez et al., 2012)
	protection contre la dessiccation	*PgEMB12, PgEMB14, PgEMB15*	late-embryogenesis-abundant (LEA) proteins ; (ABA)-responsive gene	crible de banque d'ADNc	*Picea glauca*	ES cotylédonaire	(Dong et Dunstan, 1997a)

Stade	fonction	gène/protéine	nature	technique de détection	espèce	expression	références
Maturation	protection contre la dessiccation	legumin-like storage protein (pseudotsugin)	LEA protein similaire à 11–12S storage globulins.	clones ADNc identifiés par immunomarquage avec du serum polyclonal	*Pseudotsuga menziesii*	abondant au début et au milieu des stades de l'embryogenèse dans l'axe embryonnaire ainsi que dans le megagametophyte	(Leal et Misra, 1993)
	protection contre la dessiccation	late embryogenesis abundant proteins (LEA)	protéine de réserve	SDS-PAGE + spectrométrie de masse (GC/TOFMS)	*Picea abies*	ES cotylédonaire mature (traitement saccharose)	(Businge et al., 2013)
	protection contre la dessiccation	LEA protein	protéine de réserve	microarrays	*Picea abies*	embryogenèse somatique tardive	(Stasolla et al., 2004)
	protection contre la dessiccation	three ABA-responsive LEA-like genes	protéine de réserve	high-density Southern array hybridization	*Pinus taeda*	ES mature	(Vales et al., 2007)
	protection contre la dessiccation	déhydrine	protéine LEA de classe II	SDS-PAGE + spectrométrie de masse (GC/TOFMS)	*Picea abies*	ES cotylédonaire mature (traitement saccharose)	(Businge et al., 2013)
	protection contre la dessiccation	HSPs	protéine de réponse au stress	microarrays	*Picea abies*	embryogenèse somatique tardive	(Stasolla et al., 2004)
	protection contre la dessiccation	heat shock proteins (HSPs) and small heat shock proteins (smHSPs)	proteine de réponse au stress	SDS-PAGE + spectrométrie de masse (GC/TOFMS)	*Picea abies*	ES cotylédonaire mature (traitement saccharose)	(Businge et al., 2013)
	protection contre la dessiccation	smHSP	class II LMW HSPs	crible de banque d'ADNc	*Picea glauca*	détectés dans les tissus embryogènes avant l'induction de la maturation, puis beaucoup plus abondant au stade tardif de l'embryogenèse	(Dong et Dunstan, 1996)
	réserve en acides aminés et en azote	legumin (57 kDa)	11S globuline	SDS-PAGE + spectrométrie de masse (GC/TOFMS)	*Picea abies*	ES cotylédonaire mature (traitement PEG+maltose)	(Businge et al., 2013)
	réserve en acides aminés et en azote	Legumin-like	11S globuline	RT-PCR	*Pinus taeda* et *Pinus oocarpa*	apparait au stade globulaire et maximum au stade cotylédonaire tardif	(Lara-Chavez et al., 2012)
	réserve en acides aminés et en azote	Legumin-like	11S globuline	RT-PCR	*Pinus taeda* et *Pinus oocarpa*	apparait au stade globulaire et maximum au stade cotylédonaire tardif	(Lara-Chavez et al., 2012)

Stade	fonction	gène/protéine	nature	technique de détection	espèce	expression	références
Maturation	réserve en acides aminés et en azote	légumine -like	11S- légumine	protéomique 2D	*Larix* × *eurolepis*	ES mature	(Teyssier et al., 2014)
	réserve en acides aminés et en azote	2S storage protein (20 kDa)	2S albumine	SDS-PAGE + spectrométrie de masse (GC/TOFMS)	*Picea abies*	ES cotylédonaire mature (traitement PEG+maltose)	(Businge et al., 2013)
	réserve en acides aminés et en azote	Vicilin	7S globuline	semi-quantitative RT-PCR	*Araucaria angustifolia*	ES cotylédonaire	(Schlögl et al., 2012)
	réserve en acides aminés et en azote	vicilin (50 and 51 kDa)	7S globuline	SDS-PAGE + spectrométrie de masse (GC/TOFMS)	*Picea abies*	ES cotylédonaire mature (traitement PEG+maltose)	(Businge et al., 2013)
	réserve en acides aminés et en azote	viciline -like	7S globuline	protéomique 2D	*Larix* × *eurolepis*	ES mature	(Teyssier et al., 2014)
	réserve en acides aminés et en azote	Vicilin-like	7S globuline	RT-PCR	*Pinus taeda* et *Pinus oocarpa*	apparait au stade globulaire et maximum au stade cotylédonaire tardif	(Lara-Chavez et al., 2012)
	réserve en acides aminés et en azote	seed storage protein (19 kDa)	protéine de réserve	SDS-PAGE + spectrométrie de masse (GC/TOFMS)	*Picea abies*	ES cotylédonaire mature (traitement PEG+maltose)	(Businge et al., 2013)

III.1. Embryogenèse précoce

Ces gènes interviennent au cours de l'induction de l'embryogenèse, de la proembryogenèse et de l'embryogenèse précoce. Dans l'idéal, la caractérisation de ce type de marqueurs pourrait aider à la sélection de lignées à fort potentiel embryogène.

III.1.1. Expression des gènes dans le suspenseur

Chez *Pinus taeda* de nombreux ARNm présentent une forte similarité de séquence avec des gènes d'*Arabidopsis* impliqués dans le développement du suspenseur et dans l'embryogenèse précoce (Cairney et Pullman, 2007). L'expansion cellulaire se réalise par l'augmentation de l'absorption d'eau au travers des aquaporines, qui agissent comme des canaux facilitant le transport rapide de l'eau. Les aquaporines sont surexprimées dans les régions d'élongation cellulaire (Hohmann et al., 2000). Un gène d'aquaporine spécifique du suspenseur, *PtNIP1;1* (NIP, nodulin-like intrinsic protein), a été identifié chez *Pinus taeda*. L'ARNm de *PtNIP1;1* code une aquaglycéroporine capable d'agir comme un canal hydrique (Ciavatta et al., 2001). *PtNIP1;1* est exprimé à un stade précoce chez l'ES et l'EZ et en concordance temporelle avec le développement du suspenseur. Chez *Picea abies NIP1;1* s'exprime constitutivement dans les MEs en culture, puis, lors de l'embryogenèse précoce, spécifiquement dans le suspenseur ; il n'est plus du tout exprimé lors de la maturation de l'embryon (Ciavatta et al., 2002).

Le suspenseur, persistant jusqu'au stade cotylédonaire, est ensuite dégradé par mort cellulaire programmée (Programed Cell Death, PCD).

III.1.2. Evénements précoces de mort cellulaire programmée

La morphogenèse de l'embryon est permise par l'élimination des cellules devenues obsolètes, par la mise en place d'épisodes de PCD massifs (Bozhkov et al., 2005). Deux épisodes successifs de PCD ont été mis en évidence au cours de la morphogenèse des ES chez *Picea abies*, le premier pour éliminer les cellules embryonaires et «sculpter» l'embryon, et le second pour éliminer le suspenseur (Filonova et al., 2000). La protéine mcII-Pa, dont la séquence contient de nombreuses caractéristiques de la superfamille des caspases/métacaspases/paracaspases, a été isolée dans les embryons de *Picea abies* (Suarez et al., 2004). Son activité de type caspase-6-like est forte dans les embryons à un stade précoce qui présentent une quantité élevée de PCD, et faible lorsque la morphogenèse de l'embryon est complète. De plus, son activité a été localisée dans les régions embryonnaires subissant la

PCD (Bozhkov et al., 2005). Des ES transgéniques, où le gène mcII-Pa a été inhibé par ARN interférent (ARNi), montrent une activité caspase-6-like et des PCD considérablement réduites. Ces résultats mettent en évidence le rôle central de mcII-Pa dans la PCD et le développement de l'embryon.

III.1.3. Gènes régulant le cycle cellulaire et la paroi cellulaire

L'embryogenèse somatique nécessite un contrôle spatio-temporel strict au cours de la division et de l'élongation des cellules. La polarité de l'embryon est établie par un programme précis qui contrôle la division des cellules embryogènes et l'élongation des cellules du suspenseur (Feher et al., 2003). Ainsi, l'initiation de la synthèse d'ADN et la division cellulaire sont des événements clés des étapes précoces au cours de l'embryogenèse. La caractérisation moléculaire de l'embryogenèse somatique peut ainsi être basée sur les gènes dont l'expression dépend du cycle cellulaire (Dudits et al., 1991). L'embryogenèse somatique dépend aussi des modifications de la paroi cellulaire et de la régulation des molécules signal liées à la paroi cellulaire. Ces processus peuvent entraîner des modifications essentielles à l'expression d'un ensemble de gènes responsables du cycle cellulaire et de la formation de la paroi et provoquer par la suite des changements coordonnés dans les fonctions cellulaires à différents stades de l'embryogenèse somatique.

La régulation de SEPR1 et SEPR43, respectivement homologue à la α-D-galactosidase et à la myo-inositol oxygénase, augmente tandis que la régulation de SEPR91, SEPR110 et SEPR114, qui codent des protéines de la famille des pectinestérases, diminue chez *Pinus radiata* (Aquea et Arce-Johnson, 2008). Ces gènes SEPRs sont impliqués dans la modification de la paroi cellulaire. L'identification de ces gènes soutient l'idée que des altérations spécifiques de la composition de la paroi des cellules embryogènes peuvent être importantes pour le développement de l'embryon proprement dit (Stasolla et al., 2004).

Le cytosquelette est aussi fortement remanié lors des divisions cellulaires et plusieurs expansines, actines et tubulines ont été identifiées au cours de l'embryogenèse précoce (Bishop-Hurley et al., 2003; Stasolla et al., 2004; Vestman et al., 2011).

Des enzymes de méthylation de l'ADN ont aussi été détectées, comme l'adenosyl-méthionine synthase (SAM) ; leur rôle est le remodelage de la chromatine pour mettre en place le programme d'expression de l'embryogenèse (Stasolla et al., 2004).

III.1.4. Les gènes de régulation du développement embryonnaire

Les gènes codant des facteurs de transcription et de transduction du signal sont des gènes clés pour la régulation de l'expression des gènes et pour la mise en place et la coordination de programmes cellulaires comme pour le développement embryonnaire. Il est désormais possible d'observer directement la dynamique d'expression de ces gènes par l'utilisation de colorants histochimiques et du marquage moléculaire (Yang et Zhang, 2010). Ces régulateurs peuvent intervenir à des étapes spécifiques de l'embryogenèse (Jenik et Barton, 2005) ou au cours de l'ensemble du processus (Park et Harada, 2008). Ces gènes sont considérés comme des régulateurs majeurs de l'embryogenèse des plantes supérieures, plusieurs d'entre eux ont été identifiés au cours de l'embryogenèse des conifères.

III.1.4.1. Gènes Leafy Cotylédon (LEC)

Les gènes de type LEC sont les régulateurs majeurs de l'embryogenèse les plus connus. La première classe de gènes de type LEC, représentés par LEC1 et L1L chez *Arabidospsis*, codent des facteurs de transcription de type HAP3 (Kwong et al., 2003) ; la deuxième classe de gènes, qui comprend LEC2, FUS3 et ABI3, ainsi que le facteur Viviparous1 (Vp1), code un domaine de liaison à l'ADN de type B3 (Stone et al., 2001). Ces deux classes de gènes codent des protéines de régulation impliquées dans l'embryogenèse (Kwong et al., 2003) et sont essentielles pour l'induction de développement de l'ES (Gaj et al., 2005). L'expression de LEC1 permet le maintien de l'embryogenèse et réprime le développement végétatif (Gaj et al., 2005). L'action de LEC1 est favorisée par l'auxine et le saccharose. La répression de la voie de régulation LEC1 dans des tissus post-embryonnaires, qui peut en partie être contrôlée par méthylation de l'ADN, supprime l'accumulation des produits de stockage (lipides et amidon), mais aussi la division et la différenciation cellulaire (Casson et Lindsey, 2006). La régulation négative de L1L et la régulation positive de ABI3 marque le début de la phase de maturation de l'ES chez *Picea abies* (Vestman et al., 2011). Chez *Pinus pinaster*, L1L est spécifiquement exprimé dans la phase précoce de l'embryogenèse somatique (de Vega-Bartol et al., 2013a), tandis que PaVP1, un orthologue de ABI3, est fortement surexprimé au cours de la maturation des ES de *Picea abies* (Footitt et al., 2003; Fischerova et al., 2008). LEC2 active des gènes impliqués dans la biosynthèse de l'auxine et cause des modifications d'expression de gènes répondant à l'auxine (Stone et al., 2008).

III.1.4.2. Gènes de type WUSCHEL et WUSCHEL-related homeobox (WOX)

Les protéines de type WUSCHEL font partie des facteurs de transcription présentant un motif de fixation à l'ADN très conservé appelé homéodomaine. Elles contrôlent le développement des méristèmes racinaires et caulinaires de la plante (Mayer et al., 1998), mais aussi le développement des méristèmes embryonnaires dans la graine (Gallois et al., 2004). WUSCHEL joue un rôle dans le développement des méristèmes végétaux en contrôlant l'identité cellulaire ; il permet ainsi de promouvoir la prolifération cellulaire et de réprimer la différenciation cellulaire (Gallois et al., 2004). Il est aussi nécessaire pour maintenir la division cellulaire dans les embryons et les suspenseurs. Ainsi, il est essentiel pour la croissance cellulaire durant l'embryogenèse (Wu et al., 2007).

L'étude du transcriptome des embryons de *Pinus taeda* a révélé que le gène STIMPY/WOX9 agissait dans la zone centrale du méristème caulinaire de l'embryon en régulant positivement l'activité de WUS pour maintenir l'état cellulaire indifférencié (Cairney et al., 2006). Chez *Picea. abies*, *PaWOX2* est fortement exprimé au cours du développement précoce de l'ES, mais diminue avec la maturation de l'ES. Aucune expression n'a été détectée dans les cellules non-embryogènes, indiquant qu'une forte expression de *PaWOX2* est liée à la vitesse de prolifération des cultures cellulaires embryogènes et peut être utilisée comme marqueur potentiel de l'embryogenèse. La capacité de WUS à promouvoir la transition de la cellule végétative vers la cellule embryonnaire, et finalement à former des ES, suggère que la protéine à homéodomaine jouerait également un rôle essentiel au cours de l'embryogenèse, en plus de sa fonction dans le développement du méristème (Palovaara et Hakman, 2008).

III.1.4.3. Récepteur kinase de l'embryogenèse somatique (SERK)

Le développement de l'embryon implique des événements moléculaires englobant non seulement l'expression différentielle des gènes, mais aussi différentes voies de transduction du signal d'activation ou d'inhibition de nombreux ensembles de gènes (Zimmerman, 1993). Les molécules de signalisation sont importantes pour l'induction et pour le maintien du développement embryonnaire (Quiroz-Figueroa et al., 2002; Raghavan, 2006).

De nombreuses kinases sont spécifiques des stades précoces de l'ES. Elles sont activées par phosphorylation et permettent la transduction de signaux qui activent les facteurs responsables des modifications cellulaires observées lors de la formation des ES (Jimenez et Bangerth, 2001; Shah et al., 2001).

Une de ces protéines kinases, isolée pour la 1ère fois dans les cellules de carotte (*Daucus carota*) se développant en ES, a été appelée récepteur kinase de l'embryogenèse

somatique (SERK) (Schmidt et al., 1997). Depuis, les gènes SERK ont été liés à l'embryogenèse somatique chez de nombreuses espèces. L'expression de SERK1 est liée à l'évolution du développement et de la reprogrammation cellulaire associée au développement de la plante (Nolan et al., 2009). Le gène *PmSERK1*, isolé chez *Pinus massoniana*, peut aussi être un gène marqueur de la capacité embryonnaire des cals. Il montre une séquence homologue aux gènes SERK impliqués dans l'embryogenèse somatique chez d'autres espèces (Yan et al., 2010).

III.1.5. Protéines extracellulaires

III.1.5.1. Arabinogalactane protéines, chitinases et lipochitooligosaccharides

La plupart des protéines excrétées dans les milieux de culture sont des glycoprotéines (protéines glycosylées) (Satoh et al., 1986). Parmi ces glycoprotéines, les arabinogalactanes protéines (AGPs), constituent un groupe hétérogène largement distribué dans le règne végétal. Les AGPs sont impliquées dans la signalisation cellulaire. Elles sont excrétées dans le milieu de culture et sont capables d'induire l'embryogenèse somatique lorsqu'elles sont ajoutées à une lignée cellulaire à faible pouvoir embryogène (Kreuger et Holst, 1993; von Arnold et al., 2005). La digestion des AGPs par les endochitinases produit des lipochitooligosaccharides (LCOs) qui ressemblent aux facteurs de nodulation (NOD) de *Rhizobium*, capables de stimuler l'embryogenèse somatique chez la carotte (de Jong et al., 1993) et chez *Picea abies* (Egertsdotter et von Arnold, 1998). Les relations qui existent entre les AGPs, les chitinases et les LCOs, leur implication dans la signalisation cellulaire et leur capacité commune à stimuler l'embryogenèse ont été mises en évidence chez la carotte (van Hengel et al., 2002). Un certain nombre d'AGP, de LCOs et d'endochitinases a été identifié chez les conifères à des stades précoces de l'embryogenèse somatique (**Table I.3**).

III.1.5.2. Protéines transporteur de lipides non-specifiques (LTPs)

Les LTPs représentent une famille de protéines ubiquitaires chez les plantes (Kader, 1996). Ces protéines sont caractérisées par leur capacité à transférer des phospholipides entre les membranes et pour lier les acides gras *in vitro*. Les LTPs possèdent plusieurs fonction *in vivo*, comme le transport de composés dans la cuticule (Sterk et al., 1991) et l'inhibition de la croissance d'agents pathogènes bactériens et fongiques (Molina et al., 1993). Les LTPs sont des marqueurs de l'embryogenèse somatique précoce chez la carotte (Sterk et al., 1991). Ils sont impliqués dans la formation de la couche du protoderme, qui exerce un rôle de régulation

dans le contrôle de l'expansion des cellules pendant le développement des ES et des EZ d'*Arabidospsis* (Thoma et al., 1994; Dodeman et al., 1997).

III.1.5.3. Germines et Germin-Like Proteins (GLP)

Les protéines de type germine et germin-like protein (GLPs) appartiennent à une superfamille de protéines largement répandues dans le règne végétal. Elles possèdent diverses fonctions et sont structurellement proches des protéines de la superfamille des cupines, qui possède un domaine conservé en feuillets-β (Dunwell et al., 2008). Leur nom leur a été donné après leur première identification en tant que marqueurs spécifiques de la germination du blé (Thompson et Lane, 1980; Grzelczak et Lane, 1984). Les GPLs ont ensuite été caractérisées comme des glycoprotéines ayant une activité oxalate oxydase (Lane, 2000) ; elles sont incorporées dans la matrice extracellulaire par des liaisons ioniques (Jaikaran et al., 1990). Par conséquent, elles peuvent présenter plusieurs fonctions, en tant qu'enzymes, protéines structurelles, ou récepteurs au cours de l'embryogenèse somatique, en réponse à un stress salin ou à des agents pathogènes (Dunwell et al., 2000; Bernier et Berna, 2001; Lane, 2002).

La comparaison des profils des protéines extracellulaires de lignées cellulaires non-embryogènes et embryogènes chez *Pinus caribaea* a conduit à l'identification de la première GLP impliquée dans l'embryogenèse somatique (Domon et al., 1995). Son ADNc (PcGER1) fut isolé plus tard et l'analyse de son expression a confirmé sa spécificité embryonnaire (Neutelings et al., 1998). L'expression de PcGER1 est reliée au cycle cellulaire chez *Pinus caribaea* (Mathieu et al., 2003), comme montré précédemment dans le cas de la germine du blé (Lane, 2000). Plus tard, l'ADNc d'une GLP homologue à GER1 a été isolée, et a montré un fort niveau d'expression dans les tissus embryogènes mais pas dans les tissus non-embryogènes chez *Pinus radiata* (Bishop-Hurley et al., 2003). Chez le mélèze hybride la GLP (LmGER1) joue un rôle central dans la formation de l'ES *via* la régulation du remodelage de la paroi cellulaire (Mathieu et al., 2006). Une analyse protéomique a été réalisée pour évaluer quantitativement les niveaux d'expression de protéines à travers quatre stades de maturation des ES chez *Picea glauca*. La protéine GLP montre une variation importante d'abondance dès le 7ème jour de développement de l'embryon (Lippert et al., 2005).

III.2. Embryogenèse tardive

Pendant la phase de maturation les ES subissent divers changements morphologiques et biochimiques. Les organes de stockage, les cotylédons, se développent de façon concomitante avec l'accumulation de composés, tandis que la germination est réprimée. Ces réserves vont permettrent par la suite l'acquisition de la tolérance à la déshydratation (Goldberg et al., 1989). Les ES accumulent les mêmes types de composés de réserve que les EZ, essentiellement des lipides, des sucres et des protéines (Merkle et al., 1995; Lara-Chavez et al., 2012). Dans l'idéal, les ES produits devraient pouvoir présenter la même qualité en terme de réserves que l'EZ pour pouvoir germer avec la même réussite et régénérer des plants aussi vigoureux. Cependant il s'avère que la quantité d'un composé de réserve ainsi que le moment de son accumulation, varient souvent entre l'ES et l'EZ (Klimaszewska et al., 2004; Tereso et al., 2007).

III.2.1. Protéines de réserve

La synthèse et l'accumulation de protéines de réserve sont généralement régulées par l'ABA et par la diminution de la teneur en eau lors de la déshydratation tardive de l'embryon (Dodeman et al., 1997). Chez *Picea abies*, l'accumulation des protéines de réserve n'est pas observée au cours du développement des ES de certaines lignées, cette caractéristique a donc été utilisée pour discriminer les ES « anormaux » des ES bien développés (Stasolla et Yeung, 2003). Ainsi, un ES de haute qualité accumule des protéines de réserve analogues à celles des EZ (Merkle et al., 1995), constituant d'excellents marqueurs pour vérifier la qualité des ES (Mondal et al., 2002; Klimaszewska et al., 2004; Tereso et al., 2007).

Les protéines de réserve possèdent des caractéristiques communes telles que leur solubilité, leur masse moléculaire, la structure de leurs sous-unités ainsi que leurs séquences génomiques (Gifford, 1988; Jensen et Berthold, 1989; Misra et Green, 1990b; Roberts et al., 1990; Green et al., 1991; Jensen et Lixue, 1991; Allona et al., 1992; Newton et al., 1992; Hakman, 1993a; Leal et Misra, 1993). Traditionnellement, les protéines de réserve sont classées selon leur solubilité dans l'eau (albumines), dans une solution salée (globulines), dans une solution alcaline (glutélines) ou une solution alcoolique (prolamines) (Higgins, 1984). Les protéines de réserve sont initialement synthétisées sous forme de précurseurs polypeptidiques à partir de grandes familles multigéniques, et peuvent subir des modifications post-traductionnelles, telles que le clivage du peptide signal, la glycosylation, l'association

d'oligomères, et des combinaisons intramoléculaires et intermoléculaires de ponts disulfures (Shotwell et Larkins, 1989). La plupart des protéines de réserve ne devrait pas être considérée comme une seule protéine, mais plutôt comme un complexe de protéines individuelles reliées par une combinaison de ponts disulfures intermoléculaires, de liaisons hydrogènes, de liaisons ioniques et de liaisons hydrophobes. Après modification, les protéines sont accumulées soit dans le cytoplasme soit sous forme de corps protéiques.

La majorité des protéines de réserve, contenues dans les embryons de conifères (ES et EZ), sont de type globuline. Ces globulines sont divisées en familles majeures selon leur masse moléculaire et leur coefficient de sédimentation (valeurs S). La 7S globuline (viciline), souvent abondante dans les légumineuses et la 11S globuline (légumine) sont les plus représentées chez les conifères. La majorité des études portant sur les protéines de réserve des embryons de conifères sont anciennes et basées sur des techniques de détection par migration électrophorétique et comparaison des poids moléculaires relatifs. Cependant, quelques études plus récentes ont identifié des vicilines et des légumines par spéctrométrie de masse, comme par exemple chez *Pinus strobus* (Klimaszewska et al., 2004) *Picea abies* (Businge et al., 2013), ou encore le mélèze hybride (Teyssier et al., 2014).

III.2.1.1. 11S globuline ou legumine-like

Les 11S globulines sont divisées en deux grandes classes de protéines : celles qui sont complètement solubles (par exemple, les légumines de pois et de soja) et celles qui nécessitent l'addition de dodécylsulfate de sodium (SDS), d'urée ou d'autres substances dénaturantes pour être solubilisées (par exemple, les cristalloïdes) (Bewley et Black, 1985).

Les études sur les protéines des graines d'un certain nombre d'espèces de conifères ont montré que les cristalloïdes sont les protéines de réserve majeures, par exemple chez le genre *Pinus* (Gifford, 1988; Jensen et Lixue, 1991), *Picea glauca*, *Picea engelmannii*, *Picea mariana* (Misra et Green, 1990b; Roberts et al., 1990) *Picea abies* (Hakman, 1993b). Dans les graines matures de *Picea glauca* et de *Pseudotsuga menziesii*, les cristalloïdes représentent 70 à 80% du total des protéines de réserve et sont localisées dans les corps protéiques. Elles se composent d'un complexe de 55-63 kDa se dissociant en 20-23 kDa et 32-35 kDa en conditions réductrices (Misra et Green, 1990b; Green et al., 1991). Cette structure en sous-unité a été observée pour les protéines de réserve cristalloïdes de huit espèces de *Pinus* (Gifford, 1988) et les globulines isolées chez *Picea abies* (Hakman, 1993b). Généralement, les 11S globulines, isolées à partir de graines de conifères, présentent des similarités en terme

de solubilité et de configuration des liaisons disulfures avec leurs homologues d'angiospermes.

III.2.1.2. 7S globuline ou viciline-like

Les vicilines sont généralement des protéines à chaîne unique sans liaisons disulfures. Elles s'agrègent pour former des trimères de sous-unités ayant des masses moléculaires différentes (45-53 kD), mais qui ne présentent pas de profil électrophorétique différent en condition électrophorétique réductrice. Elles sont également codées par un grand nombre de gènes (Sales, 2000). Plusieurs protéines appartenant à cette famille multigénique ont été isolées dans une étude de protéomique sur l'embryon de *Medicago truncatula* à plusieurs stade de développemnt (Gallardo et al., 2003). Trois profils électrophorétiques pouvant correspondre à des protéines de type 7S viciline ont été décrits dans des graines de plusieurs conifères dont le pin maritime (Flinn et al., 1991; Allona et al., 1992; Newton et al., 1992).

III.2.1.3. 2S albumine

Les albumines sont des protéines de réserve de bas poids moléculaire présentant une structure dimérique formée par une petite sous-unité reliée à une grande sous-unité par des ponts disulfures. Elles ont été détectées chez un certain nombre de conifères au travers de leur profil électrophorétique (Gifford, 1988; Misra et Green, 1990a; Flinn et al., 1991; Green et al., 1991; Allona et al., 1992; Hakman, 1993a). Chez *Pinus pinaster*, les albumines constituent 15% de la teneur totale en protéines dans les graines matures (Allona et al., 1992).

III.2.2. Glucides et amidon

Les glucides jouent plusieurs rôle primordiaux au cours du développement embryonnaire, servant à la fois comme source d'énergie et de squelettes carbonés, comme agent osmotique et comme régulateurs de l'expression des gènes. Au cours de la maturation de l'embryon, ils jouent un rôle de molécules protectrices contre certains stress. Leur action est ainsi primordiale dans l'établissement de la tolérance des graines à la déshydratation et aux faibles températures (Bernal-Lugo et Leopold, 1995).

Le fructose et le glucose sont des glucides à 6 carbones (hexoses) et constituent les pièces de bases du métabolisme des glucides. Le saccharose est un oligosaccharide formé par 2 hexoses alors que le raffinose et le stachyose sont des oligosaccharides formés par 3 hexoses. De plus ces derniers possèdent un groupement cétone non-oxydable, et font partie

des oligosaccharides non-réducteurs. L'amidon est un polysaccharide constitué de n hexoses ; il constitue la principale forme de stockage des glucides pour les cellules végétales. Le métabolisme des glucides comprend l'ensemble des processus biochimiques responsables de la synthèse, du catabolisme et de l'interconversion de tous ces glucides.

III.2.2.1. Métabolisme des glucides lors de l'embryogenèse

Le développement des ES et des EZ s'accompagne de changements remarquables des profils d'accumulation des glucides et de l'amidon, allant de pair avec des modifications d'activité des enzymes impliquées dans le métabolisme carboné, comme l'invertase et la saccharose synthase. Ainsi la teneur en glucides et en amidon et l'activité des enzymes clés peuvent constituer des marqueurs de la qualité de l'embryon.

Chez *Pinus pinaster* une étude comparative entre ES et EZ sur l'accumulation des composés de réserve montre que l'amidon était prédominant dans les ES matures, alors que les EZ accumulaient préférentiellement des protéines et des lipides (Jordy et Favre, 2003). L'accumulation d'amidon a lieu principalement pendant les 2 premières semaines de maturation de l'ES chez *Picea mariana* et *Picea glauca* (Iraqi et Tremblay, 2001). Chez l'ES de *Picea abies*, une forte accumulation d'amidon accompagne la période d'apparition des cotylédons, suivie d'une baisse de sa teneur lors des étapes ultérieures entrainant *in fine* une faible quantité d'amidon dans les embryons matures (Lipavskà et al., 2000). De même, les EZ matures de *Picea abies* contiennent très peu d'amidon (Gosslova et al., 2001). Ces différences d'accumulation d'amidon entre espèces et entre types d'embryons peuvent être liées aux différents modèles expérimentaux, et/ou au génotype. Chez les ES de conifères d'autres composés de réserve, les lipides et les protéines, jouent un rôle important. Par conséquent, l'accumulation et la mobilisation d'une partie spécifique du spectre des composés de réserve pourraient être dépendante de l'espèce et de sa sensibilité aux conditions de maturation (Lipavskà et Konradova, 2004).

La dégradation du saccharose est la première étape pour l'utilisation du carbone par les cellules végétales. Elle est assurée par deux enzymes, l'invertase et la saccharose synthase, permettant l'utilisation ultérieure des hexoses produits. Dans les tissus végétaux, les invertases solubles jouent un rôle important dans la régulation de la composition en glucides (Kingston-Smith et al., 1999) et dans le partitionnement du saccharose (Tang et al., 1999) alors que l'activité saccharose synthase est corrélée à l'accumulation d'amidon et la synthèse des parois

cellulaires (Winter et Huber, 2000). D'autre part, la biosynthèse du saccharose a lieu dans le cytoplasme et implique l'activité de la saccharose phosphate synthase (Peace et al., 1995).

L'étude du métabolisme des glucides au cours de la maturation de l'EZ chez *Picea abies* a révélé que des changements dans le contenu des glucides solubles sont accompagnés par des changements des activités des principales enzymes du métabolisme du saccharose. L'activité saccharose synthase au cours de la maturation augmente dans l'embryon tandis qu'elle diminue dans le mégagamétophyte. L'activité de l'invertase est très faible voir non détectable au cours de la maturation des EZ. Une troisième enzyme, la phosphoglucomutase est impliquée dans synthèse de l'amidon et catalyse la réaction : glucose-1-phosphate → glucose-6-phosphate. Son activité est stable durant tout le processus de formation de la graine (Konradova et al., 2002).

Les changements dans l'activité des enzymes impliquées dans le métabolisme des glucides sont corrélés avec une forte diminution du ratio hexose:saccharose en fin de maturation de l'EZ cotylédonaire (Konradova et al., 2002; Pullman et Buchanan, 2008).

III.2.2.2. Oligosaccharides non-réducteurs

Les oligosaccharides de la famille du raffinose (Raffinose family of oligosaccharides, RFOs) sont des résidus alpha-galactosyl dérivés du saccharose, et les plus communs sont le raffinose (trisaccharidide), le stachyose (tétrasaccharide), et la verbascose (pentasaccharidique). Les RFOs sont presque omniprésents dans le règne végétal, ils sont présents dans une grande variété de graines appartenant à des familles très diverses, et ils se classent au deuxième rang derrière le saccharose en terme d'abondance sous forme de sucres solubles.

La plupart des graines orthodoxes accumulent des glucides non réducteurs (le saccharose, les RFOs et les cyclitols) dans les derniers stades de la maturation, ce qui implique leur rôle fonctionnel dans la protection contre les dommages liés à la déshydratation par remplacement de l'eau (Black et al., 1996; Bailly et al., 2001). Cette action protectrice du saccharose et des RFOs a été attribué à leur capacité à stabiliser les membranes et les protéines en interagissant avec les groupements hydrophiles (Crowe et al., 1992). Lorsque l'eau diminue, ces glucides (et d'autres « solutés compatibles » tels que la proline et la glycine bétaïne), s'accumulent et remplacent l'eau pour maintenir l'intégrité structurale des membranes et des protéines. Chez *Pinus mariana* il a été montré que l'augmentation de la teneur en saccharose seule ne permet pas l'acquisition de la tolérance à la dessiccation de l'embryon,

alors que la présence de saccharose et de RFOs le permet (Bomal et al., 2002). En outre, le raffinose peut augmenter la disponibilité du saccharose en empêchant la cristallisation de ce dernier, renforçant ainsi son action protectrice (Oliver et al., 1998).

Dans une autre étude chez *Picea abies*, une forte accumulation d'oligosaccharides de la famille du raffinose (RFO) a également été observée en fin de maturation de la graine, constituant un événement naturel de la fin de l'embryogenèse chez cette espèce (Gosslova et al., 2001). Le raffinose et le stachyose sont fortement accumulés dans les derniers stades de développement des EZ, tandis que pour l'ES mature, seul le stachyose a été détecté au même stade (Pullman et Buchanan, 2008). De plus le ratio RFO:saccharose augmente en fin de maturation dans l'EZ chez *Pinus taeda* (Pullman et Buchanan, 2008). L'apparition des RFOs est marqueur de la déshydratation des ES et des EZ chez les conifères (Lipavskà et Konradova, 2004; Businge et al., 2013).

Les RFOs pourraient aussi jouer un rôle de réserve en source de carbone pour le développement des embryons de *Picea abies* (Konradova et al., 2003). Des résultats chez *Picea glauca* ont permis de conclure que pour certains conifères, les RFOs servent au moins de source de carbone et d'énergie facilement remplaçables pendant les premières étapes de la germination des graines (Downie et Bewley, 2000).

III.2.3. Acide abscissique (ABA)

L'acide abscissique est important pour la maturation des graines chez de nombreuses espèces d'angiospermes (Zeevaart et Creelman, 1988; Kermode, 1990; Hetherington et Quatrano, 1991), mais aussi chez les gymnospermes comme le douglas (Chiwocha et Aderkas, 2002), l'épinette blanche et le mélèze hybride (Lelu et Label, 1994; Label et Lelu, 2000; von Aderkas et al., 2001; Kong et Aderkas, 2007). L'ABA empêche la germination précoce de l'embryon, et favorise ainsi son développement (Kermode, 1990). Son rôle dans l'embryogenèse est plus vaste puisque l'activation des gènes de type rab (response sensitive ABA) par l'ABA conduit à l'accumulation d'ARNm qui vont réguler, entre autres, l'accumulation de composés de réserve tels que des protéines, des lipides, de l'amidon, et enfin des protéines abondantes de l'embryogenèse tardive (Late Embryogenesis Abundant proteins, LEA). Par l'intermédiaire de ces composés, l'ABA est impliqué dans l'acquisition de la tolérance à la déshydratation (Skriver et Mundy, 1990).

L'ABA s'accumule à partir du moment où la graine a atteint la moitié de son développement. Puis sa teneur décline souvent dans la graine mature déshydratée, permettant par la suite à l'embryon mature imbibé de germer.

.

III.2.4. Protéines abondantes de l'embryogenèse tardive (LEA)

Les protéines abondantes de l'embryogenèse tardive (LEA) ont été découvertes pour la première fois dans les graines de coton où elles s'accumulaient dans l'embryon de façon tardive (Dure et al., 1981). Elles ont ensuite été retrouvées dans les graines de nombreuses plantes et en particulier en conditions de stress abiotique (déshydratation, stress salin et froid) (Ingram et Bartels, 1996; Thomashow, 1999). Les protéines LEA sont plus abondantes à la fin de l'embryogenèse (Galau et al., 1986). Leurs gènes sont exprimés au stade cotylédonaire de l'EZ, et sont considérés comme des gènes marqueurs de la qualité de l'ES mature avant germination (Zimmerman, 1993).

Les protéines LEA des plantes, constituent une famille de protéines très diversifiées. Elles ont été classées en différents groupes basés sur la similitude de motifs caractéristiques ou de composition globale en acides aminés (Bray, 1993; Dure, 1993; Wise, 2003). Cependant, les nomenclatures de classification n'ont pas toujours été homogènes dans la littérature (Tunnacliffe et Wise, 2007). Une caractéristique commune des protéines LEA est leur composition en acide aminé leur conférant un caractère hydrophile élevé et une forte stabilité à la chaleur (Wise et Tunnacliffe, 2004). Certaines LEA ont été identifiées comme des composants du systèmes ABA-inductible (Dure et al., 1989). Leur rôle physiologique est de permettre, lors de stress abiotiques, la stabilisation des membranes (Manfre et al., 2009) et des protéines labiles et en particuliers des enzyme clés du métabolisme (Hara et al., 2001; Sanchez-Ballesta et al., 2004). Les LEA agissent en association avec d'autres molécules protectrices, dont certains sucres comme le saccharose (Kermode, 1997). Leur profil d'expression au cours de l'embryogenèse et leur capacité ABA-inductible ont conduit à la suggestion que les protéines LEA joueraient un rôle dans la tolérance à la déshydratation des EZ. L'augmentation de l'expression de certains transcrits au moment de la perte d'eau, s'explique par la nécessité d'obtenir des produits de traduction tôt lors de la germination future. Leur abondance dans la graine mature, au sein du cytoplasme et du noyau, peut représenter jusqu'à 5 % des protéines totales et environ 30% des protéines de réserve (Manfre et al., 2009).

Les déhydrines, qui appartiennent au groupe II des protéines de type LEA, sont particulièrement impliquées dans la résistance à la déshydratation des graines (Shih et al., 2008). Une caractéristique importante des déhydrines est leur caractère hydrophile extrême dû à une teneur élevée en lysine et en glycine, et la quasi-absence d'acides aminés hydrophobes. Cela leur donne une grande solubilité et la capacité de lier l'eau, pour former une enveloppe stable en association avec les membranes et les macromolécules, dans des conditions de déficit hydrique (Hoekstra et al., 2001). Les déhydrines ne se dénaturent pas à la chaleur, et forment une structure désordonnée qui peut empêcher le collapse des protéines associées au cours de la déshydratation.

Les transcrits des gènes LEA des EZ de *Pinus taeda*, qui commencent à s'accumuler au stade globulaire, montrent ensuite peu de variation au stade cotylédonaire précoce et s'accumulent fortement au stade cotylédonaire tardif (Lara-Chavez et al., 2012). De façon similaire, au cours du développement des EZ de *Picea glauca*, l'accumulation des transcrits des protéines LEA, commence avant la phase cotylédonaire tardive et atteint son maximum lorsque l'embryon est complètement mature, pour rester stable au cours de la déshydratation de la graine (Leal et Misra, 1993). En outre, l'isolement et la caractérisation de *QrEm* (un gène LEA de *Quercus robus*) a montré que son expression était initiée au cours de l'embryogenèse précoce de l'EZ, puis augmentait progressivement tout au long de la maturation, pour baisser en fin du développement de la graine (Sunderlikova et al., 2009).

III.2.5. Les protéines de réponse au choc thermique (HSP)

Une seconde classe de protéines impliquée dans l'acquisition de la tolérance à la déshydratation des graines, est celle des petites protéines de choc thermique (smHSPs) de poids moléculaire compris entre 17 et 30 kDa. L'expression de leurs gènes au cours du développement des graines coïncide souvent avec celle des protéines LEA, les deux répondant positivement à l'ABA (Kermode et Finch-Savage, 2002). Les smHSPs sont des protéines de réponse au stress (et en particulier au stress thermique) ; elles permettent le maintien ou la restauration de la structure des protéines et empêchent leur dénaturation irréversible. Elles pourraient donc jouer un rôle similaire en ce qui concerne la tolérance à la déshydratation. Un certain nombre de smHSP a été caractérisé en fin de maturation des ES de conifères (Dong et Dunstan, 1996; Stasolla et al., 2004; Businge et al., 2013).

Cependant des HSP de haut poids moléculaire (HSP70, HSP90 et HSP101) ont aussi été détectées chez le mélèze hybride dans l'ES mature (Teyssier et al., 2014). Ces HSP joueraient d'autres rôles au cours de l'embryogenèse, de par leur implication dans le repliement, la

translocation, l'assemblage et la dégradation des protéines (Sung et al., 2001; Wang et al., 2004). De nombreuses études suggèrent que leur présence est requise tout au long de l'embryogenèse : chez le Douglas, une HSP70 est spécifique du développement précoce des graines (Forward et Misra, 2000) ; chez le dattier (*Phoenix dactylifera*) une étude de protéomique comparative entre EZ et ES a montré que EZ mature accumulait plus de protéines HSP (17.6 et 70 kDa) que l'ES en fin de maturation (Sghaier-Hammami et al., 2009).

Partie IV - Méthodologies de la physiologie moléculaire

Si le processus d'embryogenèse somatique a été très bien étudié d'un point de vue technique et appliqué, il manque à ce jour des descripteurs/marqueurs moléculaires du développement (**Table I.3**, peu de données sur le pin maritime) permettant d'aborder les mécanismes moléculaires mis en œuvre au cours de l'ontogenèse de l'embryon. En effet, le développement de l'embryon demeure complexe de part les interactions qu'il entretient avec son environnement de croissance qu'il soit zygotique (la composition et l'évolution du mégagamétophyte) ou somatique (les nombreux paramètres de l'environnement de culture). Des équilibres bidirectionnels s'établissent entre l'EZ et son tissu nutritif, ou entre les cultures embryogènes et leur environnement. Ces paramètres peuvent ainsi être grandement modifiés en jouant sur la composition du milieu *in vitro* et entrainer des changements importants au cours de la maturation. Cependant, ces changements peuvent intervenir de façon rapide en début de maturation et ne sont pas forcément observables au niveau macroscopique. De la même façon, l'absence d'évolution visible des ES cotylédonaires durant plusieurs semaines en fin de maturation, interroge sur les phénomènes moléculaires agissant dans l'embryon. Nous devons développer notre compréhension des mécanismes physiologiques qui contrôllent le développement des ES et rechercher des marqueurs de leur maturité (Lelu-Walter et al., 2008). Une approche consiste à suivre et comparer l'évolution de l'accumulation des transcrits (Thibaud-Nissen et al., 2003) et des protéines (Lippert et al., 2005) dans les ES et/ou les EZ à différents stades de maturation. La réponse des ES à leur environnement de culture peut ainsi être caractérisée au niveau moléculaire par l'identification du transcriptome et du protéome des MEs au cours de la maturation. Tandis-que la maturité de l'ES en fin de développement peut être analysée qualitativement par comparaison de son protéome avec celui de l'EZ mature. En outre, les outils d'analyse protéomique ont déjà été développés pour le mélèze hybride pour des questions de recherche similaires (Teyssier et al., 2011; Teyssier et al., 2014).

Les premières approches de détection de marqueurs de l'embryogenèse somatique chez les conifères ont débuté dans les années 90 ; les techniques utilisées à l'époque étaient surtout du marquage histo-cyto-chimique, de l'électrophorèse en gel de polyacrylamide (SDS-PAGE) et du criblage de banque de clones d'ADNc (**Table I.3**, colonne « technique de détection »). Cela a limité les analyses à quelques marqueurs génétiques ou protéiques ciblés. Puis à partir des années 2000, avec l'essor des techniques de séquençage à haut débit et l'arrivée de l'identification des protéines par spectrométrie de masse, de nouvelles approches d'analyses

globales du protéome et du transcriptome ont été possible. Cependant pour l'embryogenèse du pin maritime, tout ou presque reste à faire.

IV.1. Approche de transcriptomique globale

Des méthodes de génomique fonctionnelle ont été entreprises pour identifier les gènes contribuant à l'ES chez plusieurs espèces de plantes. Les analyses de transcriptomique en particulier possèdent une forte résolution qui permet de visualiser des différences d'expression fine entre les gènes et d'établir un profil d'expression tout au long du développement des ES.

L'essor récent des méthodes d'analyse de l'expression génique, par séquençage d'EST à haut débit ou par la technologie microarray, a permis d'envisager une étude exhaustive et simultanée de l'expression des gènes au cours du développement embryonnaire. En permettant l'analyse de l'expression d'un très grand nombre de gènes, ces études de transcriptomique ont fourni les premières images globales du processus cellulaire impliqué à différentes étapes de l'embryogenèse somatique. L'efficacité de ces techniques repose notamment sur un débit élevé d'analyses. La technique à laquelle nous avons fait appel dans ce travail de thèse repose sur le pyroséquençage à très haut débit des ARNs messagers.

IV.1.1. Séquençage par terminaison cyclique réversible d'Illumina/Solexa

Les ADNc obtenus par reverse transcription des transcrits extraits de nos tissus embryogènes, sont physiquement cassés en fragments de 100 pb environ par nébulisation. La technique de séquençage de l'ADN Illumina/Solexa est basée sur la modification des nucléotides incorporés par l'ADN polymérase. Ils sont modifiés par ajout d'une molécule fluorescente en 5' et d'un terminateur labile en 3' (**Figure I.16**). Un fluorochrome marque spécifiquement chaque type de nucléotide (au nombre de 4). Une étape de lavage élimine les nucléotides modifiés non incorporés. Ensuite, les différents fluorochromes excités par un laser émettent des signaux lumineux spécifiques détectés par le séquenceur puis sont traduits en séquence de nucléotides. Par la suite, le fluorochrome et le terminateur sont clivés pour initier un nouveau cycle de séquençage. De plus, une préparation spécifique qui consiste à isoler chaque fragment d'ADN sur une plaque de verre puis à l'amplifier par réaction de PCR permet de générer de grandes quantités de données. Il va être ainsi possible de séquencer 100 à 200 millions de matrices sur une même plaque. Cette méthode permet de séquencer 18 à 35 milliards de bases en 4 à 9 jours (Metzker, 2010).

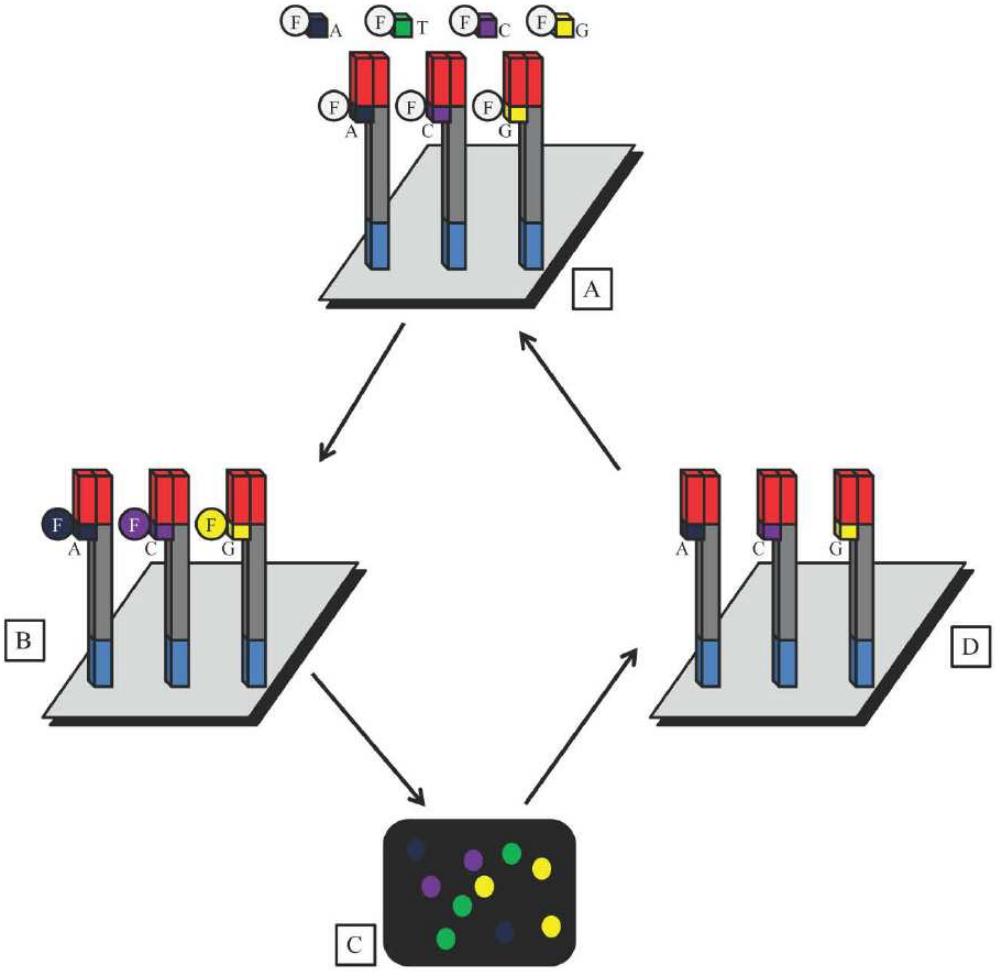


Figure I.16. Principe du séquençage Illumina/Solexa. **A.** Des nucléotides marqués à l'aide de fluorochromes différents sont apportés au mélange réactionnel. Une ADN polymérase va les insérer dans le brin complémentaire du fragment séquencé. La présence en 3' des nucléotides ne permet l'incorporation que d'un seul nucléotide. **B.** Les fluorochromes sont excités grâce à des lasers permettant l'émission d'un signal lumineux, spécifique à chaque nucléotide, qui sera détecté par une caméra. **C.** Exemple de signaux détectés par la caméra du séquenceur. Des programmes informatiques vont traduire les signaux en séquence nucléotidique. **D.** Après détection, le fluorochrome et le terminateur de chaque nucléotides sont retirés pour permettre l'initiation d'un nouveau cycle **(Metzker, 2010)**.

Ce principe de séquençage ADN a obtenu un large succès au cours de ces dernières années, ce qui ne l'exempte pas de certaines limites bien identifiées. **1)** La taille des lectures générées (35-100 pb) est plus faible que pour le pyroséquençage (400 pb au maximum) et que pour la méthode de Sanger (1 kb au maximum). **2)** La fiabilité du séquençage diminue aux extrémités de lectures. Un phénomène de déphasage contribuerait à l'accumulation d'erreurs à l'extrémité 3' de lecture, par extension incomplète du fragment séquencé ou addition de nucléotides multiples au fragment séquencé (Dohm et al., 2008). Une augmentation du bruit de fluorescence résulte de l'accumulation de ces erreurs et entraine une mauvaise interprétation des signaux lumineux par le séquenceur (Erlich et Mitra, 2008). Une fois

séquencés, tous ces fragments (plusieurs millions, en général) doivent être assemblés bioinformatiquement pour retrouver la séquence des transcrits analysés.

IV.1.2. Difficultés de l'assemblage

Les séquences répétées, qui peuvent se retrouver de multiples fois au sein d'un génome séquencé, constituent une difficulté majeure lors d'un assemblage. Lorsque 2 lectures en provenance de 2 régions éloignées se chevauchent parfaitement, l'assembleur ne fait pas de différence et considère qu'elles proviennent de la même région génomique, entrainant la formation de contigs chimériques et de trous dans l'assemblage final **(Figure I.17**). Deux parades ont été développées pour répondre à cette contrainte : **1)** Faire des lectures les plus longues possibles pour mieux définir l'entourage immédiat des séquences très répétées et augmenter la spécificité de lecture, **2)** Mettre en place des banques de séquence dite « *paired-end* ». Les lectures générées sont regroupées en paires complémentaires, chaque paire étant caractérisée par l'orientation et la distance séparant ses lectures (**Figure I.18**). L'objectif est de rajouter 2 critères d'assemblage pour la recherche de chevauchement de séquences par l'assembleur. C'est cette stratégie qui a été mise en œuvre lors de ma thèse.

A cette difficulté technique, s'ajoute les contraintes matérielles inhérentes aux masses de données croissantes générées par les nouvelles techniques de séquençage qui imposent un traitement des données toujours plus conséquent par l'intermédiaire de matériel informatique toujours plus performant. Des méthodes de calcul permettant d'alléger les ressources matérielles demandées mais aussi d'améliorer la prise en compte des répétitions sont désormais incorporées aux assembleurs pour répondre à ces contraintes.

Le nombre de lectures pour chaque paire de fragment ADN cartographiée est compté, et normalisé par la longueur du transcrit et le nombre total de lectures, afin d'obtenir les niveaux d'expression des transcrits en ratios de FPKM (Roberts et al., 2011). Le FPKM mesure l'abondance relative des transcrits en fragments par kilobase d'ARN par million de fragments cartographiés ; c'est un ratio sur lequel on peut se baser pour faire une étude statistique et révéler des différences d'abondance relative entre les transcrits. Cette technique a donc l'énorme avantage d'être quantitative et de permettre l'analyse différentielle. De plus ce type de séquençage permet une analyse fine à haut pouvoir résolutif des transcrits présents dans les tissus analysés. Ainsi même un transcrit présent en faible nombre de copies dans l'échantillon pourra être détecté. La limite est le mode de calcul de l'assemblage des contigs obtenus, qui peut apporter des faux positifs pas forcément décelables.

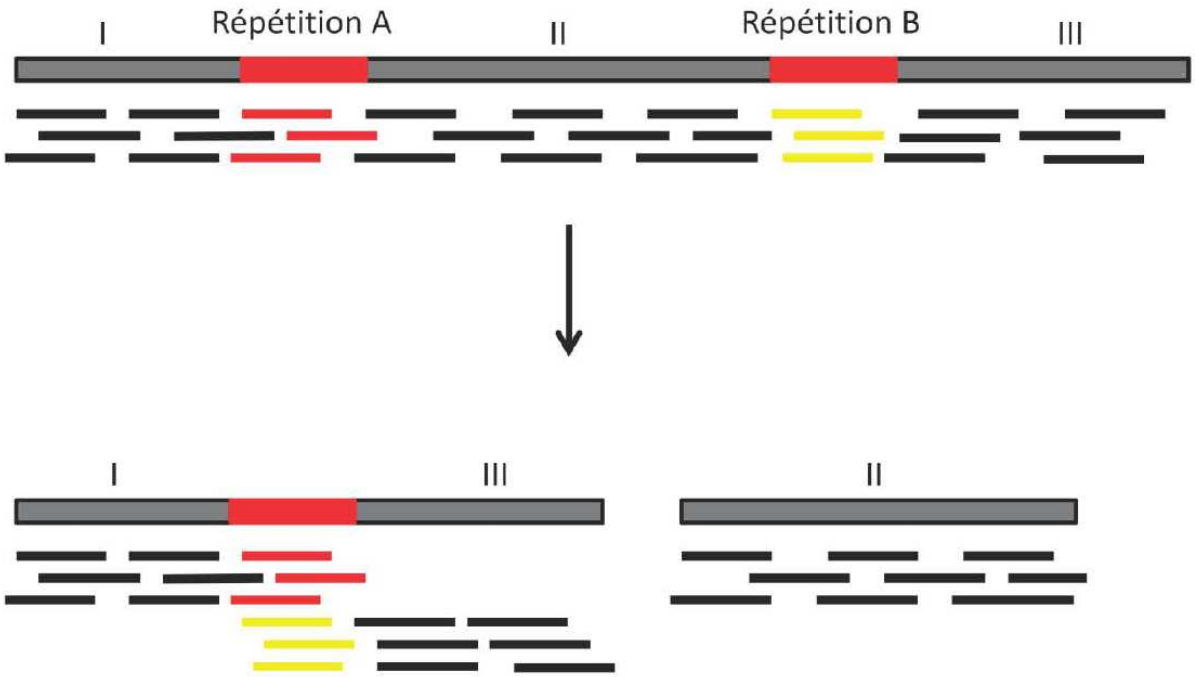


Figure I.17. Schématisation du comportement d'un assembleur vis-à-vis d'une répétition. Pour l'assembleur, les lectures provenant de la répétition A (lectures rouges) et de la répétition B (lectures jaunes) sont supposées provenir de la même région génomique. Cela peut entrainer la formation de contigs chimériques et l'apparition de trous dans l'assemblage final.

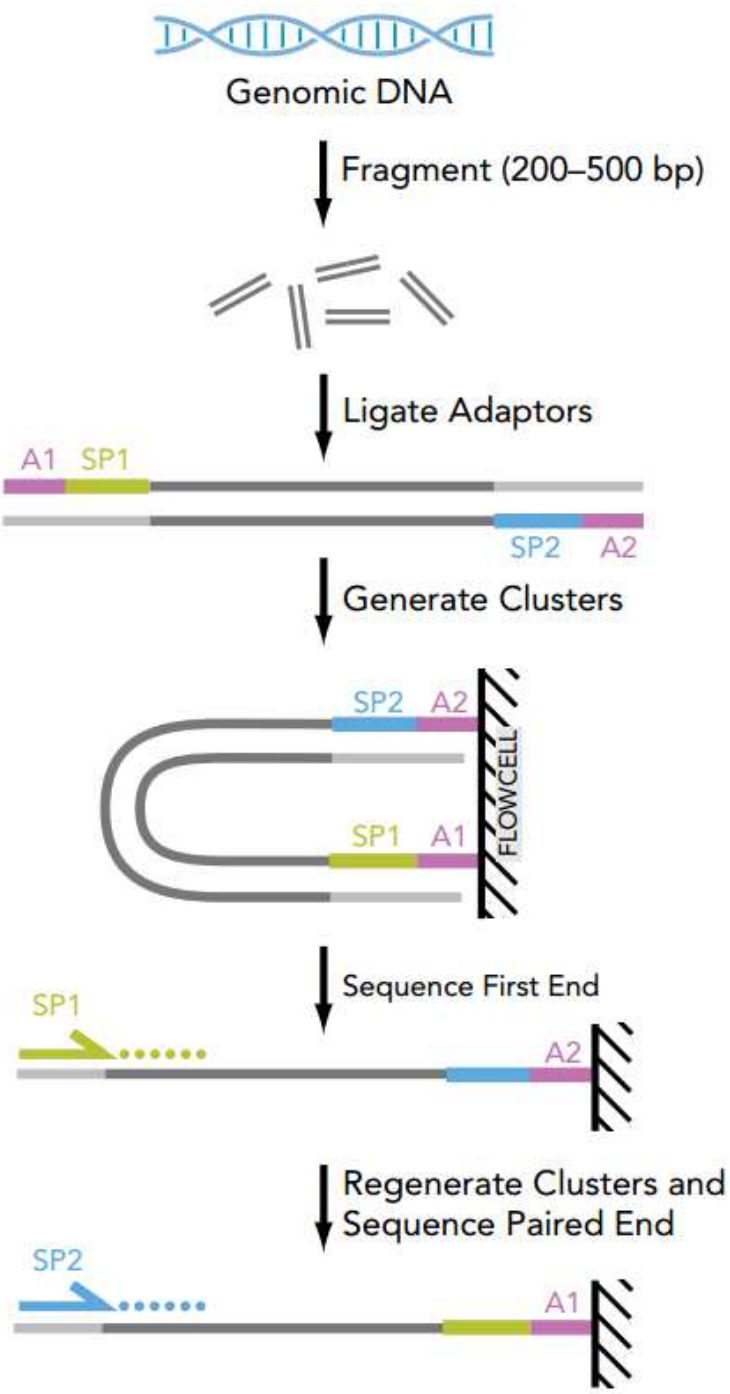


Figure I.18. Protocole de séquençage *paired-end* avec Illumina. Une petite modification à la préparation de la librairie d'ADN single-read permet la lecture sur le brin forward et le brin reverse du fragment d'ADN. Les adaptateurs contenant des séquences d'attachement (A1 et A2) et des sites d'amorce de séquençage (SP1 et SP2) sont liés sur les fragments d'ADN (ici, ADN génomique). La banque résultante de molécules uniques est attachée à une matrice. Chaque extrémité de chaque matrice est lue séquentiellement. Le protocole permet à l'utilisateur de choisir la taille d'insertion de 200 à 500bp. Un run classique en Illumina paired-end permet d'obtenir environ 200 millions de lectures de 2 x 75bp.

IV.2. Approche de protéomique globale

Le « protéome » désigne l'ensemble des protéines exprimées au sein d'un tissu biologique à partir de son génome dans des conditions données et à un moment donné. De ce fait, il existe beaucoup de protéomes potentiels pour un seul génome, selon son expression (Wasinger et al., 1995). Le protéome est une entité dynamique et les données protéomiques permettent de mieux entrevoir le fonctionnement cellulaire, complétant ainsi les données génomiques qui n'offrent qu'une indication du potentiel des cellules.

Aujourd'hui, le concept ancien du gène unique codant une protéine est dépassé. L'étude quantitative des transcrits fournit une connaissance fragmentaire de l'expression du gène, pour plusieurs raisons :

- l'épissage alternatif est la première raison
- une protéine et son ARNm n'ont pas nécessairement le même turnover (cas d'une protéine toujours abondante dans une cellule, alors que son ARNm n'est plus synthétisé) ;
- les protéines subissent des modifications post-traductionnelles (phosphorylations et glycosylations en particulier) qui influent sur leur activité et leur distribution cellulaire ;
- des synergies complexes peuvent se produire entre des regroupements de plusieurs protéines.

Ainsi des relations quantitatives et qualitatives entre le transcrit et son produit de traduction final ne peuvent pas être définies (Gygi et al., 1999; Uesono et Toh-e, 2002). C'est pourquoi, la protéomique se révèle être un outil particulièrement adapté lorsqu'il s'agit de définir des variations développementales et physiologiques en lien avec l'expression des protéines (Zivy et de Vienne, 2000).

IV.2.1. L'électrophorèse bidimensionnelle (2DE)

L'électrophorèse bidimensionnelle (2DE) est actuellement la technique la plus couramment utilisée dans le cadre d'analyses en protéomique. Perfectionnées depuis près de 40 ans (Klose, 1975; O'Farrell, 1975), elle permet de séparer et de visualiser sous forme de « spots », des milliers de protéines en fonction de leurs caractéristiques physico-chimiques. Son essor récent a été dynamisé grâce au développement des techniques ultra sensibles d'identification des protéines par spectrométrie de masse. Cette technique consiste à séparer les protéines extraites selon leur charge (pHi) par isoélectofocalisation (IEF), dans un 1er temps (1ère dimension), puis selon leur masse moléculaire (MM) au travers d'un tamis moléculaire de poly-acrylamide (SDS-PAGE) dans un second temps (2ème dimension). La

dernière étape est la coloration des gels obtenus, puis leur numérisation. Des spots sont définis par analyse informatique, chaque spot étant caractérisé par ses coordonnées de pHi et de MM, et par un volume normalisé par rapport à la coloration (calculé par la surface du spot x intensité du spot) qui traduit l'abondance relative des protéines. Une analyse statistique de ces abondances relatives permet de définir les spots d'intérêts qui seront ensuite excisés, puis identifiés par spectrométrie de masse **(Figure I.19)**.

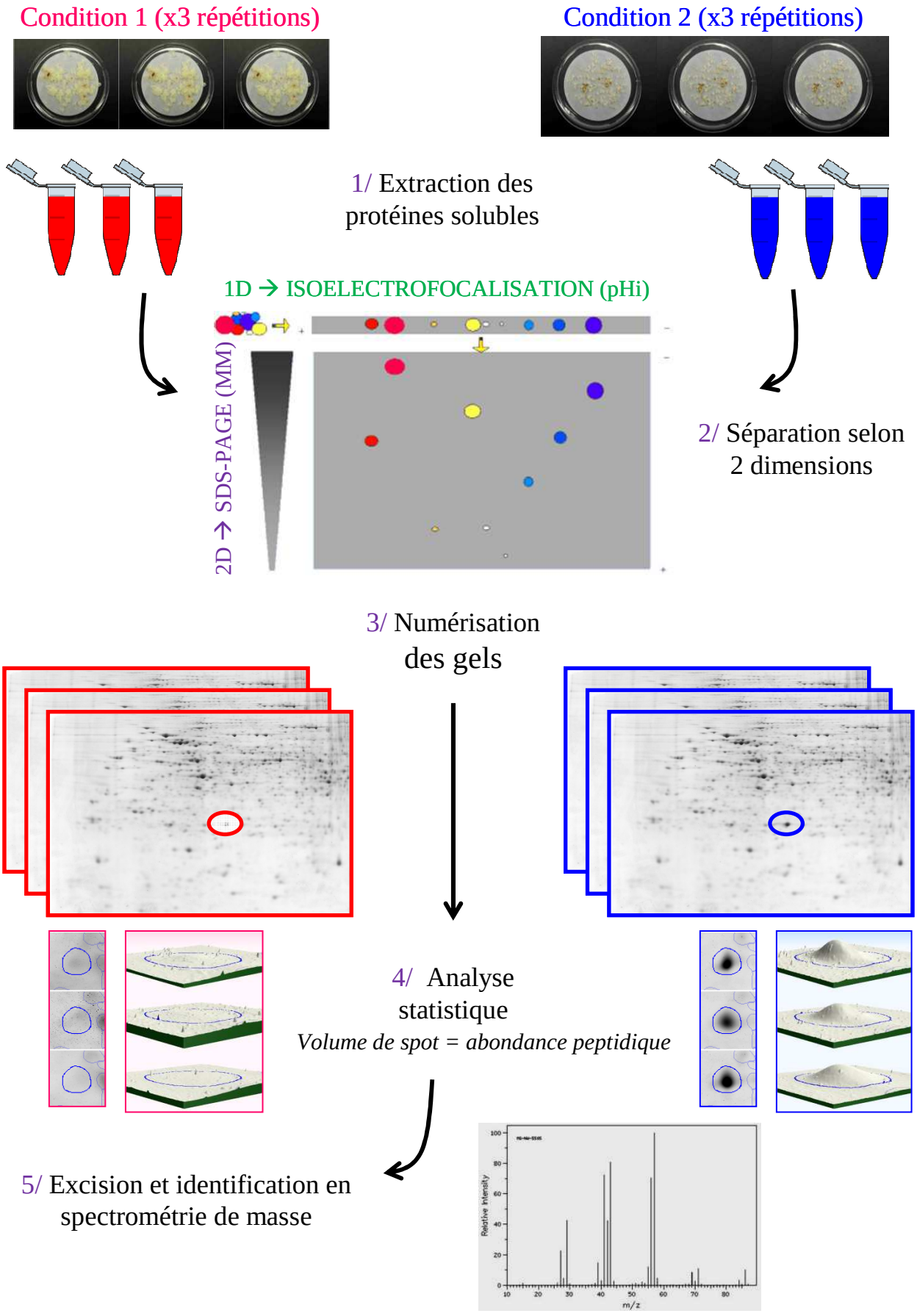


Figure I.19. Etapes de l'analyse différentielle en protéomique.

Aujourd'hui, la technique DIGE (differential in gel-electrophoresis) constitue une évolution pour la protéomique bidimensionnelle ; son principe est la co-migration sur un même gel de plusieurs échantillons protéiques (jusqu'à 3) marqués de façon covalente par trois pigments fluorescents de cyanine de spectre d'absorption différents. L'acquisition d'images du gel unique aux trois longueurs d'onde facilite grandement l'étape d'alignement des gels en augmentant la précision statistique (Unlu et al., 1997).

Une technique alternative à la protéomique bidimensionnelle existe, il s'agit de la protéomique « en vrac » (Shotgun Proteomics). Les protéines extraites sont digérées par une protéase pour produire un mélange de peptides. Le mélange de peptides est ensuite chargé directement sur une colonne micro-capillaire et les peptides sont séparés selon leur hydrophobicité et leur pHi. Les peptides élués de la colonne, sont ensuite ionisés et séparés par leur ratio masse/charge (m/z) dans une première étape de spectrométrie de masse en tandem. Les ions sélectionnés sont fragmentés par collision. Les fragments chargés sont séparés dans une seconde étape de spectrométrie de masse en tandem, permettant la caractérisation de milliers de protéines. Cette technique possède l'avantage de pouvoir détecter des protéines de valeurs de pHi extrême, faiblement abondantes, membranaires. C'est cependant une démarche très onéreuse et très gourmande en moyen d'analyse informatique. Par ailleurs, le problème majeur de cette approche « en vrac » est l'ambigüité pour assigner des peptides communs à différentes protéines, à l'inverse de l'approche par gel bi-dimensionnel où les protéines sont séparées au préalable (Washburn et al., 2001).

Chez les conifères, plusieurs approches de protéomique 2D, pour suivre les protéines accumulées au cours du développement embryonnaire, ont été réalisées chez *Picea glauca* (Lippert et al., 2005), chez *Cunninghamia lanceolata* (Shi et al., 2010), chez *Pinus massoniana* (Zhen et al., 2012) ou chez *Larix* x *eurolepis* (Teyssier et al., 2011; Teyssier et al., 2014) (la **Table I.3** ressence le faible nombre d'études protéomiques effectuées chez les conifères dans ce contexte**)**. Une approche analogue a donc été entreprise lors de mon travail de thèse afin de faire un inventaire quantitatif et qualitatif des protéines lors du développement des ES de pin maritime. Des profils d'abondance de protéines impliquées dans certains stades de développement des embryons ont été mis en évidence et des protéines spécifiques de ces stades ont été identifiées afin de mieux caractériser le développement de l'ES de pin maritime.

IV.2.2. Limites de la protéomique bidimensionnelle

Actuellement, la protéomique 2DE est la technique la plus adaptée et la plus résolutive pour l'étude comparative de l'expression des protéines, cependant elle possède certaines limites.

Plusieurs spots possédant des coordonnées bidimensionnelles différentes (pHi et/ou MM), et donc étant éloignés physiquement sur le gel, peuvent être identifiés en spectrométrie de masse comme correspondant à une protéine identique. Dans certains cas, cela peut être à cause de produits de dégradation de certaines protéines, (Gion et al., 2005; Jorge et al., 2006; Valledor et al., 2008) et qui peuvent co-migrer avec d'autres protéines. Dans d'autres cas, cela peut être à cause de la présence de plusieurs isoformes de la protéine. Ceci a été observé au cours de la maturation des graines d'orge (Finnie et al., 2006). De plus, les redondances d'identification peuvent être la conséquence de modifications post-traductionnelles des protéines, de duplication de gènes, d'épissage alternatif des transcrits, ou encore de variations alléliques, retrouvés chez plusieurs conifères (Gion et al., 2005; Lippert et al., 2005; Balbuena et al., 2009). Enfin, des protéines différentes mais de pHi et MM proches pourront quand même co-migrer et constituer un spot unique conduisant à des identifications multiples.

La technique de protéomique 2D utilisée, est limitée par la gamme de pH de fractionnement utilisée (souvent de pH4 à pH7) et par le choix du type de protéines extraites (solubles généralement). Elle n'est donc pas globale. Elle concerne rarement les protéines membranaires plus difficiles à extraire. La résolution de cette technique constitue une autre limite. Elle est bien moins forte que celle de la transcritpomique. En effet, seules sont observées les protéines les plus différentiellement exprimées parmi les plus représentatives en termes d'abondance dans un extrait donné. De plus le nombre de protéines détectables est de l'ordre du millier (contre plusieurs centaines de milliers de transcrits pour la transcriptomique).

Enfin une dernière limite concerne le grand nombre de répétitions nécessaires pour compenser **1)** des extractions pas toujours parfaitement reproductibles et **2)** des variations de migration qui entrainent des difficultés d'alignement.

IV.3. Annotations des séquences

Les progrès de la génomique des conifères ont été entravés par la grande taille de leurs génomes (30,3 Gb pour le pin maritime) (Murray, 1998). L'absence de génome séquencé chez

le pin maritime limite l'annotation des séquences obtenues pour les 2 types d'analyses (transcrits et protéines). Cependant, d'importants efforts de recherche se sont concentrés sur la caractérisation du transcriptome des conifères. Mais les premiers projets de séquençage d'EST (Expressed Sequence Tag) à grande échelle basé sur la méthode de Sanger ont montré que les transcriptomes de pins et d'épicéas sont très diversifiés et complexes (Allona et al., 1998; Pavy et al., 2005; Cairney et al., 2006; Ralph et al., 2008; Li et al., 2009). L'émergence du séquençage de nouvelle génération a profondément transformé le paysage de l'analyse du génome. Les efforts de recherche chez les conifères ont fourni des catalogues croissants d'EST pour plusieurs espèces d'importance économique et écologique, y compris *Picea. sitchensis* (Ralph et al., 2008), *Picea glauca* (Rigault et al., 2011), *Picea abies* (Chen et al., 2012) et *Pinus pinaster* (Fernandez-Pozo et al., 2011; Canales et al., 2013). Ces catalogues complets de gènes et de séquences codant pour des protéines de conifères et de pin maritime constituent pour notre étude une base de comparaison de séquence déjà très conséquente.

IV.4. Physiologie moléculaire et intégrative

Certaines protéines non-détectées possèdent un transcrit identifié et inversement. Cela peut être dû aux limites analytiques de détection inhérentes aux techniques, mais aussi à des différences de temporalité d'expression (Dembinsky et al., 2007; Faurobert et al., 2007; Pan et al., 2012). En effet, dans un même tissu analysé, le transcriptome contient l'information du futur contenu protéique de la cellule (tous les transcrits ne sont pas traduits, c'est le cas des ARNm régulés négativement, mais aussi des microARN et des petits ARNs interférents…), alors que le protéome représente un niveau de complexité donné dans l'état cellulaire présent (**Figure I.20**). La transcriptomique étant très sensible permet une vision exhaustive de tous les transcrits alors que le nombre de protéines pouvant être identifiées par la protéomique reste limité. Cependant ces 2 manières de caractériser les processus moléculaires impliqués dans la maturation du pin maritime, donnent 2 visions complémentaires peuvant être intégrées l'une avec l'autre pour donner une image plus précise des processus mis en jeu. C'est là tout l'intérêt de recouper les jeux de données de transcriptomiques et de protéomiques par des analyses bioinformatiques de comparaison de séquences. La compréhension des mécanismes ontogénétiques appelle la prise en compte des différents niveaux de complexité biologiques, dont nous avons soumis certains à l'expérimentation. De ce fait, durant ma thèse les analyses moléculaires (transcriptomique, protéomique) ont été complétées par des observations cyto-histologiques (microscopie confocale laser), des analyses physiologiques (mesures de la

teneur en eau, ABA endogène) et biochimiques, (contenu en sucres solubles, en protéines totales).

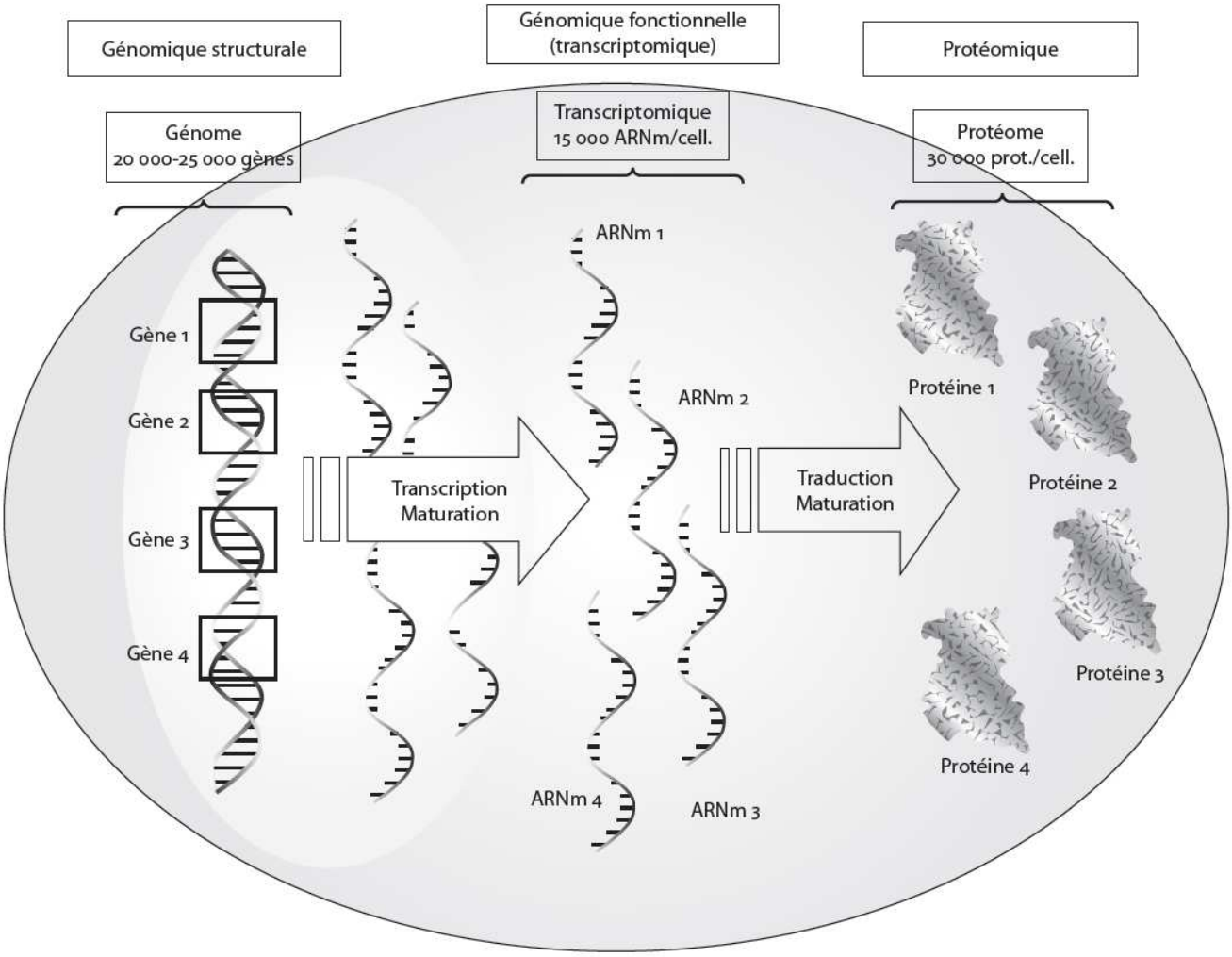


Figure I.20. Relations entre l'étude du transcriptome et du protéome à l'échelle d'une cellule (**Bidart et Lacroix, 2006**)

L'intégration de ces données aux résultats de génomique (transcriptomique et protéomique) complète notre vision sur plusieurs échelles d'étude et permet de mieux cerner les niveaux de complexité dans la cellule (gènes, transcrits, protéines, métabolites), les régulations qui les relient ; mais aussi les niveaux de complexité spatiale dans l'embryon en développement (cellule, tissus, organe, embryon), ainsi que des coordinations entre ces échelles. Les séquences de transcrits et de protéines identifiées donnent accès à des données fonctionnelles (moléculaires, métaboliques, cellulaires), qui peuvent être misent en relation directe avec les données physiologiques et biochimiques, pour reconstituer une partie des

interactions qui se créent entre les différents niveaux de complexité biologiques au cours de l'embryogenèse.

Partie V - L'embryogenèse somatique du pin maritime aujourd'hui : problématique et objectifs

Chez les pins en général, l'embryogenèse somatique est aux portes de l'application (Klimaszewska et al., 2007), d'énormes progrès ayant été réalisés. Chez le pin maritime, plus de 15 ans de travaux ont conduit à des protocoles plus simples et plus efficaces aux étapes d'initiation, de production d'ES et de germination (**Table I.2**). Cependant, l'étape de maturation de l'ES chez le pin maritime n'est pas encore maîtrisée. En effet l'obtention *in fine* en routine et en grandes quantités d'ES cotylédonaire demeure difficile. Par ailleurs, la qualité des ES régénérés reste souvent éloignée de la référence que constitue l'EZ (Jordy et Favre, 2003; Perez Rodriguez et al., 2006). Les plants régénérés à partir d'ES cotylédonaire mature, affichent un retard de croissance par rapport aux semis (Trontin et al., 2013 ; **Figure I.21**). Si les ES ont bien terminé leur programme d'embryogenèse à l'issue de l'étape de maturation, ils ne sont probablement pas complètement matures.

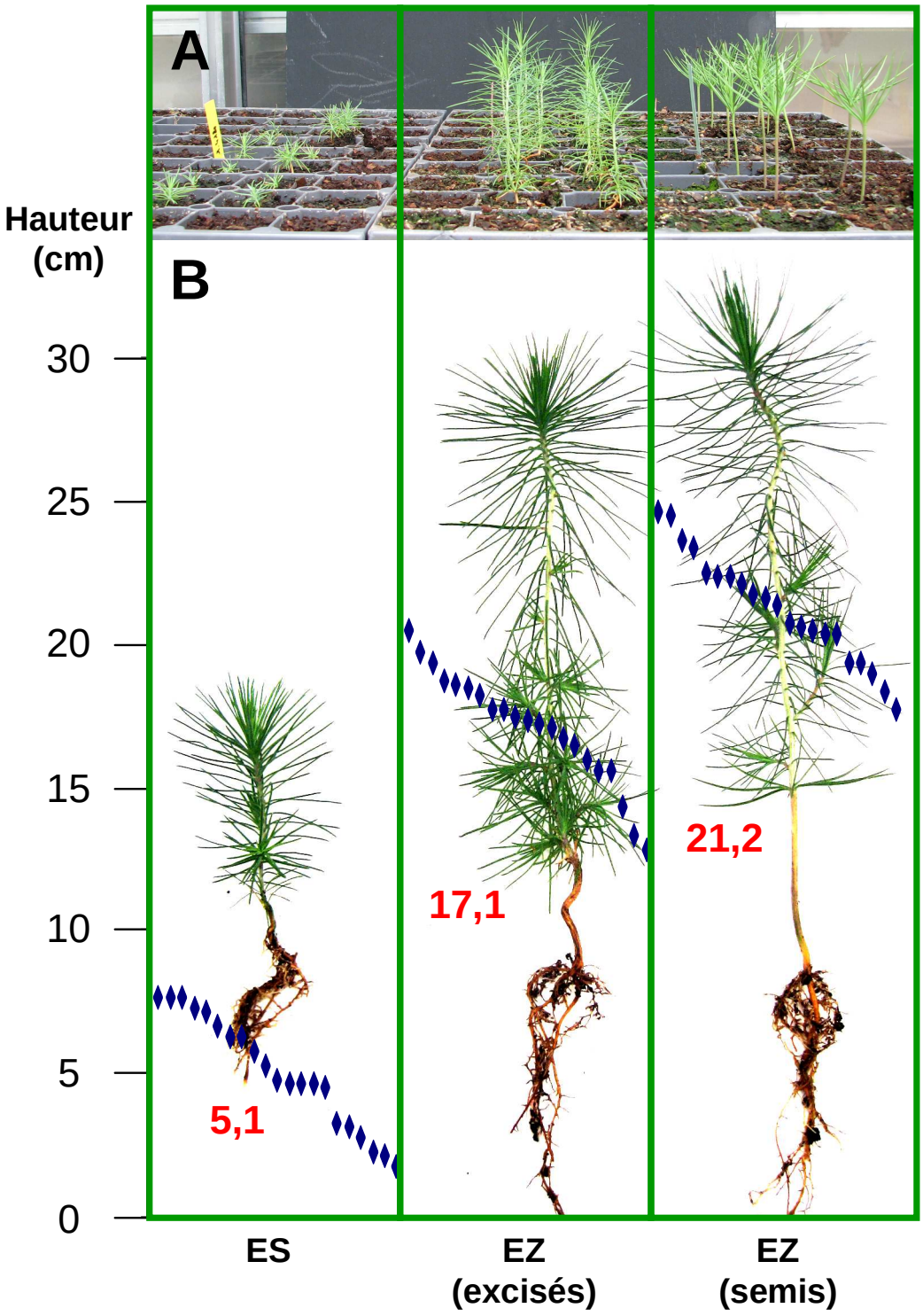


Figure I.21. Comparaison de la croissance des semis issus d'embryons somatiques ou d'embryons zygotiques. Aspect des semis (ES) et zygotiques (EZ) d'une même famille (NL) excisés ou non de la graine (semis) après 6 **(A)** et 22 semaines **(B)** de développement *ex vitro*. ES et EZ excisés ont germé 10 semaines in vitro puis ont été transférés ex vitro au moment du semis témoin. Le graphique **B** présente la variabilité (points bleus) et la hauteur moyenne (en rouge) de chaque catégorie de plants. Les différences observées sont significatives au seuil de 5%. **(Trontin et al., 2013)**.

C'est pourquoi, en regard de ces difficultés **mon premier objectif de thèse** a été d'étudier la différenciation des ES pour une meilleure compréhension des processus physiologiques relatifs à cette différenciation. Ensuite il était important d'évaluer la qualité finale des ES cotylédonaires ce qui constitue mon **second objectif de thèse.** Plus précisément, les questions de recherche liées à ces deux objectifs sont les suivantes (**Figure I.22**).

Question 1 : Comprendre les évènements moléculaires précoces associés au développement des embryons somatiques de pin maritime en condition de disponibilité en eau réduite.

Il est impératif d'augmenter la teneur en gélifiant du milieu de maturation pour obtenir le développement des ES de pin maritime. La forte teneur en gélifiant (Phytagel™, Sigma, voir Annexes) augmente la dureté du milieu et diminue la disponibilité en eau (potentiel matriciel élevé) (Klimaszewska et al., 2000). Ainsi sur le milieu à faible disponibilité en eau (**9G**), des ES cotylédonaires se développent en 12 semaines, tandis-que sur milieu à forte disponibilité en eau (**4G**), les ES restent à l'état immature (microscopiques) et prolifèrent. Nous décrivons ainsi un système inducteur de l'embryogenèse somatique (sur milieu **9G**) et son témoin non-inducteur (milieu **4G**). Ce 1er objectif de ma thèse tente d'élucider les mécanismes moléculaires qui contrôlent la mise en place de l'embryogenèse somatique du pin maritime, ce qui lui confère un aspect très fondamental. Cette étude à ce jour jamais réalisée, est aussi novatrice par les approches méthodologiques mises en œuvre (multi-échelles, du cellulaire au moléculaire). Cette question de recherche sera présentée au travers d'un article de publication dans le chapitre II.

Question 2 : Evaluation de la qualité de développement de l'ES cotylédonaire.

Le critère de prélèvement des ES cotylédonaires aptes à régénérer des plants, a été défini empiriquement sur la base de comparaisons morphologiques proches de l'EZ, constituant la référence en terme d'aptitude à la régénération de plants. Les composés de réserve tels que les protéines de réserve et les sucres, qui s'accumulent lors de la maturation de l'embryon, sont des bons indicateurs de cette aptitude à régénérer un plant. C'est pourquoi nous avons décidé de comparer qualitativement et quantitativement plusieurs composés de réserve de l'ES mature et l'EZ mature de pin maritime et de dresser une liste de marqueurs moléculaires potentiels de l'état de maturation de l'ES. Cette question de recherche représente un aspect plus appliqué de ma thèse et sera présentée au travers d'un article de publication dans le chapitre III.

Augmenter le taux de génotypes réceptifs à l'étape de maturation et porter la qualité des ES au niveau de leur contrepartie zygotique, permettrait d'exprimer au mieux l'énorme potentiel des lignées et de concourir, à terme, à la mise en œuvre de cette technologie chez le pin maritime.

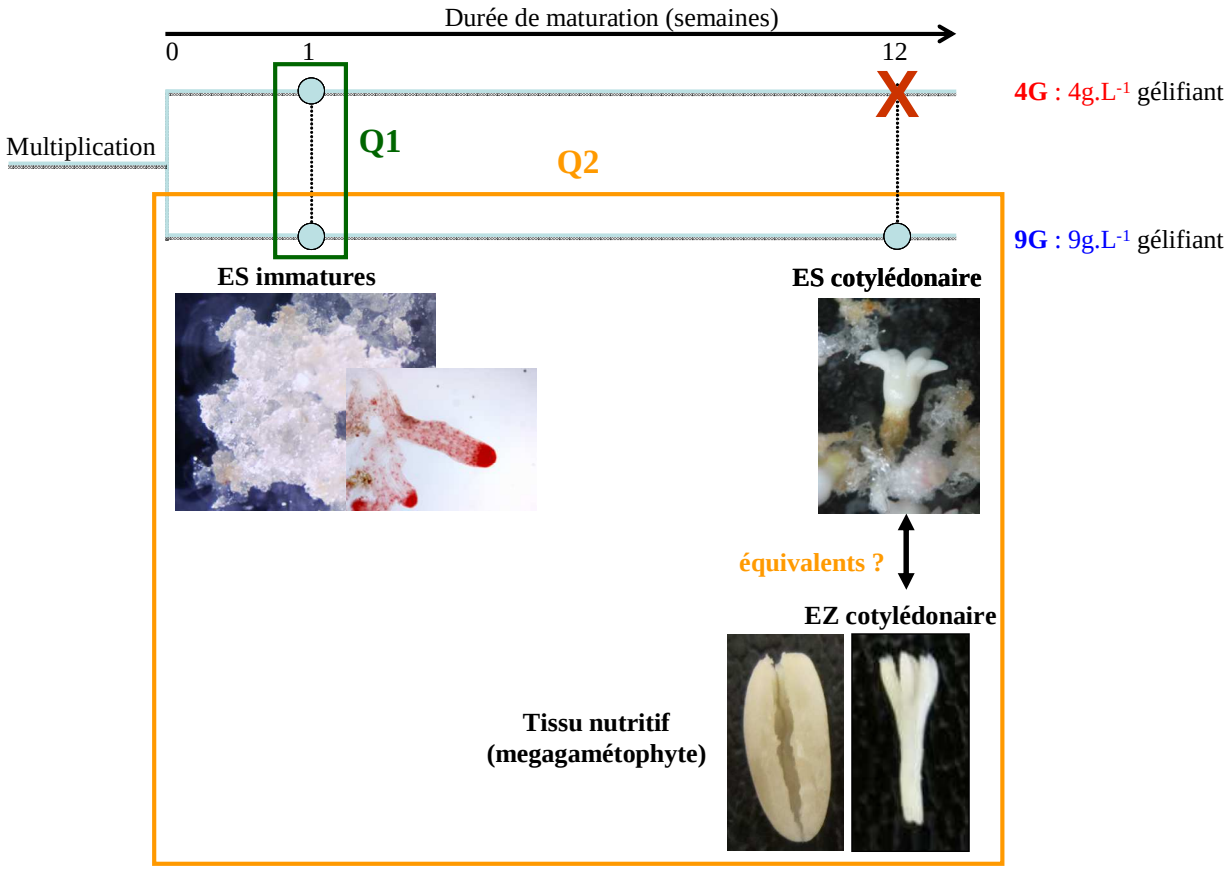


Figure I.22. Physiologie moléculaire du développement des embryons somatiques (**ES**) de pin maritime: positionnement des questions de recherche. **EZ :** embryon zygotique.

Références

Alazard, P., Canteloup, D., Crémière, L., Daubet, A., et Lesgourgues, T. (2005). Genetic breeding of the maritime pine in Aquitaine: an exemplary success story. (PG_Edition, ed.). By Groupe Pin Maritime du Futur, Cestas, France.

Alazard, P., et Raffin, A. (2002). Les gains génétiques des premières et deuxièmes générations de vergers. *In* "Groupe pin maritime du futur eds. *Le progrès génétique en foret*", pp. 27-40. PG éditions.

Alia, R., et Martin, S. (2003). "EUFORGEN Technical Guidelines for genetic conservation and use for Maritime pine (*Pinus pinaster*)," Bioversity International.

Allona, I., Casado, R., et Aragoncillo, C. (1992). Seed storage protein from *Pinus pinaster* : homology of major components with 11S proteins from angiosperms. *Plant Science* **87**, 9-18.

Allona, I., Quinn, M., Shoop, E., Swope, K., Cyr, S. S., Carlis, J., Riedl, J., Retzel, E., Campbell, M. M., et Sederoff, R. (1998). Analysis of xylem formation in pine by cDNA sequencing. *Proceedings of the National Academy of Sciences* **95**, 9693-9698.

Alvarez, J. M., Cortizo, M., Bueno, N., Rodriguez, A., et Ordas, R. J. (2013). CLAVATA1-LIKE, a leucine-rich-repeat protein receptor kinase gene differentially expressed during adventitious caulogenesis in *Pinus pinaster* and *Pinus pinea*. *Plant Cell, Tissue and Organ Culture (PCTOC)* **112**, 331-342.

Aquea, F., et Arce-Johnson, P. (2008). Identification of genes expressed during early somatic embryogenesis in *Pinus radiata*. *Plant Physiology and Biochemistry* **46**, 559-568.

Attree, S. M., et Fowke, L. C. (1993). Embryogeny of gymnosperms: advances in synthetic seed technology of conifers. *Plant Cell, Tissue and Organ Culture* **35**, 1-35 (Revue).

Bailly, C., Audigier, C., Ladonne, F., Wagner, M. H., Coste, F., Corbineau, F., et Côme, D. (2001). Changes in oligosaccharide content and antioxidant enzyme activities in developing bean seeds as related to acquisition of drying tolerance and seed quality. *Journal of Experimental Botany* **52**, 701-708.

Balbuena, T. S., Silveira, V., Junqueira, M., Dias, L. L. C., Santa-Catarina, C., Shevchenko, A., et Floh, E. I. S. (2009). Changes in the 2-DE protein profile during zygotic embryogenesis in the Brazilian Pine (*Araucaria angustifolia*). *Journal of proteomics* **72**, 337-352.

Bercetche, J., et Pâques, M. (1995). Somatic embryogenesis in maritime pine (*Pinus pinaster*). *In* "Somatic Embryogenesis in Woody Plants" (G. P. Jain S, Newton RJ, ed.), Vol. 3 Gymnosperms, pp. 269-285. Kluwer Academic Publishers, The Netherlands.

Bernal-Lugo, I., et Leopold, A. C. (1995). Seed stability during storage: Raffinose content and seed glassy state. *Seed Science Research* **5**, 75-80.

Bernier, F., et Berna, A. (2001). Germins and germin-like proteins: plant do-all proteins. But what do they do exactly? *Plant Physiology and Biochemistry* **39**, 545-554.

Bewley, J. D., et Black, M. (1985). Seed development and maturation. *In* "Seeds", pp. 29-88. Springer.

Bewley, J. D., Bradford, K., et Hilhorst, H. W. M. (2013). Development and maturation. *In* "Seeds. Physiology of development, germination and dormancy" (B. Bewley J.D., K.J, Hilhorst, H.W.M., Nonogaki, H, ed.), pp. 27-83. Springer, New-York.

Bishop-Hurley, S. L., Gardner, R. C., et Walter, C. (2003). Isolation and molecular characterization of genes expressed during somatic embryo development in *Pinus radiata*. *Plant Cell, Tissue and Organ Culture* **74**, 267-281.

Black, M., Corbineau, F., Grzesikit, M., Guyi, P., et Côme, D. (1996). Carbohydrate metabolism in the developing and maturing wheat embryo in relation to its desiccation tolerance. *Journal of Experimental Botany* **47**, 161-169.

Bomal, C., Le, V. Q., et Tremblay, F. M. (2002). Induction of tolerance to fast desiccation in black spruce (*Picea mariana*) somatic embryos: relationship between partial water loss, sugars, and dehydrins. *Physiologia Plantarum* **115**, 523-530.

Bonga, J. M., Klimaszewska, K. K., et von Aderkas, P. (2010). Recalcitrance in clonal propagation, in particular of conifers. *Plant Cell, Tissue and Organ Culture* **100**, 241-254 (Revue).

Bouffier, L., Raffin, A., et Kremer, A. (2009). La gestion de la variabilité génétique dans le programme d'amélioration du pin maritime (*Pinus pinaster* ait.). *Revue Forestière Française* **61**, 369-388.

Boyer, C., Hilbert, J.-L., et Vasseur, J. (1993). Embryogenesis-related protein synthesis and accumulation during early acquisition of somatic embryogenesis competence in *Cichorium*. *Plant Science* **93**, 41-53.

Bozhkov, P. V., Filonova, L. H., Suarez, M. F., et Gerald, P. S. (2005). Programmed Cell Death in Plant Embryogenesis. *In* "Current Topics in Developmental Biology", Vol. Volume 67, pp. 135-179. Academic Press.

Bray, E. A. (1993). Molecular responses to water deficit. *Plant physiology* **103**, 1035.

Breton, D., Harvengt, L., Trontin, J. F., Bouvet, A., et Favre, J. M. (2005). High subculture frequency, maltose-based and hormone-free medium sustained early development of somatic embryos in Maritime pine. *In Vitro Cellular & Developmental Biology - Plant* **41**, 494-504.

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

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

Cairney, J., et Pullman, G. S. (2007). The cellular and molecular biology of conifer embryogenesis. *New Phytologist* **176**, 511-536.

Cairney, J., Zheng, L., Cowels, A., Hsiao, J., Zismann, V., Liu, J., Ouyang, S., Thibaud-Nissen, F., Hamilton, J., Childs, K., Pullman, G. S., Zhang, Y. T., Oh, T., et Buell, C. R. (2006). Expressed Sequence Tags from loblolly pine embryos reveal similarities with angiosperm embryogenesis. *Plant Molecular Biology* **62**, 485-501.

Canales, J., Bautista, R., Label, P., Gomez-Maldonado, J., Lesur, I., Fernández-Pozo, N., Rueda-López, M., Guerrero-fernández, D., Castro-Rodriguez, V., Benzekri, H., Canas, R., Guevara, M.-A., Rodrigues, A., Seoane, P., Teyssier, C., Ehrenmann, F., Morel, A., Le Provost, G., Lalanne, C., Noirot, C., Klopp, C., Raymond, I., Garcia-Gutierrez, A., Trontin, J.-F., Lelu-Walter, M.-A., Miguel, C., Cervera, M. T., canton, F., Plomion, C., Harvengt, L., Avila, C., Claros, M. G., et Canovas, F. (2013). De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. *Plant Biotechnology* (sous presse).

Casson, S. A., et Lindsey, K. (2006). The turnip mutant of Arabidopsis reveals that LEAFY COTYLEDON1 expression mediates the effects of auxin and sugars to promote embryonic cell identity. *Plant physiology* **142**, 526-541.

Chen, J., Uebbing, S., Gyllenstrand, N., Lagercrantz, U., Lascoux, M., et Kallman, T. (2012). Sequencing of the needle transcriptome from Norway spruce (*Picea abies* Karst L.) reveals lower substitution rates, but similar selective constraints in gymnosperms and angiosperms. *BMC genomics* **13**, 589.

Chiwocha, S., et Aderkas, P. v. (2002). Endogenous levels of free and conjugated forms of auxin, cytokinins and abscisic acid during seed development in Douglas fir. *Plant Growth Regulation* **36**, 191-200.

Ciavatta, V. T., Egertsdotter, U., Clapham, D., Arnold, S. v., et Cairney, J. (2002). A promoter from the loblolly pine *PtNIP1;1* gene directs expression in an early-embryogenesis and suspensor-specific fashion. *Planta* **215**, 694-698.

Ciavatta, V. T., Morillon, R., Pullman, G. S., Chrispeels, M. J., et Cairney, J. (2001). An aquaglyceroporin is abundantly expressed early in the development of the suspensor and the embryo proper of loblolly pine. *Plant Physiology* **127**, 1556-1567.

Colin, A., Meredieu, C., Labbe , T., et Belouard, T. (2010). Etude rétrospective et mise à jour de la ressource en pin maritime du massif des Landes de Gascogne après la tempête Klaus du 24 janvier 2009 *Etat des lieux de la ressource du massif landais après la tempête Klaus de 2009 – IGF rapport final – décembre 2010*

Crowe, J. H., Hoekstra, F. A., et Crowe, L. M. (1992). Anhydrobiosis. *Annual Review of Physiology* **54**, 579-599.

de Jong, A. J., Schmidt, E. D. L., et de Vries, S. C. (1993). Early events in higher-plant embryogenesis. *Plant Molecular Biology* **22**, 367-377.

de Vega-Bartol, J. J., Santos, R. R., Simoes, M., et Miguel, C. M. (2013a). Normalizing gene expression by quantitative PCR during somatic embryogenesis in two representative conifer species: *Pinus pinaster* and *Picea abies*. *Plant cell reports*, 1-15.

de Vega-Bartol, J. J., Santos, R. R., Simoes, M., et Miguel, C. M. (2013b). Normalizing gene expression by quantitative PCR during somatic embryogenesis in two representative conifer species: *Pinus pinaster* and *Picea abies*. *Plant Cell Reports* **32**, 715-729.

Dembinsky, D., Woll, K., Saleem, M., Liu, Y., Fu, Y., Borsuk, L. A., Lamkemeyer, T., Fladerer, C., Madlung, J., Barbazuk, B., Nordheim, A., Nettleton, D., Schnable, P. S., et Hochholdinger, F. (2007). Transcriptomic and proteomic analyses of pericycle cells of the maize primary root. *Plant Physiology* **145**, 575-588.

Dodeman, V. L., Ducreux, G., et Kreis, M. (1997). Zygotic embryogenesis versus somatic embryogenesis - a review. *Journal of Experimental Botany* **48**, 1493-1509 (Revue).

Dogra, P. D. (1967). Seed Sterility and disturbances in Embryogeny in conifers with particular reference to seed testing and tree breeding in *Pinaceae*. *In* "Studia Forestia Suecica", Vol. 45, pp. 97, Stockholm.

Dohm, J. C., Lottaz, C., Borodina, T., et Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic acids research* **36**, e105-e105.

Domon, J.-M., Neutelings, G., Roger, D., David, A., et David, H. l. n. (2000). A Basic Chitinase-like Protein Secreted by Embryogenic Tissues of *Pinus caribaea* acts on Arabinogalactan Proteins Extracted from the same Cell Lines. *Journal of Plant Physiology* **156**, 33-39.

Domon, J. M., Dumas, B., Laine, E., Meyer, Y., David, A., et David, H. (1995). Three glycosylated polypeptides secreted by several embryogenic cell cultures of pine show highly specific serological affinity to antibodies directed against the wheat germin apoprotein monomer. *Plant Physiol* **108**, 141-8.

Dong, J., et Dunstan, D. I. (1996). Characterization of three heat-shock-protein genes and their developmental regulation during somatic embryogenesis in white spruce [*Picea glauca* (Moench) Voss]. *Planta* **200**, 85-91.

Dong, J., et Dunstan, D. I. (1997a). Characterization of cDNAs representing five abscisic acid-responsive genes associated with somatic embryogenesis in *Picea glauca*, and their responses to abscisic acid stereostructure. *Planta* **203**, 448-453.

Dong, J., et Dunstan, D. I. (1997b). Endochitinase and beta-1, 3-glucanase genes are developmentally regulated during somatic embryogenesis in *Picea glauca*. *Planta* **201**, 189-194.

Downie, B., et Bewley, J. D. (2000). Soluble sugar content of white spruce (*Picea glauca*) seeds during and after germination. *Physiologia Plantarum* **110**, 1-12.

Dudits, D., Bogre, L., et Gyorgyey, J. (1991). Molecular and cellular approaches to the analysis of plant embryo development from somatic cells *in vitro*. *Journal of Cell Science* **99**, 473-482.

Dunwell, J. M., Gibbings, J. G., Mahmood, T., et Saqlan Naqvi, S. M. (2008). Germin and germin-like proteins: evolution, structure, and function. *Critical Reviews in Plant Sciences* **27**, 342-375 (Revue).

Dunwell, J. M., Khuri, S., et Gane, P. J. (2000). Microbial Relatives of the Seed Storage Proteins of Higher Plants: Conservation of Structure and Diversification of Function during Evolution of the Cupin Superfamily. *Microbiology and Molecular Biology Reviews* **64**, 153-179 (Revue).

Dure, L. (1993). A repeating 11-mer amino acid motif and plant desiccation. *The Plant Journal* **3**, 363-369.

Dure, L., Crouch, M., Harada, J., Ho, T. H. D., Mundy, J., Quatrano, R., Thomas, T., et Sung, Z. R. (1989). Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology* **12**, 475-486.

Dure, L., Greenway, S. C., et Galau, G. A. (1981). Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. *Biochemistry* **20**, 4162-4168.

Egertsdotter, U., et Arnold, S. v. (1995). Importance of arabinogalactan proteins for the development of somatic embryos of Norway spruce (*Picea abies*). *Physiologia Plantarum* **93**, 334-345.

Egertsdotter, U., et von Arnold, S. (1998). Development of somatic embryos in Norway spruce. *Journal of Experimental Botany* **49**, 155-162.

Engelmann, F. (1991). In vitro conservation of tropical plant germplasm - a review. *Euphytica* **57**, 227-243 (Revue).

Erlich, Y., et Mitra, P. P. (2008). Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nature methods* **5**, 679-682.

FAO (2006). Global planted forests thematic study: results and analysis. FAO, A. DEL LUNGO, J. BALL and J. CARLE, Rome.

Faurobert, M., Mihr, C., Bertin, N., Pawlowski, T., Negroni, L., Sommerer, N., et Causse, M. (2007). Major proteome variations associated with cherry tomato pericarp development and ripening. *Plant Physiology* **143**, 1327-1346.

Favennec, J. (1999). Aménagement des forêts littorales: cas des forêts dunaires du littoral atlantique français. *Revue forestière française* **51**, 217-229.

Feher, A., Pasternak, T. P., et Dudits, D. (2003). Transition of somatic plant cells to an embryogenic state. *Plant Cell, Tissue and Organ Culture* **74**, 201-228 (Revue).

Fernandez-Pozo, N., Canales, J., Guerrero-Fernandez, D., Villalobos, D. P., Diaz-Moreno, S. M., Bautista, R., Flores-Monterroso, A., Guevara, M. a., Perdiguero,

P., et Collada, C. (2011). EuroPineDB: a high-coverage web database for maritime pine transcriptome. *BMC genomics* **12**, 366.

Filonova, L. H., Bozhkov, P. V., et Arnold, S. v. (2000). Developmental pathway of somatic embryogenesis in *Picea abies* as revealed by time-lapse tracking. *Journal of Experimental Botany* **51**, 249-264.

Finnie, C., Bak-Jensen, K. S., Laugesen, S., Roepstorff, P., et Svensson, B. (2006). Differential appearance of isoforms and cultivar variation in protein temporal profiles revealed in the maturing barley grain proteome. *Plant Science* **170**, 808-821.

Fischerova, L., Fischer, L., Vondrakova, Z., et Vagner, M. (2008). Expression of the gene encoding transcription factor PaVP1 differs in *Picea abies* embryogenic lines depending on their ability to develop somatic embryos. *Plant cell reports* **27**, 435-441.

Flinn, B. S., Roberts, D. R., et Taylor, I. E. P. (1991). Evaluation of somatic embryos of interior spruce. Characterization and developmental regulation of storage proteins. *Physiologia Plantarum* **82**, 624-632.

Footitt, S., Ingouff, M., Clapham, D., et von Arnold, S. (2003). Expression of the viviparous 1 (*Pavp1*) and p34cdc2 protein kinase (*cdc2Pa*) genes during somatic embryogenesis in Norway spruce (*Picea abies* [L.] Karst). *Journal of experimental botany* **54**, 1711-1719.

Forward, B. S., et Misra, S. (2000). Characterization and expression of the Douglas-fir luminal binding protein (PmBiP). *Planta* **212**, 41-51.

Gaj, M. D., Zhang, S., Harada, J. J., et Lemaux, P. G. (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of *Arabidopsis*. *Planta* **222**, 977-988.

Galau, G. A., Hughes, D. W., et Dure, L. (1986). Abscisic acid induction of cloned cotton late embryogenesis-abundant (LEA) mRNAs. *Plant Molecular Biology* **7**, 155-170.

Gallois, J.-L., Nora, F. R., Mizukami, Y., et Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes & development* **18**, 375-380.

Gifford, D. J. (1988). An electrophoretic analysis of the seed proteins from *Pinus monticola* and eight other species of pine. *Canadian journal of botany* **66**, 1808-1812.

Gion, J. M., Lalanne, C., Le Provost, G., Ferry-Dumazet, H., Paiva, J., Chaumeil, P., Frigerio, J. M., Brach, J., Barré, A., de Daruvar, A., Claverol, S., Bonneu, M., Sommerer, N., Negroni, L., et Plomion, C. (2005). The proteome of maritime pine wood forming tissue. *PROTEOMICS* **5**, 3731-3751.

Goldberg, R. B., Barker, S. J., et Perez-Grau, L. (1989). Regulation of gene expression during plant embryogenesis. *Cell* **56**, 149-160.

Gosslova, Monika, Svobodova, Hana, Lipavska, Helena, Albrechtova, Jana, Vreugdenhil, et Dick (2001). Comparing carbohydrate status during norway spruce seed development and somatic embryo formation. *In Vitro Cellular & Developmental Biology-Plant* **37**, 24-28.

Green, M. J., McLeod, J. K., et Misra, S. (1991). Characterization of Douglas fir protein body composition by SDS-PAGE and electron microscopy. *Plant physiology and biochemistry* **29**, 49-55.

Grzelczak, Z. F., et Lane, B. G. (1984). Signal resistance of a soluble protein to enzymic proteolysis. An unorthodox approach to the isolation and purification of germin, a rare growth-related protein. *Canadian journal of biochemistry and cell biology* **62**, 1351-1353.

Gygi, S. P., Rochon, Y., Franza, B. R., et Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Molecular and cellular biology* **19**, 1720-1730.

Hakman, I. (1993a). Embryology in Norway spruce (*Picea abies*). An analysis of the composition of seed storage proteins and deposition of storage reserves during seed development and somatic embryogenesis. *Physiologia Plantarum* **87**, 148-159.

Hakman, I. (1993b). Embryology in Norway spruce (*Picea abies*). Immunochemical studies on transport of a seed storage protein. *Physiologia Plantarum* **88**, 427-433.

Hakman, I., Fowke, L. C., Arnold, S. v., et Eriksson, T. (1985). The development of somatic embryos in tissue cultures initiated from immature embryos of *Picea abies* (Norway spruce). *Plant Science, Irish Republic* **38**, 53-59.

Hakman, I., Hallberg, H., et Palovaara, J. (2009). The polar auxin transport inhibitor NPA impairs embryo morphology and increases the expression of an auxin efflux facilitator protein PIN during *Picea abies* somatic embryo development. *Tree physiology* **29**, 483-496.

Hara, M., Terashima, S., et Kuboi, T. (2001). Characterization and cryoprotective activity of cold-responsive dehydrin from *Citrus unshiu*. *Journal of Plant Physiology* **158**, 1333-1339.

Helal, N. A. S. (2011). The green revolution via synthetic (artificial) seeds: a review. *Research Journal of Agriculture and Biological Sciences* **7**, 464-477 (Revue).

Hetherington, A. M., et Quatrano, R. S. (1991). Mechanisms of action of abscisic acid at the cellular level. *New Phytologist* **119**, 9-12.

Higgins, T. J. V. (1984). Synthesis and regulation of major proteins in seeds. *Annual Review of Plant Physiology* **35**, 191-221 (Revue).

Hjortswang, H. I., Filonova, L. H., Vahala, T., et Arnold, S. v. (2002). Modified expression of the *Pa18* gene interferes with somatic embryo development in Norway spruce. *Plant Growth Regulation* **38**, 75-82.

Hoekstra, F. A., Golovina, E. A., Tetteroo, F. A. A., et Wolkers, W. F. (2001). Induction of Desiccation Tolerance in Plant Somatic Embryos: How Exclusive Is the Protective Role of Sugars? *Cryobiology* **43**, 140-150.

Hohmann, S., Nielsen, S. r., Weig, A., et Eisenbarth, D. (2000). Role of Aquaporins During Elongation Growth of Castor Bean Seedlings. *In* "Molecular Biology and Physiology of Water and Solute Transport", pp. 357-364. Springer US.

Humanez, A., Blasco, M., Brisa, C., Segura, J., et Arrillaga, I. (2012). Somatic embryogenesis from different tissues of Spanish populations of maritime pine. *Plant Cell, Tissue and Organ Culture (PCTOC)* **111**, 373-383.

Ingouff, M., Farbos, I., Lagercrantz, U., et von Arnold, S. (2001). PaHB1 is an evolutionary conserved HD-GL2 homeobox gene expressed in the protoderm during *Norway spruce* embryo development. *Genesis* **30**, 220-230.

Ingram, J., et Bartels, D. (1996). The molecular basis of dehydration tolerance in plants. *Annual review of plant biology* **47**, 377-403 (Revue).

Iraqi, D., et Tremblay, F. M. (2001). Analysis of carbohydrate metabolism enzymes and cellular contents of sugars and proteins during spruce somatic embryogenesis suggests a regulatory role of exogenous sucrose in embryo development. *Journal of Experimental Botany* **52**, 2301-2311.

Isik, F., Li, B. L., et Frampton, J. (2003). Estimates of additive, dominance and epistatic genetic variances from a clonally replicated test of loblolly pine. *Forest Science* **49**, 77-88.

Jaikaran, A. S., Kennedy, T. D., Dratewka-Kos, E., et Lane, B. G. (1990). Covalently bonded and adventitious glycans in germin. *Journal of Biological Chemistry* **265**, 12503-12512.

Jarlet-Hugues, E. (1989). Recherches sur l'aptitude à l'embryogenèse somatique de matériel juvénile et de matériel issu d'arbres adultes de *Pinus pinaster* Sol., Doctorat d'Université Pierre et Marie Curie, Paris VI, 135 pp.

Jenik, P. D., et Barton, M. K. (2005). Surge and destroy: the role of auxin in plant embryogenesis. *Development* **132**, 3577-3585.

Jensen, U., et Berthold, H. (1989). Legumin-like proteins in gymnosperms. *Phytochemistry* **28**, 1389-1394.

Jensen, U., et Lixue, C. (1991). Abies seed protein profile divergent from other *Pinaceae*. *Taxon*, 435-440.

Jimenez, V. M., et Bangerth, F. (2001). *In vitro* culture and endogenous hormone levels in immature zygotic embryos, endosperm and callus cultures of normal and high-lysine barley genotypes. *Journal of Applied Botany* **75**, 1-7.

Jordy, M. N., et Favre, J. M. (2003). Spatio-Temporal Variations in Starch Accumulation During Germination and Post-Germinative Growth of Zygotic and Somatic Embryos of *Pinus pinaster*. *Biologia Plantarum* **46**, 507-512.

Jorge, I., Navarro, R. M., Lenz, C., Ariza, D., et Jorran, J. (2006). Variation in the holm oak leaf proteome at different plant developmental stages, between provenances and in response to drought stress. *Proteomics* **6**, S207-S214.

Kader, J.-C. (1996). Lipid-transfer proteins in plants. *Annual review of plant biology* **47**, 627-654 (Revue).

Kermode, A. R. (1990). Regulatory mechanisms involved in the transition from seed development to germination. *Critical Reviews in Plant Sciences* **9**, 155-195 (Revue).

Kermode, A. R. (1997). Approaches to elucidate the basis of desiccation-tolerance in seeds. *Seed Science Research* **7**, 75-95.

Kermode, A. R., et Finch-Savage, B. E. (2002). Desiccation sensitivity in orthodox and recalcitrant seeds in relation to development. *In* "Desiccation and survival in plants: drying without dying", pp. 149-184. Cabi, Wallingford, UK.

Kingston-Smith, A. H., Walker, R. P., et Pollock, C. J. (1999). Invertase in leaves: Conundrum or control point ? *Journal of Experimental Botany* **50**, 735-743.

Klimaszewska, K., Bernier-Cardou, M., Cyr, D., et Sutton, B. (2000). Influence of gelling agents on culture medium gel strength, water availability, tissue water potential, and maturation response in embryogenic cultures of *Pinus strobus* L. *In Vitro Cellular & Developmental Biology - Plant* **36**, 279-286.

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

Klimaszewska, K., Noceda, C., Pelletier, G., Label, P., Rodriguez, R., et Lelu-Walter, M. (2009). Biological characterization of young and aged embryogenic cultures of *Pinus pinaster* (Ait.). *In Vitro Cellular & Developmental Biology - Plant* **45**, 20-33.

Klimaszewska, K., Trontin, J.-F., Becwar, M., Devillard, C., Park, Y.-S., et Lelu-Walter, M.-A. (2007). Recent progress in somatic embryogenesis of four *Pinus spp*. *Tree For. Sci. Biotech.* **1**, 11-25 (Revue).

Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. *Humangenetik* **26**, 231-243.

Kong, L., et Aderkas, P. v. (2007). Genotype effects on ABA consumption and somatic embryo maturation in interior spruce (*Picea glauca* * engelmanni). *Journal of Experimental Botany* **58**, 1525-1531.

Konradova, H., Gricova, M., et Lipavska, H. (2003). Cold-induced accumulation of raffinose family oligosaccharides in somatic embryos of Norway spruce (*Picea abies*). *In Vitro Cellular & Developmental Biology - Plant* **39**, 425-427.

Konradova, H., Lipavska, H., Albrechtova, J., et Vreugdenhil, D. (2002). Sucrose metabolism during somatic and zygotic embryogeneses in Norway spruce: content of soluble saccharides and localisation of key enzyme activities. *Journal of Plant Physiology* **159**, 387-396.

Kremer, A. (1992). Predictions of age-age correlations of total height based on serial correlations between height increments in Maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics* **85**, 152-158.

Kreuger, M., et Holst, G.-J. (1993). Arabinogalactan proteins are essential in somatic embryogenesis of *Daucus carota* L. *Planta* **189**, 243-248.

Kwong, R. W., Bui, A. Q., Lee, H., Kwong, L. W., Fischer, R. L., Goldberg, R. B., et Harada, J. J. (2003). LEAFY COTYLEDON1-LIKE Defines a Class of Regulators Essential for Embryo Development. *The Plant Cell Online* **15**, 5-18.

Label, P., et Lelu, M.-A. (2000). Exogenous abscisic acid fate during maturation of hybrid larch (*Larix* x *leptoeuropaea*) somatic embryos. *Physiologia Plantarum* **109**, 456-462.

Lane, B. G. (2000). Oxalate oxidases and differentiating surface structure in wheat: germins. *The Biochemical journal* **349**, 309-321.

Lane, B. G. (2002). Oxalate, Germins, and High Plant Pathogens. *IUBMB life* **53**, 67-75.

Lara-Chavez, A., Egertsdotter, U., et Flinn, B. S. (2012). Comparison of gene expression markers during zygotic and somatic embryogenesis in pine. *In Vitro Cellular & Developmental Biology-Plant* **48**, 341-354.

Larkins, B., Vasil, I., Schwartz, B., Vernon, D., et Meinke, D. (1997). Development of the Suspensor: Differentiation, Communication, and Programmed Cell Death During Plant Embryogenesis. *In* "Cellular and Molecular Biology of Plant Seed Development", Vol. 4, pp. 53-72. Springer Netherlands.

Le, B. H., Cheng, C., Bui, A. Q., Wagmaister, J. A., Henry, K. F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R., et Horvath, S. (2010). Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences* **107**, 8063-8070.

Leal, I., et Misra, S. (1993). Molecular cloning and characterization of a legumin-like storage protein cDNA of Douglas fir seeds. *Plant molecular biology* **21**, 709-715.

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

Lelu-Walter, M.-A., Thompson, D., Harvengt, L., Sanchez, L., Toribio, M., et Pâques, L. (2013). Somatic embryogenesis in forestry with a focus on Europe: state-of-the-art, benefits, challenges and future direction. *Tree Genetics & Genomes* **9**, 883-899 (Revue).

Lelu-Walter, M. A., Bernier-Cardou, M., et Klimaszewska, K. (2006). Simplified and improved somatic embryogenesis for clonal propagation of *Pinus pinaster* (Ait.). *Plant Cell Rep* **25**, 767-76.

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

Lelu, M. A., et Label, P. (1994). Changes in the levels of abscisic acid and its glucose ester conjugate during maturation of hybrid larch (*Larix* × *leptoeuropaea*) somatic embryos, in relation to germination and plantlet recovery. *Physiologia Plantarum* **92**, 53-60.

Li, X., Wu, H. X., Dillon, S. K., et Southerton, S. G. (2009). Generation and analysis of expressed sequence tags from six developing xylem libraries in *Pinus radiata* D. Don. *BMC genomics* **10**, 41.

Linkies, A., Graeber, K., Knight, C., et Leubne-Metzger, G. (2010). The evolution of seeds. *New Phytologist* **186**, 817-831.

Lipavskà, H., et Konradova, H. (2004). Somatic embryogenesis in conifers: the role of carbohydrate metabolism. *In Vitro Cellular & Developmental Biology-Plant* **40**, 23-30 (Revue).

Lipavskà, H., Svobodová, H., Albrechtová, J., Kumstýřová, L., Vágner, M., et Vondráková, Z. (2000). Carbohydrate status during somatic embryo maturation in Norway spruce. *In Vitro Cellular & Developmental Biology - Plant* **36**, 260-267.

Lippert, D., Zhuang, J., Ralph, S., Ellis, D. E., Gilbert, M., Olafson, R., Ritland, K., Ellis, B., Douglas, C. J., et Bohlmann, J. (2005). Proteome analysis of early somatic embryogenesis in *Picea glauca*. *Proteomics* **5**, 461-73.

Manfre, A. J., LaHatte, G. A., Climer, C. R., et Marcotte, W. R. (2009). Seed dehydration and the establishment of desiccation tolerance during seed maturation is altered in the *Arabidopsis thaliana* mutant atem6-1. *Plant and cell physiology* **50**, 243-253.

Mathieu, M., Lelu-Walter, M. A., Blervacq, A. S., David, H., Hawkins, S., et Neutelings, G. (2006). Germin-like genes are expressed during somatic embryogenesis and early development of conifers. *Plant Molecular Biology* **61**, 615-627.

Mathieu, M., Neutelings, G., Hawkins, S., Grenier, E., et David, H. (2003). Cloning of a pine germin-like protein (GLP) gene promoter and analysis of its activity in transgenic tobacco Bright Yellow 2 cells. *Physiologia Plantarum* **117**, 425-434.

Mayer, K. F. X., Schoof, H., Haecker, A., Lenhard, M., Jurgens, G., et Laux, T. (1998). Role of WUSCHEL in Regulating Stem Cell Fate in the *Arabidopsis* Shoot Meristem. *Cell* **95**, 805-815.

Merkle, S. A., Parrott, W. A., et Flinn, B. S. (1995). Morphogenic aspects of somatic embryogenesis. *In* "In vitro embryogenesis in plants", pp. 155-203. Springer.

Merzeau, D., Canteloup, D., et Raffin, A. (2002). Les gains génétiques des premières et deuxième générations de vergers. Le progrès génétique en forêt. *In* "Groupe Pin Maritime du Futur", pp. 49-56. PG Edition.

Metzker, M. L. (2010). Sequencing technologie : the next generation. *Nature Reviews Genetics* **11**, 31-46 (Revue).

Miguel, C., Goncalves, S., Tereso, S., Marum, L., Maroco, J., et Oliveira, M. M. (2004). Somatic embryogenesis from 20 open-pollinated families of Portuguese plus trees of maritime pine. *Plant Cell, Tissue and Organ Culture* **76**, 121-130.

Misra, S., et Green, M. J. (1990a). Developmental gene expression in conifer embryogenesis and germination. I, Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* [moench] voss.). *Plant Science* **68**, 163-173.

Misra, S., et Green, M. J. (1990b). Developmental gene expression in conifer embryogenesis and germination. I. Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* [Moench] Voss.). *Plant Science* **68**, 163-173.

Molina, A., Segura, A., et Garcia-Olmedo, F. (1993). Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. *FEBS letters* **316**, 119-122.

Mondal, T. K., Bhattacharya, A., Sood, A., et Ahuja, P. S. (2002). Factors affecting germination and conversion frequency of somatic embryos of Tea [*Camellia sinensis* (L.) O. Kuntze]. *Journal of plant physiology* **159**, 1317-1321.

Murray, B. G. (1998). Nuclear DNA amounts in gymnosperms. *Annals of Botany* **82**, 3-15.

Neutelings, G., Domon, J., Membré, N., Bernier, F., Meyer, Y., David, A., et David, H. (1998). Characterization of a germin-like protein gene expressed in somatic and zygotic embryos of pine (*Pinus caribaea* Morelet). *Plant Molecular Biology* **38**, 1179-1190.

Newton, C. H., Flinn, B. S., et Sutton, B. C. S. (1992). Vicilin-like seed storage proteins in the gymnosperm interior spruce (*Picea glauca/engelmanii*). *Plant molecular biology* **20**, 315-322.

Nolan, K. E., Kurdyukov, S., et Rose, R. J. (2009). Expression of the SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE1 (SERK1) gene is associated with developmental change in the life cycle of the model legume *Medicago truncatula*. *Journal of experimental botany* **60**, 1759-1771.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *Journal of biological chemistry* **250**, 4007-4021.

Oliver, A. E., Crowe, L. M., et Crowe, J. H. (1998). Methods for dehydration-tolerance: depression of the phase transition temperature in dry membranes and carbohydrate vitrification. *Seed Science Research* **8**, 211-221.

Palmisano, A. M., Ambrogioni, L., et Caroppo, S. (1994). First record of a Bursaphelenchus species from *Pinus pinaster* in Italy1. *EPPO Bulletin* **24**, 467-474.

Palovaara, J., et Hakman, I. (2008). Conifer *WOX*-related homeodomain transcription factors, developmental consideration and expression dynamic of *WOX2* during *Picea abies* somatic embryogenesis. *Plant Molecular Biology* **66**, 533-549.

Palovaara, J., Hallberg, H., Stasolla, C., et Hakman, I. (2010). Comparative expression pattern analysis of *WUSCHEL-related homeobox 2* (*WOX2*) and *WOX8/9* in developing seeds and somatic embryos of the gymnosperm *Picea abies*. *New Phytologist* **188**, 122-135.

Pan, Z., Zeng, Y., An, J., Ye, J., Xu, Q., et Deng, X. (2012). An integrative analysis of transcriptome and proteome provides new insights into carotenoid biosynthesis and regulation in sweet orange fruits. *Journal of Proteomics* **75**, 2670-2684.

Park, S., et Harada, J. J. (2008). *Arabidopsis* embryogenesis. *In* "Plant Embryogenesis", pp. 3-16. Springer.

Park, Y. S., Lelu-Walter, M. A., Harvengt, L., Trontin, J. F., MacEacheron, I., Klimaszewska, K., et Bonga, J. M. (2006). Initiation of somatic embryogenesis in *Pinus banksiana*, *P. strobus*, *P. pinaster*, and *P. sylvestris* at three laboratories in Canada and France. *Plant Cell, Tissue and Organ Culture* **86**, 87-101.

Pavy, N., Paule, C., Parsons, L., Crow, J., Morency, M.-J., Cooke, J., Johnson, J., Noumen, E., Guillet-Claude, C., et Butterfield, Y. (2005). Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC genomics* **6**, 144.

Peace, E. A., Lea, P. J., et Darrall, N. M. (1995). The effect of open-air fumigation with SO2 and O3 on carbohydrate metabolism in Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*). *Plant, Cell & Environment* **18**, 277-283.

Pédezert, G. (2009). La Gironde en bref - Raport INSEE - octobre 2009.

Perez Rodriguez, M. J., Suarez, M. F., Heredia, R., Avila, C., Breton, D., Trontin, J. F., Filonova, L., Bozhkov, P., Arnold, S. v., Harvengt, L., et Canovas, F. M. (2006). Expression patterns of two glutamine synthetase genes in zygotic and somatic pine

embryos support specific roles in nitrogen metabolism during embryogenesis. *New Phytologist* **169**, 35-44.

Pignard, G., Bélouard, T., et Stach, N. (2001). Suivi de la ressource pin maritime en Aquitaine, Union Européenne PDZR, Derf, Conseil Régional d'Aquitaine. *Inventaire Français National*.

Plomion, C., Chagne, D., Pot, D., Kumar, S., Wilcox, P. L., Burdon, R. D., Prat, D., Peterson, D. G., Paiva, J., Chaumeil, P., Vendramin, G. G., Sebastiani, F., Nelson, C. D., Echt, C. S., Savolainen, O., Kubisiak, T. L., Cervera, M. T., Maria, N. d., et Islam-Faridi, M. N. (2007). Pines. *In* "Forest trees", pp. 29-92. Springer-Verlag GmbH, Heidelberg, Germany.

Pot, D., Chantre, G., Rozenberg, P., Rodrigues, J. C., Jones, G. L., Pereira, H., Hannrup, B., Cahalan, C., et Plomion, C. (2002). Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). *Annals of Forest Science* **59**, 563-575.

Pullman, G. S., et Buchanan, M. (2008). Identification and quantitative analysis of stage-specific carbohydrates in loblolly pine (*Pinus taeda*) zygotic embryo and female gametophyte tissues. *Tree Physiology* **28**, 985-96.

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

Quiroz-Figueroa, F., Fuentes-Cerda, C., Rojas-Herrera, R., et Loyola-Vargas, V. (2002). Histological studies on the developmental stages and differentiation of two different somatic embryogenesis systems of *Coffea arabica*. *Plant Cell Reports* **20**, 1141-1149.

Raghavan, V. (2006). Can carrot and *Arabidopsis* serve as model systems to study the molecular biology of somatic embryogenesis? *CURRENT SCIENCE-BANGALORE* **90**, 1336.

Rai, M. K., Shekhawat, N. S., Harish, Gupta, A. K., Phulwaria, M., Ram, K., et Jaiswal, U. (2011). The role of abscisic acid in plant tissue culture: a review of recent progress. *Plant Cell, Tissue and Organ Culture* **106**, 179-190 (Revue).

Ralph, S. G., Chun, H. J. E., Kolosova, N., Cooper, D., Oddy, C., Ritland, C. E., Kirkpatrick, R., Moore, R., Barber, S., et Holt, R. A. (2008). A conifer genomics resource of 200,000 spruce (*Picea* spp.) ESTs and 6,464 high-quality, sequence-finished full-length cDNAs for Sitka spruce (*Picea sitchensis*). *BMC genomics* **9**, 484.

Ramarosandratana, A., Harvengt, L., Bouvet, A., Calvayrac, R., et Pâques, M. (2001a). Effects of carbohydrate source, polyethylene glycol and gellan gum concentration of embryonal-suspensor mass (ESM) proliferation and maturation of maritime pine somatic embryos. *In Vitro Cellular & Developmental Biology - Plant* **37**, 29-34.

Ramarosandratana, A., Harvengt, L., Bouvet, A., Calvayrac, R., et Pâques, M. (2001b). Influence of the embryonal-suspensor mass (ESM) sampling on development and proliferation of maritime pine somatic embryos. *Plant Science* **160**, 473-479.

Renauld, F. (1995). "La fructification du pin maritime : influence du climat, du mode de conduite en verger et outils de prévision de production des cônes." Rapport de fin d'étude ISAB.

Rigault, P., Boyle, B., Lepage, P., Cooke, J. E. K., Bousquet, J., et MacKay, J. J. (2011). A white spruce gene catalog for conifer genome analyses. *Plant physiology* **157**, 14-28.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L., et Pachter, L. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* **12**, R22.

Roberts, D. R., Flinn, B. S., Webb, D. T., Webster, F. B., et Sutton, B. (1990). Abscisic acid and indole-3-butyric acid regulation of maturation and accumulation of storage proteins in somatic embryos of interior spruce. *Physiologia Plantarum* **78**, 355-360.

Sabala, I., Elfstrand, M., Farbos, I., Clapham, D., et von Arnold, S. (2000). Tissue-specific expression of *Pa18*, a putative lipid transfer protein gene, during embryo development in Norway spruce (*Picea abies*). *Plant molecular biology* **42**, 461-478.

Sanchez-Ballesta, M. T., Rodrigo, M. J., Lafuente, M. T., Granell, A., et Zacarias, L. (2004). Dehydrin from citrus, which confers in vitro dehydration and freezing protection activity, is constitutive and highly expressed in the flavedo of fruit but responsive to cold and water stress in leaves. *Journal of agricultural and food chemistry* **52**, 1950-1957.

Satoh, S., Kamada, H., Harada, H., et Fujii, T. (1986). Auxin-controlled glycoprotein release into the medium of embryogenic carrot cells. *Plant physiology* **81**, 931-933.

Schlögl, P. S., dos Santos, A. L. W., do Nascimento Vieira, L., Floh, E. I. S., et Guerra, M. P. (2012). Gene expression during early somatic embryogenesis in Brazilian pine (Araucaria angustifolia (Bert) O. Ktze). *Plant Cell, Tissue and Organ Culture (PCTOC)* **108**, 173-180.

Schmidt, E. D., Guzzo, F., Toonen, M. A., et de Vries, S. C. (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* **124**, 2049-2062.

Schwarzerova, K., Vondrakova, Z., Fischer, L., Borikova, P., Bellinvia, E., Eliasova, K., Havelkova, L., Jindriska, F., Vagner, M., et Opatrny, Z. (2010). The role of actin isoforms in somatic embryogenesis in Norway spruce. *BMC plant biology* **10**, 89.

Sghaier-Hammami, B., Drira, N., et Jorrin-Novo, J. V. (2009). Comparative 2-DE proteomic analysis of date palm (*Phoenix dactylifera* L.) somatic and zygotic embryos. *J Proteomics* **73**, 161-77.

Shah, K., Gadella Jr, T. W. J., van Erp, H., Hecht, V., et de Vries, S. C. (2001). Subcellular Localization and Oligomerization of the *Arabidopsis thaliana* Somatic Embryogenesis Receptor Kinase 1 Protein. *Journal of Molecular Biology* **309**, 641-655.

Shelbourne, C. (1997). Genetics of adding value to the end-products of radiata pine. *In* "IUFRO 97 Genetics of Radiata Pine, Proceedings of NZ FRI-IUFRO Conference 1-5 December" (M. J. Burdon RD, ed.), Vol. 203, pp. 129-141. FRI Bull, Rotorua, New Zealand.

Shi, J., Zhen, Y., et Zheng, R.-H. (2010). Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. *Journal of Experimental Botany* **61**, 2367-2381.

Shih, M.-D., Hoekstra, F. A., et Hsing, Y.-I. C. (2008). Late embryogenesis abundant proteins. *Advances in Botanical Research* **48**, 211-255.

Shotwell, M. A., et Larkins, B. A. (1989). The biochemistry and molecular biology of seed storage proteins. *The biochemistry of plants: a comprehensive treatise*.

Singh, H. (1978). Embryology of gymnosperms. *In* "Handbuch der Pflanzenanatomie (Encyclopedia of Plant Anatomy)", Vol. 10, Part 2. Berlin, Germany Gebruder Borntraeger.

Skriver, K., et Mundy, J. (1990). Gene expression in response to abscisic acid and osmotic stress. *The Plant Cell* **2**, 503.

Stasolla, C., Belmonte, M. F., Zyl, L. v., Craig, D. L., Liu, W. B., Yeung, E. C., et Sederoff, R. R. (2004). The effect of reduced glutathione on morphology and gene expression of white spruce (*Picea glauca*) somatic embryos. *Journal of Experimental Botany* **55**, 695-709.

Stasolla, C., et Yeung, E. C. (2003). Recent advances in conifer somatic embryogenesis: improving somatic embryo quality. *Plant Cell, Tissue and Organ Culture* **74**, 15-35 (Revue).

Steiner, N., Santa-Catarina, C., Guerra, M. P., Cutri, L., Dornelas, M. C., et Floh, E. I. S. (2012). A gymnosperm homolog of SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE-1 (*SERK1*) is expressed during somatic embryogenesis. *Plant Cell, Tissue and Organ Culture* **109**, 41-50.

Sterk, P., Booij, H., Schellekens, G. A., Van Kammen, A., et de Vries, S. C. (1991). Cell-specific expression of the carrot EP2 lipid transfer protein gene. *The Plant Cell Online* **3**, 907-21.

Steward, F. C., Marion, O., Smith, M., et Smith, J. (1958). Growth and Organized Development of Cultured Cells. I. Growth and Division of Freely Suspended Cells. *American Journal of Botany* **45**, 693-703.

Stone, S. L., Braybrook, S. A., Paula, S. L., Kwong, L. W., Meuser, J., Pelletier, J., Hsieh, T.-F., Fischer, R. L., Goldberg, R. B., et Harada, J. J. (2008). *Arabidopsis* LEAFY COTYLEDON2 induces maturation traits and auxin activity: implications for somatic embryogenesis. *Proceedings of the National Academy of Sciences* **105**, 3151-3156.

Stone, S. L., Kwong, L. W., Yee, K. M., Pelletier, J., Lepiniec, L. c., Fischer, R. L., Goldberg, R. B., et Harada, J. J. (2001). LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proceedings of the National Academy of Sciences* **98**, 11806-11811.

Suarez, M. F., Filonova, L. H., Smertenko, A., Savenkov, E. I., Clapham, D. H., Arnold, S. v., Zhivotovsky, B., et Bozhkov, P. V. (2004). Metacaspase-dependent programmed cell death is essential for plant embryogenesis. *Current Biology* **14**, R339-R340.

Sunderlikova, V., Salaj, J., Matusikova, I., et Wilhelm, E. (2009). Isolation and characterization of an embryo-specific *Em*-like gene of pedunculate oak (*Quercus robur* L.) and its temporal and spatial expression patterns during somatic and zygotic embryo development. *Trees* **23**, 135-144.

Sung, D. Y., Kaplan, F., et Guy, C. L. (2001). Plant Hsp70 molecular chaperones: protein structure, gene family, expression and function. *Physiologia plantarum* **113**, 443-451.

Sutton, B. (2002). Commercial delivery of genetic improvement to conifer plantations using somatic embryogenesis. *Ann. For. Sci.* **59**, 657-661.

Tahir, M., Law, D. A., et Stasolla, C. (2006). Molecular characterization of *PgAGO*, a novel conifer gene of the ARGONAUTE family expressed in apical cells and required for somatic embryo development in spruce. *Tree physiology* **26**, 1257-1270.

Tang, G.-Q., Lüscher, M., et Sturm, A. (1999). Antisense Repression of Vacuolar and Cell Wall Invertase in Transgenic Carrot Alters Early Plant Development and Sucrose Partitioning. *The Plant Cell Online* **11**, 177-190.

Tautorus, T. E., Fowke, L. C., et Dunstan, D. I. (1991). Somatic embryogenesis in conifers. *Canadian Journal of Botany* **69**, 1873-1899 (Revue).

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

Teyssier, C., Grondin, C., Bonhomme, L., Lomenech, A.-M., Vallance, M., Morabito, D., Label, P., et Lelu-Walter, M.-A. (2011). Increased gelling agent concentration promotes somatic embryo maturation in hybrid larch (*Larix×eurolepsis*): a 2-DE proteomic analysis. *Physiologia Plantarum* **141**, 152-165.

Teyssier, C., Maury, S., Beaufour, M., Grondin, C., Delaunay, A., Le Metté, C., Ader, K., Cadene, M., Label, P., et Lelu-Walter, M.-A. (2014). In search of markers for somatic embryo maturation in hybrid larch (*Larix— eurolepis*): global DNA methylation and proteomic analyses. *Physiologia Plantarum* **150**, 271-291.

Thibaud-Nissen, F., Shealy, R. T., Khanna, A., et Vodkin, L. O. (2003). Clustering of Microarray Data Reveals Transcript Patterns Associated with Somatic Embryogenesis in Soybean. *Plant Physiology* **132**, 118-136.

Thoma, S., Hecht, U., Kippers, A., Botella, J., de Vries, S., et Somerville, C. (1994). Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from *Arabidopsis*. *Plant Physiology* **105**, 35-45.

Thomashow, M. F. (1999). Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annual review of plant biology* **50**, 571-599 (Revue).

Thompson, E. W., et Lane, B. G. (1980). Relation of protein synthesis in imbibing wheat embryos to the cell-free translational capacities of bulk mRNA from dry and imbibing embryos. *Journal of Biological Chemistry* **255**, 5965-5970.

Trontin, J. F., Canlet, F., Reymond, I., Debille, S., Durandeau, K., Harvengt, L., Rousseau, J., de Boisseson, J., Fraysse, J., Alazard, P., Bailly, A., Teyssier, C., Le Metté, C., Morel, A., Label, P., et Lelu-Walter, M. (2013). L'embryogenèse somatique : une méthode de multiplication végétative du pin maritime pour demain ? *FCBA - INFO - Août 2013*.

Tunnacliffe, A., et Wise, M. J. (2007). The continuing conundrum of the LEA proteins. *Naturwissenschaften* **94**, 791-812.

Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T. C., McElver, J., Aux, G., et Patton, D. (2004). Identification of genes required for embryo development in *Arabidopsis*. *Plant Physiology* **135**, 1206-1220.

Uddenberg, D., Valladares, S., Abrahamsson, M., Sundstrom, J., Sundas-Larsson, A., et von Arnold, S. (2011). Embryogenic potential and expression of embryogenesis-related genes in conifers are affected by treatment with a histone deacetylase inhibitor. *In* "BMC Proceedings", Vol. 5, pp. P151. BioMed Central Ltd.

Uesono, Y., et Toh-e, A. (2002). Transient inhibition of translation initiation by osmotic stress. *Journal of Biological Chemistry* **277**, 13848-13855.

Unlu, M., Morgan, M. E., et Minden, J. S. (1997). Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071-2077.

Vales, T., Feng, X., Ge, L., Xu, N., Cairney, J., Pullman, G. S., et Peter, G. F. (2007). Improved somatic embryo maturation in loblolly pine by monitoring ABA-responsive gene expression. *Plant cell reports* **26**, 133-143.

Valledor, L., Castillejo, M. A., Lenz, C., Rodriguez, R., Canal, M. J., et Jorrin, J. (2008). Proteomic analysis of *Pinus radiata* needles: 2-DE map and protein identification by LC/MS/MS and substitution-tolerant database searching. *Journal of Proteome Research* **7**, 2616-2631.

van Hengel, A. J., van Kammen, A., et de Vries, S. C. (2002). A relationship between seed development, Arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis. *Physiologia Plantarum* **114**, 637-644.

Vestman, D., Larsson, E., Uddenberg, D., Cairney, J., Clapham, D., Sundberg, E., et von Arnold, S. (2011). Important processes during differentiation and early development of somatic embryos of Norway spruce as revealed by changes in global gene expression. *In* "BMC Proceedings", Vol. 5, pp. P78. BioMed Central Ltd.

von Aderkas, P., Bonga, J., Klimaszewska, K., et Owens, J. (1991). Comparison of Larch Embryogeny *In Vivo* and *In Vitro*. *In* "Woody Plant Biotechnology", Vol. 210, pp. 139-155. Springer US.

von Aderkas, P., et Bonga, J. M. (2000). Influencing micropropagation and somatic embryogenesis in mature trees by manipulation of phase change, stress and culture environment. *Tree Physiology* **20**, 921-928 (Revue).

von Aderkas, P., Lelu, M. A., et Label, P. (2001). Plant growth regulator levels during maturation of larch somatic embryos. *Plant Physiology and Biochemistry* **39**, 495-502.

von Arnold, S., Bozhkov, P., Clapham, D., Dyachok, J., Filonova, L., Högberg, K.-A., Ingouff, M., et Wiweger, M. (2005). Propagation of Norway spruce via somatic embryogenesis. *Plant Cell, Tissue and Organ Culture* **81**, 323-329.

von Arnold, S., Sabala, I., Bozhkov, P., Dyachok, J., et Filonova, L. (2002). Developmental pathways of somatic embryogenesis. *Plant Cell, Tissue and Organ Culture* **69**, 233-249.

Wang, W., Vinocur, B., Shoseyov, O., et Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in plant science* **9**, 244-252.

Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., et Humphery-Smith, I. (1995). Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* **16**, 1090-1094.

Winter, H., et Huber, S. C. (2000). Regulation of Sucrose Metabolism in Higher Plants : Localization and Regulation of Activity of Key Enzymes. *Critical Reviews in Biochemistry and Molecular Biology* **35**, 253-289 (Revue).

Wise, M. J. (2003). LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *Bmc Bioinformatics* **4**, 52.

Wise, M. J., et Tunnacliffe, A. (2004). POPP the question: what *do* LEA proteins do? *Trends in plant science* **9**, 13-17.

Wiweger, M., Farbos, I., Ingouff, M., Lagercrantz, U., et Von Arnold, S. (2003). Expression of *Chia4-Pa* chitinase genes during somatic and zygotic embryo development in Norway spruce (*Picea abies*): similarities and differences between gymnosperm and angiosperm class IV chitinases. *Journal of experimental botany* **54**, 2691-2699.

Wu, X. L., Chory, J., et Weigel, D. (2007). Combinations of *WOX* activities regulate tissue proliferation during *Arabidopsis* embryonic development. *Developmental Biology* **309**, 306-316.

Yan, G., Xi, M., Wang, G., Yang, L., et Shi, J. (2010). Molecular characteristics and expression analysis of *PmSERK1* during somatic embryogenesis in Masson pine. *Molecular Plant Breeding* **8**, 53-58.

Yang, X., et Zhang, X. (2010). Regulation of Somatic Embryogenesis in Higher Plants. *Critical Reviews in Plant Sciences* **29**, 36-57 (Revue).

Zeevaart, J. A. D., et Creelman, R. A. (1988). Metabolism and physiology of abscisic acid. *Annual review of plant physiology and plant molecular biology* **39**, 439-473.

Zhen, Y., Zhao, Z.-Z., Zheng, R.-H., et Shi, J. (2012). Proteomic analysis of early seed development in *Pinus massoniana* L. *Plant Physiology and Biochemistry* **54**, 97-104.

Zimmerman, J. L. (1993). Somatic embryogenesis: a model for early development in higher plants. *Plant Cell* **5**, 1411-1423 (Revue).

Zivy, M., et de Vienne, D. (2000). Proteomics: a link between genomics, genetics and physiology. *Plant molecular biology* **44**, 575-580.

CHAPITRE II - Événements moléculaires précoces impliqués dans le développement des embryons somatiques de *Pinus pinaster* Ait. en réponse à la disponibilité réduite de l'eau : analyses transcriptomique et protéomique

II.1. Présentation de la publication

La première question de recherche de ma thèse concerne l'effet de la disponibilité en eau dans leur milieu de maturation sur la différenciation des ES. Le développement *in fine* d'ES cotylédonaires de pin sur milieu de maturation ne s'obtient qu'en présence d'une concentration élevée en gélifiant (9 g L^{-1}, condition notée **9G**). A l'inverse, en présence d'une faible concentration (4 g L^{-1}, condition notée **4G**), les ES restent à l'état immature et prolifèrent. Chez *Pinus strobus*, il a été montré qu'une forte concentration en gélifiant a pour conséquence de diminuer la disponibilité en eau du milieu de maturation vers les ES immatures (Klimaszewska et al., 2000). À ce jour, les mécanismes moléculaires sous-jacents de l'impact de la teneur en gélifiant sur le développement des ES ne sont pas élucidés car pas étudiés. Notre première question scientifique est de comprendre les évènements moléculaires précoces associés au développement des ES de pin maritime en condition de disponibilité en eau réduite. Notre objectif est double : **1)** comprendre comment très tôt les ES immatures s'adaptent à leur environnement de culture, et parviennent à se développer et, **2)** identifier les mécanismes moléculaires qui contrôlent le développement de l'ES. Pour cela, nous disposons d'un système d'étude reposant sur l'effet d'un seul facteur clé : la teneur en gélifiant dans le milieu de culture. Une forte teneur en gélifiant (**9G**) permet la différenciation des ES immatures, contrairement à une teneur plus faible (**4G**). Notre stratégie d'échantillonnage a été conçue pour décrire les mécanismes moléculaires précoces qui se mettent en place (au niveau des transcrits et des protéines) en réponse aux changements de conditions de culture (**Figure II.1**).

Plusieurs marqueurs du développement précoce des embryons des conifères ont été décrits :

- Les variations des spectres d'accumulation des glucides et de l'amidon sont révélatrices de changements métaboliques profonds (Lipavskà et Konradova, 2004; Tereso et al., 2007; Pullman et Buchanan, 2008).
- Les variations d'abondance relative de certaines protéines clés vont constituer des marqueurs spécifiques des changements majeurs dans les processus moléculaires et cellulaires (Lippert et al., 2005; Shi et al., 2010; Zhen et al., 2012).
- L'acide abscissique est un PGR (Plant Growth Regulator) régulant de nombreux processus physiologiques, les variations de son contenu endogène révèlent des changements physiologiques importants (Kong et Aderkas, 2007; Vales et al., 2007).

L'étude de la régulation des gènes au niveau des transcrits nous donne un aperçu fragmentaire des réponses des ES aux conditions de culture. Elle a donc été complétée par une étude protéomique, la combinaison des deux approches permettant de caractériser plus précisément cette réponse. Les 2 systèmes de culture (**9G** et **4G**) ont ainsi été comparés au niveau de leur transcriptome (séquençage illumina des ADNc) et de leur protéome (protéomique bidimensionnelle). De plus des caractérisations physiologique (teneur en eau), histologique (ultrastructure microscopique) et biochimiques (sucres solubles, amidon, protéines totales et ABA endogène) ont complété l'étude. L'intégration de l'ensemble de ces données a permis d'obtenir une vision multi-échelle des mécanismes biologiques mis en jeu précocement au cours de la différenciation des ES.

Chez le mélèze hybride, des ES immatures cultivées sur milieu à forte teneur en gélifiant (8 g L^{-1}) montraient une diminution de leur teneur en eau et une augmentation de leur potentiel osmotique, contrairement à des ES cultivées sur milieu à faible teneur en gélifiant (4 g L^{-1}), soulignant la déshydratation des ES qui s'opérait (Teyssier et al., 2011). Toujours chez le mélèze hybride, une augmentation transitoire de l'ABA endogène des ES est détectée rapidement seulement 6h après qu'ils aient été soumis à un traitement de déshydratation (Dronne et al., 1997). Nous supposons donc que les ES immatures de pin maritime en différenciation devraient répondre rapidement à leurs nouvel environnement hydrique. C'est pourquoi nous avons analysé la teneur en eau et en ABA endogène au cours des 4 premières semaines de maturation sur milieux **9G** et **4G**. Par ailleurs, nous savons que les ES immatures

de pin maritime après 1 semaine de maturation sur milieux **9G** ou **4G**, affichent des différences macroscopiques flagrantes. De plus chez le mélèze hybride la comparaison protéomique d'ES immatures après 1 semaine de maturation sur milieu à 4 g L^{-1} ou 8 g L^{-1} de gélifiant, a mis en évidence des marqueurs protéiques associés à la réponse au stress abiotique sur le milieu à faible teneur en gélifiant (métabolisme secondaire, réponse au stress oxydatif et HSP. D'après ces données, nous supposons qu'à un stade très précoce de maturation, les ES immatures de pin maritime exprimeront un programme génétique différent selon leur environnement de maturation (**9G** et **4G**), et que ces différences seront visibles en transcriptomique et en protéomique. Ainsi, nous avons choisi de focaliser nos analyses moléculaires à un stade précoce de 1 semaine de maturation. Le plan expérimental présenté dans la **Figure II.1** résume l'ensemble des analyses réalisées pour répondre à notre premier objectif.

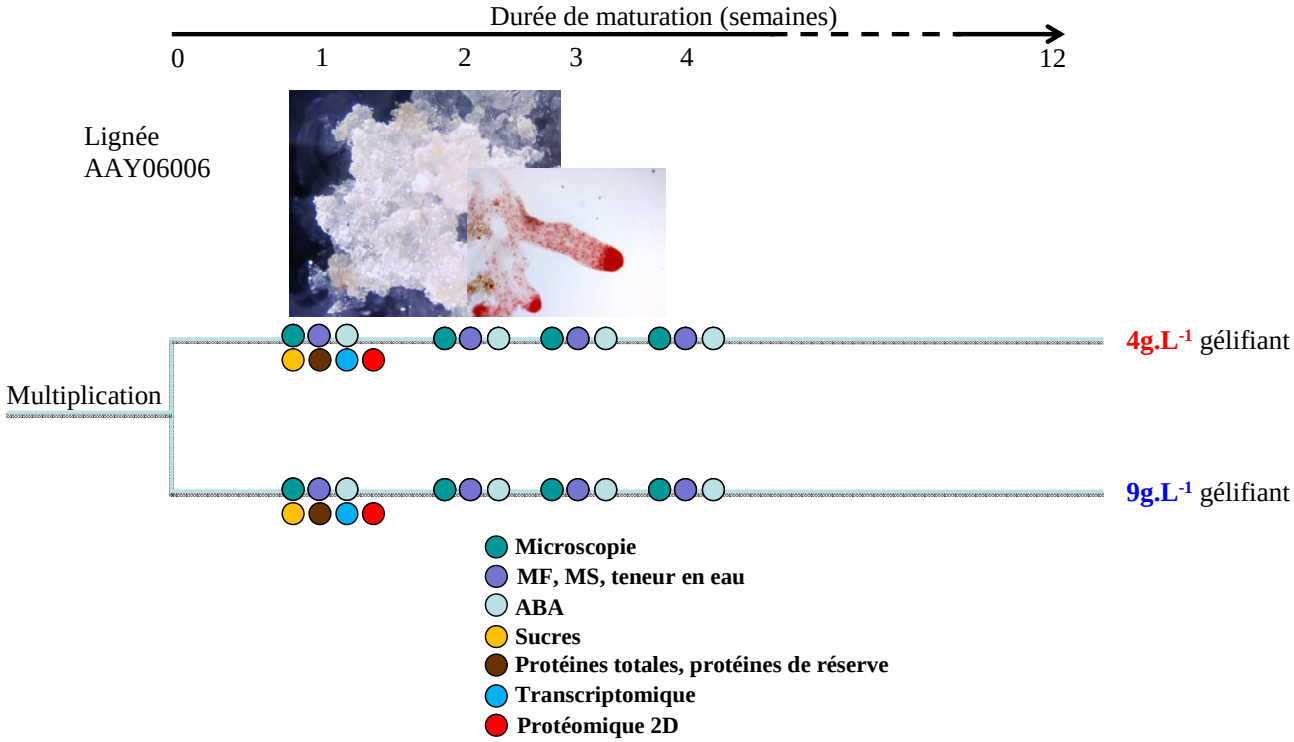


Figure II.1. Plan expérimental, analyses réalisées pour la question de recherche 1 : *Comprendre les évènements moléculaires précoces associés au développement des embryons somatiques de pin maritime en condition de disponibilité en eau réduite.*
MF : masse fraîche ; **MS :** masse sèche, **ABA :** acide abscissique.

II.2. Publication

Sous presse dans *Physiologia Plantarum* : doi: 10.1111/ppl.12158

Résumé

Les embryons somatiques de pin maritimes exigent une réduction de la disponibilité de l'eau (forte concentration en facteur gélifiant dans le milieu de maturation) pour atteindre le stade cotylédonaire. Ce phénomène spécifique aux espèces de pins, n'est pas encore bien compris. Pour faciliter l'utilisation de l'embryogenèse somatique pour la propagation de masse des conifères, nous avons besoin d'une meilleure compréhension du développement de l'embryon. Les comparaisons du transcriptome (séquençage ARN Illumina) et du protéome (2D-SDS-PAGE avec identification en LC/MS-MS) d'embryons somatiques immatures, cultivées soit sur forte (9G) ou faible (4G) concentration en gélifiant, ont été effectuée, ainsi que l'analyse de la teneur en eau, la masse sèche, la teneur en ABA endogène (GC- MS), en sucres solubles (HPLC), en amidon et des observations en microscopie confocale à laser.

L'intégration de ces analyses multi-échelles, a été utilisée pour identifier les événements moléculaires et physiologiques impliqués dans le développement précoce de l'embryon somatique. Dans des conditions défavorables (4G), le métabolisme de la glycolyse est favorisé, certainement en relation avec la prolifération des cellules qui peut être antagoniste du développement de l'embryon somatique. Dans des conditions favorables (9G), les embryons somatiques s'adaptent à la contrainte culturale en activant des voies spécifiques de protection, ainsi que des réponses moléculaires et physiologiques, ABA dépendantes, favorisant le développement de l'embryon. Nos résultats suggèrent que des protéines de type ubiquitine-ligase et germin-like pourraient être utilisées comme marqueurs prédictifs du développement de l'embryon somatique tandis que la protéine phosphatase 2C pourrait être un biomarqueur des réponses adaptatives aux conditions de culture.

C'est la première caractérisation des mécanismes moléculaires précoces impliqués dans le développement d'embryons somatiques de pin suite à une augmentation de la concentration en gélifiant dans le milieu de maturation, et c'est également la première étude sur l'embryogenèse somatique des conifères combinant des données de type transcriptomique et protéomique.

Physiologia Plantarum
An International Journal for Plant Biology

 © 2014 Scandinavian Plant Physiology Society, ISSN 0031-9317

Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water availability: transcriptomic and proteomic analyses

Alexandre Morel[a], Caroline Teyssier[a], Jean-François Trontin[b], Kateřina Eliášová[c], Bedřich Pešek[c], Martine Beaufour[d], Domenico Morabito[e], Nathalie Boizot[a], Claire Le Metté[a], Leila Belal-Bessai[a], Isabelle Reymond[b], Luc Harvengt[b], Martine Cadene[d], Françoise Corbineau[f], Martin Vágner[c], Philippe Label[a,†,‡] and Marie-Anne Lelu-Walter[a,*,‡]

[a]INRA, UR 0588 Unité Amélioration, Génétique et Physiologie Forestières, 2163 Avenue de la Pomme de Pin, CS 4001, Ardon, F-45075 Orléans Cedex 2, France
[b]FCBA, Pôle Biotechnologie et Sylviculture Avancée, Equipe Génétique et Biotechnologie, Campus Forêt-Bois de Pierroton, 71 route d'Arcachon, F-33610 Cestas, France
[c]Institute of Experimental Botany ASCR, Rozvojová 263, Praha 6-Lysolaje 165 02, Czech Republic
[d]CNRS, UPR4301, Centre de Biophysique Moléculaire, Rue Charles Sadron, F-45100 Orléans, France
[e]Université d'Orléans, INRA, LBLGC, EA1207, F-45067 Orléans, France
[f]Sorbonne Universités, UPMC Université Paris 06, UMR 7622, Biologie du Développement, 4 place Jussieu, F-75005 Paris, France

Correspondence
*Corresponding author,
e-mail: marie-anne.lelu-walter@orleans.inra.fr

Received 16 October 2013;
revised 19 December 2013

doi:10.1111/ppl.12158

Maritime pine somatic embryos (SEs) require a reduction in water availability (high gellan gum concentration in the maturation medium) to reach the cotyledonary stage. This key switch, reported specifically for pine species, is not yet well understood. To facilitate the use of somatic embryogenesis for mass propagation of conifers, we need a better understanding of embryo development. Comparison of both transcriptome (Illumina RNA sequencing) and proteome [two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis with mass spectrometry (MS) identification] of immature SEs, cultured on either high (9G) or low (4G) gellan gum concentration, was performed, together with analysis of water content, fresh and dry mass, endogenous abscisic acid (ABA; gas chromatography-MS), soluble sugars (high-pressure liquid chromatography), starch and confocal laser microscope observations. This multiscale, integrated analysis was used to unravel early molecular and physiological events involved in SE development. Under unfavorable conditions (4G), the glycolytic pathway was enhanced, possibly in relation to cell proliferation that may be antagonistic to SE development. Under favorable conditions (9G), SEs adapted to culture

Abbreviations – 2D, two-dimensional; ABA, abscisic acid; ABA-GE, ABA glucose ester; ADH, alcohol dehydrogenase; CI, confidence interval; DW, dry weight; EC, enzyme commission; EMs, embryonal masses; EST, expressed sequence tag; F3H, flavanone 3-hydroxylase; FDA, fluorescein diacetate; FDR, false discovery rate; FW, fresh weight; GC, gas chromatography; GO, gene ontology; HPLC, high-pressure liquid chromatography; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC, liquid chromatography; MS, mass spectrometry; PAGE, polyacrylamide gel electrophoresis; PDC, pyruvate decarboxylase; PI, propidium iodide; PP2C, protein phosphatase 2C; SE, somatic embryo; SUMO, small ubiquitin-related modifier.

[†]Present address: INRA-Université Blaise Pascal, UMR_A 547 PIAF, Les Cézeaux, 24 Avenue des Landais, 63177 Aubière cedex, France
[‡]These authors contributed equally to this work.

constraint by activating specific protective pathways, and ABA-mediated molecular and physiological responses promoting embryo development. Our results suggest that on 9G, germin-like protein and ubiquitin-protein ligase could be used as predictive markers of SE development, whereas protein phosphatase 2C could be a biomarker for culture adaptive responses. This is the first characterization of early molecular mechanisms involved in the development of pine SEs following an increase in gellan gum concentration in the maturation medium, and it is also the first report on somatic embryogenesis in conifers combining transcriptomic and proteomic datasets.

Introduction

Maritime pine, *Pinus pinaster* Ait., is industrially important in Southwestern Europe, especially in France, Portugal and Spain. Maintenance of plantation productivity together with increasing forest resilience to major biotic and abiotic risks are thus major objectives for this species. In order to bypass the current limitations of classical methods for accelerating tree domestication, technological innovation is required, and this may result from any of a number of biotechnological inputs into improvement programs (Harfouche et al. 2012).

Biotechnology offers new opportunities for improving selection efficiency (balancing genetic gain and genetic diversity), conservation and deployment of new varieties for reforestation. As a high-performance, clonal propagation system, somatic embryogenesis is a powerful tool for scaling up production of genetically improved varieties (Park 2002). During the last decade, methodology for somatic embryogenesis in conifers has undergone rapid and significant progress (see review by Lelu-Walter et al. 2013). However, somatic embryo (SE) maturation is a challenging step (Stasolla et al. 2002) toward rapid production of uniform and consistent emblings quality up to the accepted standard by foresters for conventional seedlings in commercial deployment of new varieties. Since the early protocols were developed for spruce species (Attree and Fowke 1993), it was assumed that they would be applicable to pines; however, it subsequently became apparent that pines were less responsive to these approaches. When effective, somatic embryogenesis could also result in qualitative and quantitative differences in gene expression between somatic and reference zygotic embryos (Pérez Rodríguez et al. 2006, Lara-Chavez et al. 2012). Such distinctive features raised the important question of achieving in pine similar SE development to high-quality seedlings.

With regard to zygotic embryogenesis, the genus *Pinus* differs markedly from *Picea*, the main difference in *Pinus* being the type of cleavage that takes place at the late embryogeny stage (Dogra 1978). In pine somatic embryogenesis, recovery of cotyledonary SE occurs only if the availability of water to the embryonal masses (EMs) is reduced (Klimaszewska and Smith 1997). This is achieved by increasing the gel strength of the medium, by using a high gellan gum concentration (Klimaszewska et al. 2000). In maritime pine, SE maturation, which is a protracted process (requiring 12 weeks to reach the cotyledonary stage), is obtained only on high gellan gum concentration, i.e. $9-10\,g\,l^{-1}$ (Ramarosandratana et al. 2001, Lelu-Walter et al. 2006). In contrast, under low gellan gum concentrations, e.g. $4\,g\,l^{-1}$ as used for spruce species, EMs undergo excessive proliferation, which interferes with embryo maturation.

In both angiosperms (Black 1991) and gymnosperms (Kong et al. 1997), abscisic acid (ABA) is involved in embryogenesis. Apart from certain species such as *Pinus taeda* (Kapik et al. 1995), ABA content increases during seed maturation, reaches a maximum and then declines as the seed dehydrates (reviewed by King 1982). In *Picea glauca* the main source of ABA for the developing zygotic embryo is from the surrounding tissue, the megagametophyte (Kong et al. 1997). Interestingly, for conifers exogenous ABA promotes SE maturation through accumulation of storage products such as proteins (Gutmann et al. 1996). ABA can regulate the switch from EM proliferation to embryo development (Stasolla et al. 2002).

Many stress treatments, such as water stress, are known to induce *de novo* synthesis of ABA and therefore to enhance endogenous ABA levels. These increases generally occur during the first days of the treatment. Desiccation of hybrid larch SE was found to result in reduced water content associated with a transient increase in ABA (Dronne et al. 1997). ABA signaling involves complex transduction mechanisms in which there is cross-talk between serine/threonine kinase (SRK2) and protein phosphatase 2C (PP2C) during ontogenesis (Umezawa et al. 2009).

In addition to water availability and ABA, carbohydrate supply is of great importance for pine somatic embryogenesis. In *P. pinaster*, maturation of SE is

enhanced by raising the sucrose concentration to 0.2 *M* (Lelu-Walter et al. 2006). Clearly, endogenous carbohydrate status varies throughout somatic embryogenesis (see review by Lipavská and Konrádová 2004). In general, during the proliferation phase, EMs have high hexose content and almost no sucrose, resulting in a high hexose/sucrose ratio. On the other hand, maturation results in a decrease in hexose content, whereas a gradual increase in sucrose content is observed in the SE (Lipavská et al. 2000). Starch accumulates mainly during the first weeks of the maturation process.

In addition to these physiological studies, efforts have recently been made to describe SE development in conifers at the molecular level, but studies so far remain limited. Proteomic studies, combining two-dimensional (2D) gel electrophoresis with a high degree of accuracy in protein identification by mass spectrometry (MS), are available only for *Cupressus sempervirens* (Sallandrouze et al. 1999), *P. glauca* (Lippert et al. 2005) and *Larix* × *eurolepis* (Teyssier et al. 2014). These studies have contributed to the identification of proteins involved in the embryogenesis process. For gene expression, only a few reports are available, mainly for *Picea abies*. In this species, transcript levels have been compared between different lines (embryogenic vs developmentally arrested) at different SE developmental stages (Stasolla et al. 2004). More recently, changes in expression of genes with products involved in auxin biosynthesis and gibberellin-mediated signaling were reported in a study that made the assumption that conifer genes are homologous to those in *Arabidopsis* (Vestman et al. 2011). In support of this assumption, genes related to embryogenesis in model plants have sequence similarity to expressed sequence tags (ESTs) from *P. taeda* and *Larix leptolepis* (Cairney and Pullman 2007, Zhang et al. 2012).

This study provides new insights into early molecular events regulating the development of *P. pinaster* SE under reduced water availability. Our original approach consists in integrated multiscale analyses, including comparisons of the transcriptome and proteome with histological, biological and biochemical observations. This is the first report combining transcriptomic and proteomic data for conifer somatic embryogenesis. The molecular mechanisms underlying SE response to reduced water availability are discussed in detail. We have found novel candidate predictive markers for conifer SE development as well as for adaptive responses, which will facilitate practical application of the knowledge gained in this study to monitor early responses of embryogenic tissue to maturation conditions.

Materials and methods

Pinus pinaster somatic embryogenesis

Experiments were conducted with an embryogenic line (AAY06006) of *P. pinaster*. EMs were initiated from immature zygotic embryos (Park et al. 2006) originating from a Landes × Morocco polycross. EMs were subcultured and matured in darkness according to Lelu-Walter et al. (2006). Briefly, filter paper discs with dissociated EMs were transferred onto maturation medium that consisted of modified Litvay medium (Litvay et al. 1985) supplemented with 0.2 *M* sucrose, 80 µ*M* *cis-trans* (±) ABA (Sigma, St Louis, MO) and either 4 or 9 g l^{-1} gellan gum (Phytagel™, Sigma), indicated here as 4G and 9G, respectively. After 12 weeks of maturation, cotyledonary SEs were counted and embryogenic potential [number of SE per g fresh weight (FW)] was estimated.

FW, DW and water content of EMs

Measurements were performed after 0, 1, 2, 3 and 4 weeks of maturation on both media. Each week, each filter paper disc with dispersed EMs was weighed and replaced on the maturation medium. FW of EMs was obtained by subtracting the mass of the filter paper disc from the total FW. Dry weight (DW) was determined after oven drying at 70°C for 24 h and calculated by multiplying the DW to FW ratio by 100 (Teyssier et al. 2011). Water content was calculated as (FW − DW)/DW and expressed as g H_2O g^{-1} DW (Dronne et al. 1997). Finally, water availability from the maturation media was determined by measuring the amount of water absorbed by the filter paper disc (Teyssier et al. 2011). Each assay (FW, DW, water content and water availability) was carried out on 10 biological replicates for each medium at each time point.

Histology and microscopy

EMs were collected after 0, 1, 2 and 3 weeks of maturation on both media. EM viability was determined by double staining with fluorescein diacetate (FDA) (Sigma-Aldrich) and propidium iodide (PI) (Sigma-Aldrich, Steinheim, Germany) according to modified protocols (Jones and Senft 1985, Petřek et al. 2005, Vondráková et al. 2010). A LSM5 Duo confocal laser scanning microscope (Zeiss, Jena, Germany) equipped with an Argon/2 laser (FDA excitation at 488 nm, emission filter-set BP 505–550) and a DPSS laser (PI excitation at 561 nm, emission filter-set LP 650) was used to visualize stained EMs. Viable cells exhibit bright green fluorescence, whereas non-viable or dead cells show bright red fluorescence. Starch grains in squashed fresh

EMs were stained with iodine/potassium iodide and observed with a Jenaval transmission light microscope (Zeiss) equipped with a DS-5M digital camera (Nikon, Tokyo, Japan). All images were processed using the computer image analysis system NIS Elements, version 3.2 (Laboratory Imaging, Prague, Czech Republic).

ABA and ABA glucose ester (ABA-GE) measurements

Sample preparation and determination of ABA and ABA-GE

EMs were collected after 0, 1, 2, 3 and 4 weeks of maturation. Samples (each approximately 0.5 g FW) were ground to powder in liquid nitrogen using a pestle and mortar, and each sample was mixed with 3 ml of modified Bieleski solution; a further 2×3 ml of the solution was used to rinse the mortar and the three extracts were combined. ABA and ABA-GE labeled with stable isotopes were added to the extraction mixture as internal standards and the mixture was left overnight to extract. The extract was centrifuged and applied on C18 SPE column (Strata C18-T, Grace, Lokeren, Belgium) to eliminate non-polar compounds. The eluate was partly evaporated by the rotary vacuum evaporator to approximately 3 ml and acidified by adding of 1 ml of 1 *M* formic acid. The eluate was applied on the MCX SPE column (Oasis MCX, Waters, Milford, MA), cleaned by 2 ml of 1 *M* formic acid, then eluted by 5 ml of 100% methanol and evaporated on the rotary vacuum concentrator to dryness. Dried samples were diluted in 100 µl of 15% (v/v acetonitrile:water) solution and filtered on Nylon 0.2-µm Micro-Spin (Grace) filters. The filtrates were divided into two parts: one for determination of ABA and the second for determination of ABA-GE.

ABA determination by GC-MS

ABA was analyzed by gas chromatography (GC)-MS/MS (Vandenbussche et al. 2010). The MS detector (ion trap) was operated in MS/MS mode: MS/MS precursor 190.2, product full scan 65–200 amu for ABA; MS/MS precursor 194.2, product full scan 70–200 amu for labeled ABA. At the first stage, MS full scan 50–300 amu was added to confirm the identity of the substance. A calibration graph for each substance was constructed to cover the range 1–1000 ng ml^{-1} of final extract.

ABA-GE determination by LC-MS

A 5 µl sample of filtrate was injected into a liquid chromatography (LC)-MS system consisting of an HTS PAL autosampler (CTC Analytics, Zwingen, Switzerland), Rheos 2200 gradient high-pressure liquid chromatography (HPLC) pump (Flux, Basel, Switzerland) and Quantum Ultra AM triple-quadrupole mass spectrometer (Thermo Instruments, San Jose, CA) equipped with an electrospray ionization (ESI) interface. Separation was performed on an HPLC column, 50×2.1 mm^2 Kinetex C18 (Phenomenex, Torrance, CA) using gradient elution with a mobile phase consisting of acetonitrile, water and 1% acetic acid. The mass spectrometer was set to record selected reaction monitoring (SRM) transitions $425 \rightarrow 153$ for ABA-GE and, as internal standard, $430 \rightarrow 157$ for D5-ABA-GE. A calibration graph was constructed to cover the concentration range 1–100 ng ml^{-1} of extract.

Determination of carbohydrate and total protein content

Maturation was performed as described above and EMs were collected after 1 week.

Carbohydrate and starch assay

Soluble sugars were extracted as described by Black et al. (1996). EMs (approximately 40 mg FW) were ground in 1 ml of 80% aqueous ethanol containing melezitose as an internal standard and lipids were removed using chloroform (1:1, v/v). An aliquot of the aqueous phase was freeze-dried and the dried extract was dissolved in distilled water, filtered through a 0.45-µm pore filter (Nalgene, Rochester, NY) and then analyzed by HPLC. Samples were eluted with 80:20 (v/v) acetonitrile/H_2O at a flow rate of 1 ml min^{-1} on a 200×4.6 mm^2 Spherisorb-NH2 column (Thermo Fisher Scientific, France). The eluates were analyzed with a differential refractometer (RI Plus Detector, Thermo Fisher Scientific) and the peak areas were electronically integrated using ChromQuest 5.0 (Thermo Fisher Scientific). The various sugars were identified by co-elution with standards (Sigma). For starch, undissolved sample pellets resulting from soluble carbohydrate extraction were dried and incubated in 0.02 *M* NaOH for 2 h at 95°C. After cooling to room temperature, starch was reduced to glucose; amyloglucosidase (35 U) was added and samples were incubated at 50°C for 1 h with stirring. After centrifugation, aliquots of the clear supernatants were assayed for glucose (Bergmeyer et al. 1974).

Total protein assay

Total protein extracts were prepared from five replicates for each type of sample (25–50 mg FW of frozen EMs) as previously described (Teyssier et al. 2014). Protein

concentrations were determined using the Bradford assay with bovine serum albumin as a standard.

Results, which were expressed as $\mu g\,g^{-1}$ DW, corresponded to the mean of 3 (carbohydrates), 5 (total protein) or 6 (starch) replicates ± confidence interval (CI) at the 5% error level.

Transcriptomic analysis

Maturation was carried out as described above. EMs were collected after 1 week of maturation on 4G and 9G, frozen in liquid nitrogen and ground for 5 min with a mortar and pestle to a fine powder. Frozen powder samples were divided into two equal portions (each about 400 mg FW) for transcriptomic and proteomic analyses. For transcriptomic analysis, three biological replicates were sampled for each maturation conditions, producing a total of six samples.

Sequencing

Library construction and sequencing was done by a commercial service provider (GATC Biotech, Germany). Total RNA was extracted, mRNAs purified and cDNA synthesized according to the Smart Kit (Clontech) manufacturer's recommendations. The six samples were paired-end sequenced on an Illumina Genome Analyzer II producing 100-bp paired-end short reads. Primer–adapter removal, quality filtering and ambiguous base trimming produced sequences 82 ± 6 bp long. The six samples each produced 5 586 402 ± 1 514 289 paired-end sequence pairs, giving a total of 33 518 411 paired-end sequence pairs. Original data were deposited in the NCBI's Sequence Read Archive under accession number SRR609713.

Sequences (termed reference sequences) obtained from the Pine3 database (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster3/index.html) were used for functional annotation and gene ontology (GO) analysis and also for mapping the illumina reads. Seventy-six percent of our Illumina reads mapped onto these reference sequences. Reference sequences numbered 212 928, N50 was 691 bp and average length was 818 bp.

Functional characterization and GO annotation

GO annotations to retrieve molecular function, biological process and cellular component terms were performed using BLAST2GO (http://www.blast2go.org/) (Conesa et al. 2005). The sequences were loaded into the BLAST2GO program, and BLASTX with a minimum *E*-value of 1.0 was performed by the program prior to mapping. The mapping step was used to annotate the sequences with GO terms from the GO database. Fisher's exact test was performed within BLAST2GO to determine whether GO identifiers occurred more often in a group than would appear by chance. For GO analysis, biological process GO identifiers were considered only if they occurred more than five times.

Functional classification by KEGG

Annotation against the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/) was performed in order to exploit enzyme data to integrate genomic, chemical and network information (Ogata et al. 1999) using the BLAST2GO program. All mapped sequences were annotated to KEGG database accessions to obtain their enzyme commission (EC) numbers. The EC numbers were then mapped onto KEGG to obtain the corresponding KEGG Pathway Maps.

2D gel proteomic analysis

2D PAGE

Total protein extracts were prepared from five replicates for each type of sample (approximately 400 mg FW) as previously described by Teyssier et al. (2014). Briefly, extraction was realized with phenol and precipitation with ammonium acetate in ethanol and then the proteins were assayed. First, dimension separation was performed with 300 µg protein loaded onto 24-cm IPG strips, pH 4–7 (Protean IEF Cell system, BioRad, France), whereas the 2D polyacrylamide gel electrophoresis (PAGE) was performed with 11% polyacrylamide gels. Gels were stained with Colloidal Brillant Blue G (CBB-G). The numerical analysis of the gels was performed with PROGENESIS software (Nonlinear Dynamics, UK) as described by Teyssier et al. (2014).

Gel spot processing

The destain, wash and proteolysis steps and sample preparation for MS are described elsewhere (Teyssier et al. 2014).

NanoLC-MS/MS analysis

Liquid chromatographic separations of peptides were performed using an UltiMate 3000 nano-ultraLC system (Dionex) after preconcentration on a 15-cm reverse-phase column (Acclaim Pepmap RSLC C18). Online analysis via a distal-coated PicoTip emitter (New Objective) and a nanospray source was performed on an HCT Ultra ion trap (Bruker, Germany), using HYSTAR 3.2

software (Bruker) for integration of liquid chromatography and MS. The spray voltage was set between 1800 and 2200 V through EsquireControl 6.2 (Bruker) and the ion trap was scanned in MS mode from m/z 100 to 1500. Peaks with the four highest intensities were selected for CID MS/MS fragmentation within a 4-Da isolation window, using an exclusion m/z list and dynamic exclusion as previously described (Beaufour et al. 2012). Fragment ions were scanned from m/z 100 to 3000.

Protein identification by database sequence assignment

The procedure was as described elsewhere (Teyssier et al. 2014), except for the database used. MS/MS spectra were searched against the *P. pinaster* contig database (see following paragraph). Only proteins with a minimum of four peptides reaching individual scores greater than 60% were taken into account. Protein lists for each spot were curated manually to remove redundancies. The false discovery rate (FDR) at the peptide level was calculated by combining all data and applying the decoy feature of the identification engine. A FDR value of 0.57% was found for peptide scores equal or greater than 60%.

Statistical analysis

Statistical analysis was carried out with R software (version 2.8.0; R Development Core Team 2011). Effects of the treatments on the FW, DW, water content, sugar analysis, total protein analysis and ABA measurements were evaluated using one-way analysis of variance (ANOVA). Variations in these parameters during maturation in relation to the gellan gum concentration in the medium were analyzed with multiple comparisons of means with Tukey contrasts ($P < 0.05$). For analysis of transcriptomic data, each paired-end sequence was mapped onto the *P. pinaster* transcriptome assembly (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster3/index.html) with Burrows-Wheeler Aligner (BWA) (Li and Durbin 2009). Paired-end matches per transcript were counted, and corrected for transcript length and sample sequencing depth to produce transcript expression levels in Fragments Per Kilobase of transcript per Million fragments mapped (FPKM) (Roberts et al. 2011). Differential expression was assessed by a *t*-test using biological replicates of each treatment. *P*-values were adjusted for multiple hypothesis testing by the Benjamini–Hochberg correction (Benjamini and Hochberg 1995) to control for the FDR. Transcripts with an adjusted *P*-value below 0.05 were considered to be differentially expressed. Finally, for 2D PAGE analysis, the intensity change for each spot was analyzed with Student's *t*-test on the basis of the normalized spot volume ($P < 0.05$).

Results

Histological and biological characteristics of maritime pine EMs

EMs from proliferation medium consisted of single early SEs each being composed of several meristematic cells and a few voluminous suspensor cells with starch grains present (Fig. 1A). After transfer onto 4G maturation medium, embryos started to develop; however, after 2 weeks dead cells were detected in meristematic centers (Fig. 1B). In the third week, dying or dead cells were observed in the majority of suspensors (Fig. 1C). On 9G, meristematic centers gradually enlarged, and suspensors became well organized, comprising a number of long narrow cells after 3 weeks (Fig. 1D). Vital staining showed that dead cells were mainly confined to very distal parts of suspensors. A difference in starch grain distribution was clear after 2 weeks of maturation. On 4G, starch grains spread throughout the suspensors and there were a few in basal parts of meristems (Fig. 1E). On 9G, starch grains accumulated to a noticeably higher extent in suspensors and parts of meristems (Fig. 1F). After 12 weeks, EMs on 4G showed overgrowth all over the filters but no embryos (Fig. 2A), whereas those on 9G yielded cotyledonary SE (Fig. 2B), with the embryogenic potential being estimated at 191 ± 23 (mean $\pm$ CI at the 5% error level).

Compared with EMs matured on 4G, those transferred onto 9G showed no FW increase (Fig. 3A), but there was a significant increase in their DW ($P < 0.05$, Fig. 3B) and decrease in their water content ($P < 0.05$, Fig. 3C). After 4 weeks, EMs matured on 9G had a significantly higher DW content (10.1% FW) and lower water content (9.0 g $H_2O\,g^{-1}$ DW) than those on 4G (5.1% FW and 19.0 g $H_2O\,g^{-1}$ DW; Fig. 3B, C, respectively). Availability of water from the maturation medium was significantly lower on 9G (0.580 g $H_2O \pm 0.009$, mean $\pm$ CI at the 5% error level) than on 4G (0.635 g $H_2O \pm 0.008$).

ABA content

At the beginning of the maturation period, the ABA content of EMs was low (2.1 nmol g^{-1} DW). EMs subsequently showed a significant increase in endogenous ABA (Fig. 4), whereas ABA-GE conjugate remained undetectable whatever the conditions tested (data not shown). On 9G, ABA content increased slowly during the first 3 weeks of maturation but peaked at 21.1 nmol g^{-1} DW after 4 weeks. A similar initial increase in ABA content was observed in EMs matured on 4G, but after

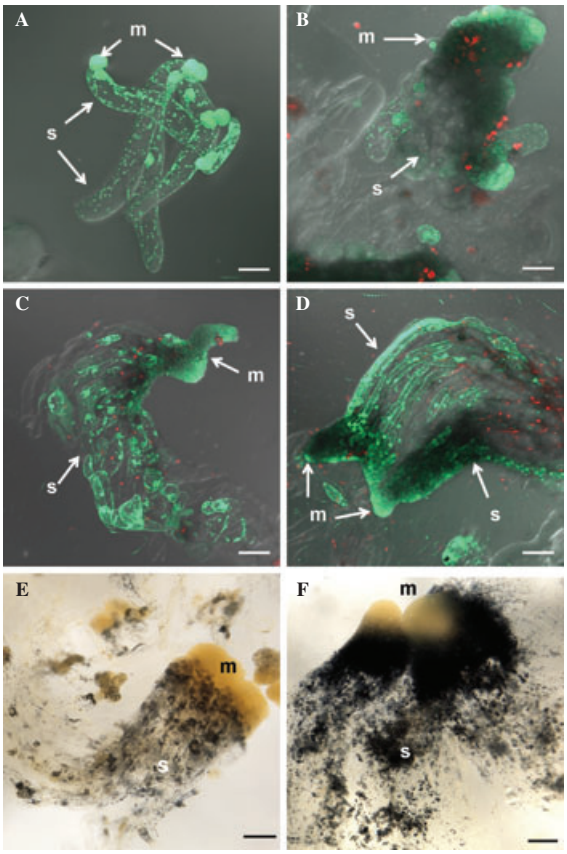


Fig. 1. Viability and starch accumulation in immature SEs of maritime pine matured on either 4 or 9 g l^{-1} gellan gum. Confocal images (projections of optical sections) of embryos stained with FDA and PI show green fluorescein fluorescence in viable cells and red fluorescence of PI–DNA complexes in dead cells (A–D). Light transmission microscope images show staining of starch grains with Lugol solution (E, F). (A) Very early embryos consisting of a few meristematic cells and one or two suspensor cells before transfer onto maturation media. (B) Embryo matured for 2 weeks on 4 g l^{-1} gellan gum with dead cells apparent in the meristem. (C) Embryo matured for 3 weeks on 4 g l^{-1} gellan gum with the disorganized suspensor composed of short viable cells; dead cells were present throughout the suspensor structure. (D) Embryos matured for 3 weeks on 9 g l^{-1} gellan gum with well-organized suspensors consisting of long narrow viable cells; dead cells were present in the distal parts of suspensors. (E) Embryos matured for 2 weeks on 4 g l^{-1} gellan gum accumulated starch grains around nuclei and in the cytoplasm of suspensor cells. (F) Embryos matured for 2 weeks on 9 g l^{-1} gellan gum accumulated starch grains in suspensor cells and in the basal part of the embryo proper; a much greater extent of starch accumulation was evident. Scale bar = 100 µm (A, B) and 200 µm (C–F). m, meristematic cells; s, suspensor cells.

4 weeks of maturation, the ABA content was less than half (9.0 nmol g^{-1} DW) that on 9G.

Carbohydrate and total protein contents

At 1 week of maturation, similar amounts of total protein were observed for EMs developed on 4G or 9G

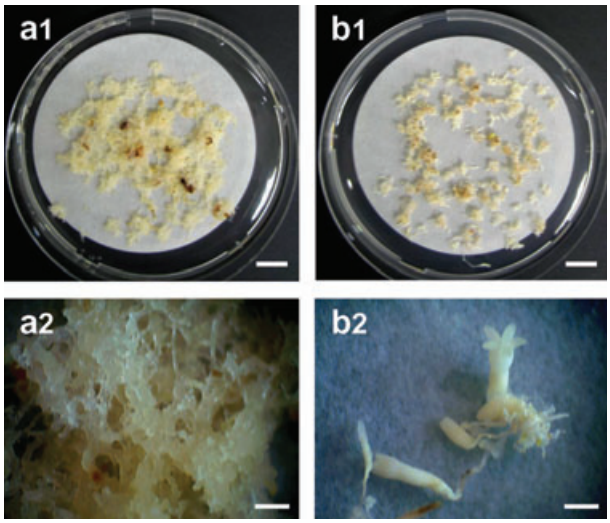


Fig. 2. Morphological appearance of EMs of maritime pine matured for 12 weeks on either 4 g l^{-1} (A) or 9 g l^{-1} (B) gellan gum. Note EM proliferation on 4G (a1, a2) and presence of cotyledonary SEs on 9G (b1, b2). Scale bar = 1 cm (a1, b1) and 1 mm (a2, b2).

(Table 1). In contrast, significant differences in carbohydrate contents were observed. EMs matured on 4G had a greater amount of fructose (Fru, 27.4 µg g^{-1} DW) and glucose (Glu, 26.6 µg g^{-1} DW) but a similar sucrose content (Su, 6.88 µg g^{-1} DW) to that of EMs matured on 9G. In consequence, the ratio (Fru + Glu)/Su was significantly higher for EMs matured on 4G than for those matured on 9G (7.9 and 4.4, respectively). Other carbohydrates such as stachyose, raffinose, maltose and mannose were not detected whatever the conditions. Finally, EMs matured on 9G contained a significantly larger amount of starch (46.30 µg g^{-1} DW) than those matured on 4G (32.70 µg g^{-1} DW).

Transcriptomic analysis

A total of 1103 transcripts were considered to be differentially expressed, 632 for EMs matured on 9G and 471 for EMs matured on 4G (Table S1, Supporting Information).

Data annotation by enrichment analysis and functional classification

Using the BLAST2GO program, significant transcripts were annotated against the Plants/*Arabidopsis thaliana* section of the protein sequences database. Next, the annotated significant transcripts were mapped against the GO database to retrieve the relevant GO terms. In total, 4865 GO terms were assigned to the 529 mapped transcripts with an average of five GO terms assigned to

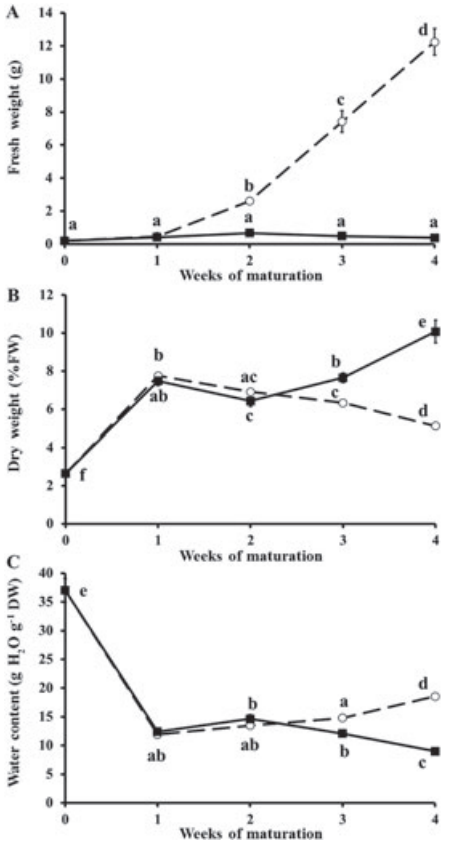


Fig. 3. Biological characteristics of immature SEs of maritime pine during maturation in the presence of either low (4 g l^{-1}, ○) or high (9 g l^{-1}, ■) gellan gum. (A) FW, (B) DW and (C) water content. Bars represent CI at the 5% error level. Letters represent statistical groups defined by the multiple comparisons of means method ($P < 0.05$, N = 10).

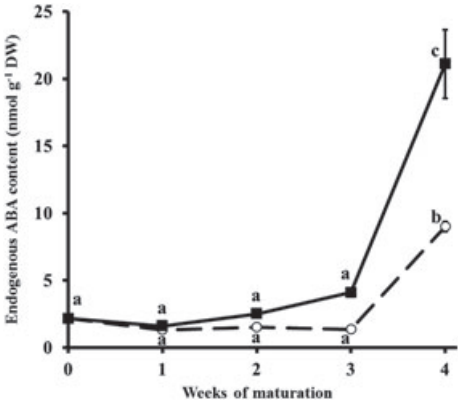


Fig. 4. Endogenous ABA content of immature SEs of maritime pine during maturation in the presence of either low (4 g l^{-1}, ○) or high (9 g l^{-1}, ■) gellan gum. Bars represent CI at the 5% error level. Letters represent statistical groups defined by the multiple comparisons of means method ($P < 0.05$, N = 10).

Table 1. Carbohydrate and protein content in maritime pine immature SEs matured for 1 week in the presence of either 4 or 9 g l^{-1} gellan gum. [a]Values are means of 3 (carbohydrates), 5 (total protein) or 6 (starch) measurements ± CI at the 5% error level. Within table, different letters represent statistically significant differences.

	Gellan gum concentration (g l^{-1})	
Compound (μg g^{-1} DW)[a]	4	9
Carbohydrate		
Fructose (Fru)	27.44 ± 2.22[a]	17.27 ± 1.39[b]
Glucose (Glu)	26.60 ± 2.86[a]	19.29 ± 0.71[b]
Sucrose (Su)	6.88 ± 0.14[a]	8.29 ± 1.24[a]
Total soluble carbohydrates	60.91 ± 5.17[a]	44.85 ± 2.65[b]
Ratio ([Fru + Glu]/Su)	7.9[a]	4.4[b]
Starch	32.70 ± 0.5[a]	46.30 ± 0.92[b]
Total protein	654.40 ± 81.72[a]	733.10 ± 100.4[a]

each transcript. Of these GO terms, the majority were assigned to the biological process category (3118, 64%), followed by molecular function (1217, 25%) and cellular component (530, 11%).

GO analysis at level 2, illustrating the biological process categories, is presented in the 'Hit description' column of Table S1. Within the biological process ontology, proteins involved in cellular process, metabolic process and response to stimulus predominated.

Figure 5 shows the results of GO enrichment analysis, illustrating the distribution of biological process GO terms that were differentially enriched for EMs matured on 4G vs 9G. Twenty-one functional categories were represented by 298 sequences originating from EMs matured on 4G and 239 sequences from 9G. The majority of the significant categories for EMs matured on 4G were involved in carbohydrate metabolism, energy metabolism and cellular respiration, whereas for EMs matured on 9G the significant categories were cell growth and embryo-related development, regulation response to stimulus and secondary metabolic processes.

In order to identify the biological pathways active in early EM maturation, the assembled ESTs were used to obtain the corresponding EC numbers by annotation against the KEGG database. A total of 207 transcripts were assigned to 40 ECs. The ECs were then grouped into 56 biochemical pathways, of which 33 pathways were found to be involved in metabolism. Most enzymes active in EMs matured on 4G participate in the glycolysis/gluconeogenesis pathway (23) and methane metabolism pathway (16).

Transcript expression levels

Among the 471 transcripts overexpressed in EMs grown on 4G, the most prominent were nine encoding pyruvate decarboxylase with a mean of 12.0 FPKM, seven en-

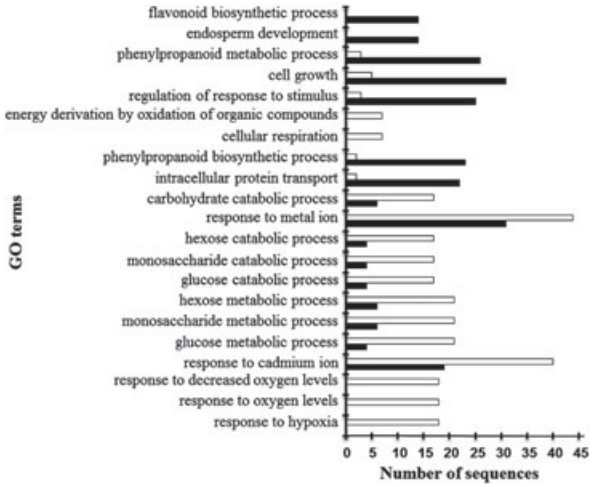


Fig. 5. Gene ontology (biological process categories at level 2) terms found to be significantly enriched after 1 week of maturation with either low (4 g l^{-1}, □) or high (9 g l^{-1}, ■) gellan gum. Only GO categories for which more than five genes were represented are included. To test whether GO terms were enriched in a test group, when compared with a reference group, Fisher's exact test with multiple testing correction of FDR (FDR < 0.05, Benjamini and Hochberg 1995) was used. Numbers of sequences: numbers of sequences involved in each biological process.

coding alcohol dehydrogenase (8.3 FPKM), seven annotated as bifunctional enolase 2 and related to a glucose catabolic process (15.2 FPKM), one annotated as 6-phosphofructokinase 2 involved in glycolysis (7.5 FPKM) and one annotated as sucrose synthase 3 responsive to oxygen levels (9.8 FPKM).

For EMs grown on 9G, the most interesting transcripts among the 632 that were overexpressed are 20 annotated as either SUMO-conjugating enzyme or ubiquitin-conjugating enzyme (25.6 FPKM) with an ontology related to endosperm development, 5 encoding naringenin,2-oxoglutarate 3-dioxygenase, also known as flavanone 3-hydroxylase (F3H) and associated with a flavonoid biosynthetic process (8.1 FPKM), 3 encoding chitinase (18.6 FPKM), 2 transcripts for PP2C 16, which is putatively involved in negative regulation of the ABA-mediated signaling pathway (28.3 FPKM), and 3 serine/threonine-protein kinases transcripts reported as being involved in the positive regulation of ABA-mediated signaling pathway, SRK2E (55.2 FPKM) and SRK2H (18.3 FPKM). In addition, one transcript annotated as pollen Ole e 1 allergen and extensin family protein, but with a weak similarity of 47%, was the most highly overexpressed transcript, at 230.4 FPKM.

2D gel proteomic analysis

Changes in EM protein profiles were investigated using 2D gel electrophoresis. Separation with high resolution was achieved, allowing detection of a total of 1428 spots (representative gel shown in Fig. S2). The gellan gum concentration had no significant effect (*t*-test, $P < 0.05$) on the normalized volumes of most spots. Significant differences were detected for only 83 spots, 35 with greater abundance on 4G and 48 on 9G. Of the significant spots, 58 were picked for analysis by MS. Protein identification was successful for 56 spots, 27 of which were overproduced on 4G and 29 on 9G (Table 2). Detailed information on the statistical parameters used for protein identification is given in Table S2. All identified proteins were grouped into functional categories within the biological process ontology (level 2) using BLAST2GO (Fig. S3). A striking result of this classification was that around 50% of all spots identified were placed in only three categories: metabolic process, cellular process and response to stimulus.

Discussion

We determined key molecular, physiological and biological parameters in order to develop a comprehensive picture of SE maturation in maritime pine. In addition to SE growth and water content, we measured endogenous ABA, total protein, soluble carbohydrate and starch contents. We also tentatively identified as many protein and transcript levels as possible in order to gain information about the cellular environment as a whole. For this study, we used a culture medium parameter specifically appropriate for pine species, namely gellan gum concentration, to trigger the maturation process. Transcriptomic and proteomic analyses were performed at week 1 after initiation of cultures, whereas other parameters were monitored weekly. This sampling strategy was designed to focus mainly on early molecular mechanisms (as reflected in transcript and protein profiles) that could be involved in subsequent physiological modifications promoting SE development.

As the *P. pinaster* genome has yet not been completely sequenced, we are aware that it is not currently possible to infer numbers of genes per gene family. Hence, we cannot estimate whether variant transcripts detected in this study represent alternative transcripts from individual members of gene families or are products of paralogs. However, we can state with confidence that each short-read pair that we obtained mapped specifically to one reference transcript, and the results indicate some diversity in the range of transcripts deriving from different gene families; further characterization of these families at the genome level is still necessary. We decided to limit data analysis to the genes that were most differentially expressed

Table 2. Characteristics of differentially expressed proteins in immature SEs of maritime pine matured for 1 week on either 4 or 9 g l^{-1} gellan gum. Excised spot number refers to the spot marked in Fig. S1. The protein assignments and accession numbers were retrieved from the GenoToul database (http://*genotoul*-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster2/index.html) by the PEAKS identification engine. The most homologous protein in the UniProt database (http://www.uniprot.org/) and its species affiliation are given. Gellan gum: protein relative abundance (9 > 4: protein overexpressed on 9G, 4 > 9: protein overexpressed on 4G). th., theoretical; exp., expected.

Spot number	Accession number	Gellan gum	Protein name	UniProt accession	Organism	MW th./ exp.	pHi th./ exp.
5	SC_RP12A_ARATH.1.6	4 > 9	26S proteasome regulatory particle subunit	Q0GEA1	*Camellia sinensis*	31/26.3	4.95/4.96
56	SC_STAD_RICCO.1.4	4 > 9	Acyl-[acyl-carrier-protein] desaturase	A9NVD4	*Picea sitchensis*	46.9/44.3	5.65/5.5
36	SC_00001024.1.1	4 > 9	Alanine aminotransferase	Q84UX4	*Oryza sativa*	53.6/52.3	8.01/5.9
30	SC_ADH3_SOLTU.1.15	4 > 9	Alcohol dehydrogenase	Q43025	*Pinus banksiana*	40.5/44.6	6.12/6.02
34	SC_ADH_FRAAN.16.30	4 > 9	Alcohol dehydrogenase	Q43025	*Pinus banksiana*	40.5/46.4	6.12/5.86
32	SC_00002407.1.1	4 > 9	At5g54100 (similarity to stomatin-like protein)	Q9LVW0	*Arabidopsis thaliana*	43.6/47.2	9.11/5.96
59	SC_00001003.1.1	4 > 9	BAHD family acyltransferase, clade V	D8T9Y4	*Selaginella moellendorffii*	49.4/55.4	5.91/5.5
10	SC_RRAA3_ARATH.2.3	4 > 9	Cell division protein FtsQ	A3TJS8	*Oryza sativa*	18/9.8	5.6/5.7
55	SC_CH62_CUCMA.1.9	4 > 9	Chaperonin CPN60-2, mito.	Q05046	*Cucurbita maxima*	57.6/52.7	5.35/5.67
62	SC_CH62_CUCMA.2.9	4 > 9	Chaperonin CPN60-2, mito.	Q05046	*Cucurbita maxima*	57.6/61.5	5.35/5.47
60	SC_CH62_CUCMA.2.9	4 > 9	Chaperonin CPN60-2, mito.	Q05046	*Cucurbita maxima*	57.6/60.9	5.35/5.52
38	SC_00001092.1.1	4 > 9	DEAD-box ATP-dependent RNA helicase 2	Q5VNM3	*Oryza sativa*	45.6/48.7	5.98/5.96
78	SC_MCM31_MAIZE.1.5	4 > 9	DNA replication licensing factor MCM3 homolog 3	Q9SX03	*Zea mays*	85.2/86	5.92/5.91
33	SC_EFTM_ARATH.1.6	4 > 9	Elongation factor Tu. mito.	Q9ZT91	*Arabidopsis thaliana*	44.1/45.1	5.93/5.9
45	SC_ENO1_HEVBR.7.31	4 > 9	Enolase 1	Q9LEJ0	*Picea sitchensis*	48.1/52.5	5.79/5.75
58	SC_GLGS_ARATH.1.1	4 > 9	Glucose-1-phosphate adenylyltransferase	B8LNV7	*Picea sitchensis*	57.9/51.2	5.99/5.46
37	SC_MDARS_CUCSA.1.10	4 > 9	Monodehydroascorbate reductase	A5JPK7	*Vitis vinifera*	47.3/47.1	5.93/5.94
29	SC_LOC100274026_MAIZE.1.3	4 > 9	NAD-dependent isocitrate dehydrogenase c	Q6ZI55	*Oryza sativa*	40.6/42.3	8.14/6.02
28	SC_PGKH1_ARATH.1.6	4 > 9	Phosphoglycerate kinase	F4MKM1	*Pinus pinaster*	52.9/44.6	8.83/6.11
31	SC_PGKH1_ARATH.1.6	4 > 9	Phosphoglycerate kinase	F4MKM1	*Pinus pinaster*	52.9/45	8.83/5.99
75	SC_PREP1_ARATH.1.3	4 > 9	Presequence protease 1. chloro./mito.	Q9LJL3	*Arabidopsis thaliana*	111.5/92	5.03/5.35
50	SC_PDC2_ORYSI.1.19	4 > 9	Pyruvate decarboxylase isozyme 2	A2XFI3	*Oryza sativa*	64.8/63.7	5.6/5.74
51	SC_PDC2_ORYSI.1.19	4 > 9	Pyruvate decarboxylase isozyme 2	A2XFI3	*Oryza sativa*	64.8/63.2	5.6/5.75
52	SC_PDC2_ORYSI.3.19	4 > 9	Pyruvate decarboxylase isozyme 2	Q10MW3	*Oryza sativa*	64.7/63	5.53/5.72
53	SC_PDC2_ORYSI.3.19	4 > 9	Pyruvate decarboxylase isozyme 2	Q10MW3	*Oryza sativa*	64.7/63.3	5.53/5.71
44	SC_METK1_PICSI.34.55	4 > 9	*S*-adenosylmethionine synthase 1	A9NUH8	*Picea sitchensis*	42.9/47.4	5.42/5.7
46	SC_00154823.1.1	4 > 9	UDP-glucose 6-dehydrogenase	Q6RK07	*Cinnamomum osmophloeum*	52.9/55.8	5.99/5.72
76	SC_DPE2_ARATH.1.1	9 > 4	4-Alpha-glucanotransferase DPE2 (amylomaltase)	Q69Q02	*Oryza sativa*	108.3/90	6/5.61
43	SC_00000649.1.1	9 > 4	6-Phosphogluconate dehydrogenase. decarboxylating	B9RVA7	*Ricinus communis*	54.2/55.9	6.09/5.94
40	SC_00000578.1.1	9 > 4	Acetyl-CoA carboxylase (biotin carboxylase)	O23960	*Glycine max*	52.3/57.6	5.82/6.11

Table 2. (*Continued*)

Spot number	Accession number	Gellan gum	Protein name	UniProt accession	Organism	MW th./ exp.	pHi th./ exp.
42	SC_SAHH_NICSY.1.26	9 > 4	Adenosylhomocysteinase	B8LLL7	*Picea sitchensis*	53.2/55.8	5.75/5.96
35	SC_ADH1_PETHY.12.16	9 > 4	Alcohol dehydrogenase	Q43300	*Pinus banksiana*	40.4/47.3	5.76/5.82
6	SC_00003988.1.1	9 > 4	AT3g23600/MDB19_9	Q9LUG8	*Arabidopsis thaliana*	25.9/23.4	5.17/5.12
22	SC_00002431.1.1	9 > 4	Benzoquinone reductase	A3F7Q3	*Picea sitchensis*	21.7/21.4	5.94/6.17
70	SC_CALR_RICCO.2.19	9 > 4	Calreticulin	Q9FYV2	*Pinus taeda*	49.6/50.1	4.6/4.84
69	SC_CALR_RICCO.2.19	9 > 4	Calreticulin	Q9FYV2	*Pinus taeda*	49.6/0.4	4.6/4.86
49	SC_CH62_CUCMA.1.9	9 > 4	Chaperonin CPN60-2. mito.	Q05046	*Cucurbita maxima*	57.6/60.6	5.35/5.75
17	SC_00004509.1.1	9 > 4	Expansin S2	Q39626	*Cucumis sativus*	25.4/21.2	9.39/6.02
64	SC_00000119.1.1	9 > 4	FAM10 family protein At4g22670	Q93YR3	*Arabidopsis thaliana*	46.6/57.5	4.83/5.31
4	SC_GLGL1_BETVU.1.2	9 > 4	Glucose-1-P adenylyltransferase large subunit	P55233	*Beta vulgaris*	57.7/32.9	6.05/4.76
24	SC_00007308.1.1	9 > 4	GTP-binding nuclear protein Ran-A1	P41918	*Nicotiana tabacum*	25/23.8	6.38/6.12
47	SC_ILV5_SPIOL.1.21	9 > 4	Ketol-acid reductoisomerase. chloro.	O82043	*Pisum sativum*	57.8/61.7	5.85/5.84
9	SC_UMC2374_MAIZE.1.3	9 > 4	Probable 6-phosphogluconolactonase 4. chloro.	Q69NG5	*Oryza sativa*	28.3/23.8	5.65/5.34
1	SC_PROF1_PHAVU.6.6	9 > 4	Profilin	A9NMR7	*Picea sitchensis*	14.3/6.1	4.61/4.81
26	SC_00151055.1.1	9 > 4	Proteasome subunit beta type 1	P42742	*Arabidopsis thaliana*	40.3/42.3	9.46/6.19
27	SC_00151055.1.1	9 > 4	Proteasome subunit beta type 1	P42742	*Arabidopsis thaliana*	40.3/42.8	9.46/6.14
71	SC_PDI14_ARATH.1.13	9 > 4	Protein disulfide isomerase	A7KY78	*Oldenlandia affinis*	54.6/70.9	5.1/4.98
72	SC_PDI14_ARATH.1.13	9 > 4	Protein disulfide isomerase	A7KY78	*Oldenlandia affinis*	54.6/73.5	5.1/4.98
57	SC_PDI23_ORYSJ.1.5	9 > 4	Protein disulfide isomerase family	A9CPA7	*Glycine max*	47.7/53.9	5.41/5.58
81	SC_XYL1_ARATH.1.6	9 > 4	Putative alpha-xylosidase	Q8VWV9	*Pinus pinaster*	100.6/91	6.1/6.09
16	SC_EXB14_ORYSJ.1.2	9 > 4	Putative expansin-B14	Q6H677	*Oryza sativa*	25.5/27.2	5.3/5.97
23	SC_00146588.1.1	9 > 4	Putative germin-like protein 9-1	Q652Q1	*Vitis vinifera*	23.5/22.4	6.4/6.21
15	SC_SODF_NICPL.4.4	9 > 4	Superoxide dismutase	Q6QJL0	*Pinus pinaster*	28.6/21.1	7.18/5.89
21	SC_SODM_NICPL.1.5	9 > 4	Superoxide dismutase	A9NNA1	*Picea sitchensis*	25.5/20.4	8.7/6.08
66	SC_TBB2_ELEIN.1.4	9 > 4	Tubulin beta-2 chain	Q9ZPN9	*Eleusine indica*	50.4/54.8	4.71/5.15
8	SC_00009054.1.1	9 > 4	UMP–CMP kinase	O04905	*Arabidopsis thaliana*	22.5/25.4	5.79/5.19

according to FPKM ratio statistics. These genes were grouped on the basis of GO annotation at the biological process level. From this data analysis, it appears that EMs expressed markedly different set of genes according to the maturation medium used.

Does gellan gum concentration control maturation of SEs?

When compared to the lower concentration, 4G, it appears that on 9G, maritime pine SE differentiated better, based on macroscopical and histological observations (Figs 1 and 2), and on measurements of DW and water content (Fig. 3). Our results reinforce findings previously reported for somatic embryogenesis in white pine (Klimaszewska et al. 2000) and Scots pine (Lelu-Walter et al. 2008), where high gellan gum concentrations were linked to lower water availability and enhanced SE development. Low water and carbon source availability also had a major impact on SE maturation and germination in carrot (Lee et al. 2001), suggesting that the induction of embryo development by temporary growth starvation observed in natural conditions can be mimicked *in vitro* in a broad range of species. Our results indicate that gellan gum concentration and associated modulation of water availability in the medium are critical in controlling SE development in maritime pine.

What happens in EMs when SE maturation is hindered?

Microscopic observations (Fig. 1) showed that EMs grown on 4G produced less structured immature embryos and more undifferentiated tissues through cell proliferation (Fig. 2), an observation corroborated by the increase in FW due mainly to water uptake throughout the culture period (Fig. 3). In general, for pine species maturation in the presence of low concentrations of gelling agent sustains proliferation of immature SE but is detrimental to SE development (Klimaszewska et al. 2000).

In 4G-grown EMs, endogenous synthesis of ABA is limited and there is no detectable ABA conjugation, at least via ABA-GE synthesis. We avoided exogenous application of ABA at levels high enough to obscure endogenous ABA synthesis by several orders of magnitude because of the occurrence of high levels of ABA conjugation (Daie et al. 1984) as reported for larch (Label and Lelu 2000).

Carbohydrate contents in EMs cultured on 4G are characterized by a high hexose:sucrose ratio and fairly low starch accumulation, which is in accordance with microscopic observations (Fig. 1E). These results suggest stimulation of the glycolytic pathway, possibly related to cell proliferation (Fig. 2A). A high hexose:sucrose ratio has been observed in proliferating conifer EMs (Lipavská and Konrádová 2004).

Transcriptionally, our results show an increased level of expression of alcohol dehydrogenase (ADH) and pyruvate decarboxylase (PDC) genes in 4G-grown EMs, along with overexpression of sucrose synthase. At the proteomic level, ADH, phosphoglycerate kinase and PDC isozyme 2 are also significantly upregulated. The two datasets are consistent and support possible stimulation of the glycolytic pathway, as also suggested by the results of carbohydrate analysis. However, a recent article on *Arabidopsis* reported that ADH and PDC enzymes and their corresponding transcripts could be modulated to overcome unfavorable conditions such as anoxia brought about by submergence in culture medium (Mithran et al. 2014). In addition, utilization of sucrose, provided in our study by the culture medium, is important in anoxia tolerance mediated by alcoholic fermentation (Loreti et al. 2005). These findings, taken together with our results, suggest that for EMs, 4G is (1) a condition unfavorable for SE differentiation and maturation, possibly owing to an excess of water at this low gellan gum concentration and (2) a suboptimal condition for proliferation (indicated by the expression of glycolysis pathway genes) that could result in the rapid onset (within 1 week) of alcoholic fermentation (expression of anoxia-related genes) and cell senescence.

Additionally, in EMs grown on 4G we observed overexpression of several bifunctional enolase-2 transcripts that may be related to abiotic stress-regulated transcriptional networks (Chinnusamy et al. 2007). Bifunctional enolase 2, also known as low expression of osmotically responsive genes 2 (LOS2), appears to be a negative regulator of transcription factors involved in abiotic stress responses. This result again suggests that 4G may be a stressful culture condition for both cell proliferation and (to an even greater extent) differentiation.

Our proteomic results revealed a high level of protein synthesis after 1 week of maturation on 4G. This is probably a prerequisite for cell division. We identified many transcripts (cell division control protein 48-e, growth-regulating factor 5 and mitogen-activated protein kinase 6) and proteins (DNA replication licensing factor MCM3 homolog 3 and cell division protein FtsQ) involved in both DNA replication and cell division on 4G. Because of the increase in FW and DW and the histological results (Fig. 1), we propose that these genes and their products may be involved in cell proliferation. One of the most significantly overexpressed proteins on 4G is related to a 26S proteasome regulatory subunit that is involved in selective breakdown of proteins (Vierstra 2003) but also in many other aspects of cell regulation including embryogenesis and cell senescence. This is an interesting point, because we observed necrotic cells in meristematic centers as early as 2–3 weeks after the start of maturation (Fig. 1). High 26S proteasome subunit activity suggests both active regulation of cell proliferation and senescence on 4G. No related transcripts could be detected in the same samples, but transcript abundance is not always well correlated with levels of the final products (Dembinsky et al. 2007, Pan et al. 2012).

Which components appear during maturation of SEs?

ABA is synthesized at a low level during the first 3 weeks of culture (Fig. 4), but peaks at week 4 when SEs are differentiating, while the water content remains relatively high. On 9G the EMs are not much dehydrated, even though water availability is relatively limited compared to that on 4G. The major increase in ABA at week 4 suggests that there is an ontogenetic signal for SE development at this stage, as has been observed in larch (Label and Lelu 2000). For instance, many ubiquitin-protein ligase transcripts are detected in 9G. More recent findings indicate that this enzyme is associated with the activation of the small ubiquitin-related modifier (SUMO), leading to specific post-transcriptional modifications of chromatin and thereby influencing gene

transcription (Cubenas-Potts and Matunis 2013). Our results suggest that in differentiated EMs, gene expression at week 1 may be undergoing large-scale reorganization to promote SE development. In addition to the ubiquitin-protein ligase transcripts, two proteasome subunit beta type 1 proteins were identified among the proteins over-expressed on 9G. The ubiquitin/proteasome pathway is reported to be a major factor in controlled proteolysis, in which ubiquitinized proteins are recognized and catabolized by the 26S proteasome complex. This pathway supplies amino acids for the biosynthesis of new proteins (Vierstra 2003). A recent article (Rode et al. 2012) links high levels of enzymes involved in the proteasome-dependent proteolysis pathway and the switch from callus to globular embryos in cyclamen.

During somatic embryogenesis in gymnosperms, chitinases are reportedly excreted into the culture medium. These enzymes are involved in the regulation of the EM-to-SE transition and massive programmed cell death (Wiweger et al. 2003, Vestman et al. 2011). Our results showed overexpression of transcripts for chitinase in EMs grown on 9G. In Norway spruce, programmed cell death is responsible for the degradation of EMs when early SEs differentiate (Vestman et al. 2011). The timing of programmed cell death can also be modulated by disulfide isomerase, which was upregulated on 9G. This enzyme seems to be required for proper embryo development in *Arabidopsis* seeds (Andème-Ondzighi et al. 2008). Expansin S2 and putative expansin-B14 are known to play roles in embryogenesis. In *Arabidopsis*, they are involved in the cell cycle as key regulators of wall and cell expansion (Thibaud-Nissen et al. 2003, Spencer et al. 2007). The proteins tubulin beta-2 chain and GTP-binding nuclear protein Ran-A1 participate in cytoskeletal processes such as nucleocytoplasmic transport (Vernoud et al. 2003, Pan et al. 2009). All these proteins, which are upregulated in 9G, may direct cell division toward cell remodeling for embryogenesis, a hypothesis consistent with our histological observations.

Several serine/threonine kinase (SRK2) transcripts are highly overexpressed in EMs growing on 9G. SRK2s are plant homologs of AMPK/SNF1 genes, which in other eukaryotes are considered to be crucial in linking stress and metabolic responses (Kulik et al. 2011). They are reported to be key players in osmotic stress responses and ABA-dependent plant development (Fujii et al. 2011). Endogenous ABA synthesis was significantly enhanced on 9G, i.e. when water availability was reduced. In addition, the detection of highly overexpressed PP2C transcripts in 9G-grown EMs reinforces the likelihood of such a – possibly adaptive – molecular response (Sheen 1998). PP2Cs are reported to be key protein–protein interactors mediating ABA responses in plants (Yoshida et al. 2006). Our results suggest that immature SEs respond to reduced availability of water in the maturation medium through ABA-mediated molecular mechanism.

Two superoxide dismutases are overexpressed at the proteomic level, as is one germin-like protein. The latter is also detected at the transcriptomic level, suggesting an ongoing process of antioxidant protein production through mRNA synthesis in EMs growing on 9G. The role of germin-like proteins in plant development and defense is well documented (Bernier and Berna 2001). Their involvement as markers of early stages of somatic embryogenesis in both conifers (Neutelings et al. 1998, Mathieu et al. 2006) and angiosperms (Caliskan et al. 2004, Kim and Triplett 2004) makes them especially valuable in this context. Superoxide dismutase is involved in the regulation of oxidative stress. It destroys free radicals that are normally produced within the cells and are toxic to biological systems. In a proteomic study on the early stages of seed development in *Pinus massoniana*, a correlation was established between the level of this protein and the transition from the proembryo to the columnar embryo stage (Zhen et al. 2012). It has been suggested that the ability of embryos to develop an efficient redox homeostasis system may be key for successful regulation and maintenance of cell differentiation.

Our results also show that many transcripts related to F3H, as well as several other transcripts related to flavonol metabolism, are overexpressed in 9G-grown EMs. F3H catalyzes hydroxylation of flavanones to dihydroflavonols. Flavonoids are ubiquitous in plants, and some of these compounds play specific roles in seed development (Lepiniec et al. 2006). Among these roles, stress resistance is a key biological function that leads to the synthesis of several secondary compounds all related to flavanones hydroxylation and subsequent metabolism down to condensed tannins (Winkel-Shirley 2002). Our results suggest that EMs cultivated on 9G may overcome constraints imposed by culture conditions by synthesizing protective secondary metabolites. This does not seem to occur in EMs grown on 4G, where many stress response-related transcripts were identified as discussed above.

At the proteomic level, two enzymes (glucose-1-phosphate adenylyltransferase AGPP; 4-alpha-glucano-transferase DPE2, disproportionating enzyme) reportedly involved in starch biosynthesis (Keeling and Myers 2010) were found to be overexpressed on 9G. This is consistent with the high starch content of immature SE cultivated in these favorable conditions (Fig. 1, Table 1). It is interesting to note that glucose and fructose contents were correlated with the difference in starch grain

accumulation on 4G and 9G (Fig. 1E, F, Table 1). In maritime pine, a low abundance of starch granules was reported in proliferating EMs (Breton et al. 2005), whereas starch was found to accumulate throughout developing embryos during early maturation (Tereso et al. 2007). A similar picture was reported during larch SE maturation (Gutmann et al. 1996).

Conclusions

This is the first study on the early molecular mechanisms involved in somatic embryogenesis of pine following an increase in gellan gum concentration in the medium, and it is also the first report combining transcriptomic and proteomic data on somatic embryogenesis in conifers. Our transcriptomic and proteomic results indicate that glycolysis is enhanced under conditions unfavorable for SE maturation (4G), which lead to EM proliferation. Under favorable conditions (9G), we observed adaptive, ABA-mediated molecular and physiological responses to reduced water availability resulting in early transition of EMs from proliferation to the SE developmental pathway (indicated by active protein synthesis, and overexpression of proteins involved in cell division, embryogenesis and starch synthesis). Specific pathways (synthesis of protective secondary metabolites and regulation of oxidative stress) are concomitantly activated, apparently to overcome constraints due to culture conditions.

On the basis of both our transcriptomic and proteomic data and related information available in the literature, we proposed the germin-like protein and ubiquitin-protein ligase as potential marker genes and proteins for early evidence of effective SE development on maturation medium. PP2C may also be a suitable marker for monitoring adaptive molecular responses to favorable maturing conditions that may have a beneficial impact on regeneration capacity of cultures. The identification of relevant biomarkers at the molecular level needs validation experiments but has clear practical implications for implementation of somatic embryogenesis in both conifer fundamental research and commercial applications. Functional genomics based on reverse genetics in maritime pine is a long process requiring transgenic plant regeneration through somatic embryogenesis and cryopreservation of selected transgenic lines (Trontin et al. 2002). Labor intensiveness would be significantly decreased if the regeneration capacity of transgenic lines could be predicted from early molecular monitoring of rapid maturation tests (1 week). Furthermore, untargeted (genome-wide) or targeted transcriptomic and proteomic profiling (candidate genes and protein markers) may be invaluable for qualitative and quantitative selection of elite embryogenic lines that are receptive to maturing conditions and productive enough for breeding application (clonal and varietal tests) and new variety deployment in the frame of multivarietal forestry.

Author contributions

A. M. participated in the acquisition of the data, carried out all types of analyses and drafted the manuscript. C. T. conceived the design of the study, participated in 2D gel proteomic analysis and drafted the manuscript. J.-F. T. conceived the design of the study, participated in somatic embryogenesis and drafted the manuscript. K. E. carried out histological and microscopical analysis and drafted the manuscript. B. P. carried out ABA and ABA-GE analyses. M. B. carried out gel spot processing and protein identification by database sequence assignment. D. M. participated in the design of the study and helped to draft the manuscript. N. B. carried out carbohydrate analyses and helped to draft the manuscript. C. L. M. carried out somatic embryogenesis and collected the material. L. B.-B. participated in 2D gel proteomic analyses. I. R. carried out somatic embryogenesis and collected the material. L. H. participated in the design of the study and helped to draft the manuscript. M. C. participated in gel spot processing and protein identification by database sequence assignment and helped to draft the manuscript. F. C. participated in study design, participated in carbohydrate analysis and drafted the manuscript. M. V. participated in study design and drafted the manuscript. P. L. conceived the design of the study, carried out transcriptomic analysis, performed the statistical analysis and drafted the manuscript. M.-A. L.-W. conceived the design of the study and its coordination, and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements – This work was supported by grants from the French 'Conseil Régional de la Région Centre' (EMBRYOME project, contract 33639), the 'Conseil Régional de la Région Aquitaine' (Embryo2011, contract 09012579-045) and the French Ministry of Foreign Affairs and the French Ministry of Higher Education and Research through the France/Czech Republic Science Cooperation BARRANDE Programme. Data analysis was made possible through the involvement of INRA's GenoToul bioinformatics platform in Toulouse (France).

References

Andème-Ondzighi C, Christopher DA, Cho EJ, Chang S-C, Staehelin LA (2008) Arabidopsis protein disulfide

isomerase-5 inhibits cysteine proteases during trafficking to vacuoles before programmed cell death of the endothelium in developing seeds. Plant Cell 20: 2205–2220
Attree SM, Fowke LC (1993) Embryogeny of gymnosperms: advances in synthetic seed technology of conifers. Plant Cell Tissue Organ Cult 35: 1–35
Beaufour M, Godin F, Vallée B, Cadene M, Bénédetti H (2012) Interaction proteomics suggests a new role for the Tfs1 protein in yeast. J Proteome Res 11: 3211–3218
Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57: 289–300
Bergmeyer HU, Bernt E, Schmidt F, Stork H (1974) D-glucose. In: Bergmeyer HU (ed) Methods of Enzymatic Analysis, Vol. III. Weinheim, Verlag Chemie, pp 1196–1201
Bernier F, Berna A (2001) Germins and germin-like proteins: plant do-all proteins. But what do they do exactly? Plant Physiol Biochem 39: 545–554
Black M (1991) Involvement of ABA in the physiology of developing and mature seeds. In: Davies WJ, Jones HG (eds) Abscisic Acid: Physiology and Biochemistry. BIOS Scientific Publisher, Oxford, pp 99–124
Black M, Corbineau F, Grzesik M, Guy P, Côme D (1996) Carbohydrate metabolism in the developing and maturing wheat embryo in relation to its desiccation tolerance. J Exp Bot 47: 161–169
Breton D, Harvengt L, Trontin J-F, Bouvet A, Favre J-M (2005) High subculture frequency, maltose-based and hormone-free medium sustained early development of somatic embryos in maritime pine. In Vitro Cell Dev Biol Plant 41: 494–504
Cairney J, Pullman GS (2007) The cellular and molecular biology of conifer embryogenesis. New Phytol 176: 511–536
Caliskan M, Turet M, Cuming A (2004) Formation of wheat (*Triticum aestivum* L.) embryogenic callus involves peroxide-generating germin-like oxalate oxidase. Planta 219: 132–140
Chinnusamy V, Zhu J, Zhu J-K (2007) Cold stress regulation of gene expression in plants. Trends Plant Sci 12: 444–451
Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676
Cubenas-Potts C, Matunis MJ (2013) SUMO: a multifaceted modifier of chromatin structure and function. Dev Cell 24: 1–12
Daie J, Wyse R, Hein M, Brenner ML (1984) Abscisic acid metabolism by source and sink tissues of sugar beet. Plant Physiol 74: 810–814
Dembinsky D, Woll K, SaleemM LY, Fu Y, Borsuk LA, Lamkemeyer T, Fladerer C, Madlung J, Barbazuk B, Nordheim A, Nettleton D, Schnable PS, Hochholdinger F (2007) Transcriptomic and proteomic analyses of pericycle cells of the maize primary root. Plant Physiol 145: 575–588
Dogra PD (1978) Morphology, development and nomenclature of conifer embryo. Phytomorphology 28: 307–322
Dronne S, Label P, Lelu M-A (1997) Desiccation decreases abscisic acid content in hybrid larch (*Larix* × *leptoeuropaea*) somatic embryos. Physiol Plant 99: 433–438
Fujii H, Verslues PE, Zhu J-K (2011) Arabidopsis decuple mutant reveals the importance of SnRK2 kinases in osmotic stress responses *in vivo*. Proc Natl Acad Sci USA 108: 1717–1722
Gutmann M, Von Aderkas P, Label P, Lelu M-A (1996) Effects of abscisic acid on somatic embryo maturation of hybrid larch. J Exp Bot 47: 1905–1917
Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, Scarascia Mugnozza G (2012) Accelerating the domestication of forest trees in a changing world. Trends Plant Sci 17: 64–72
Jones KH, Senft JA (1985) An improved method to determine cell viability by simultaneous staining with fluorescein diacetate-propidium iodide. J Histochem Cytochem 33: 77–79
Kapik RH, Dinus RJ, Dean JFD (1995) Abscisic acid and zygotic embryogenesis in *Pinus taeda*. Tree Physiol 15: 485–490
Keeling PL, Myers AM (2010) Biochemistry and genetics of starch synthesis. Annu Rev Food Sci Technol 1: 271–303
Kim H, Triplett B (2004) Cotton fiber germin-like protein. I. Molecular cloning and gene expression. Planta 218: 516–524
King RW (1982) Abscisic acid in seed development. In: Khan AA (ed) Physiology and Biochemistry of Seed Development, Dormancy and Germination. Elsevier Biomedical Press, Amsterdam, pp 157–181
Klimaszewska K, Smith D (1997) Maturation of somatic embryos of *Pinus strobus* is promoted by a high concentration of gellan gum. Physiol Plant 100: 949–957
Klimaszewska K, Bernier-Cardou M, Cyr DR, Sutton BCS (2000) Influence of gelling agents on culture medium gel strength, water availability, tissue water potential, and maturation response in embryogenic cultures of *Pinus strobus* L. In Vitro Cell Dev Biol Plant 36: 279–286
Kong L, Attree SM, Fowke LC (1997) Changes of endogenous hormone levels in developing seeds, zygotic embryos and megagametophytes in *Picea glauca*. Physiol Plant 101: 23–30
Kulik A, Wawer I, Krzywinska E, Bucholc M, Dobrowolska G (2011) SnRK2 protein kinases – key regulators of plant response to abiotic stresses. OMICS 15: 859–872

Label P, Lelu MA (2000) Exogenous abscisic acid fate during maturation of hybrid larch (*Larix* × *leptoeuropaea*) somatic embryos. Physiol Plant 109: 456–462
Lara-Chavez A, Egertsdotter U, Flinn BS (2012) Comparison of gene expression markers during zygotic and somatic embryogenesis in pine. In Vitro Cell Dev Biol Plant 48: 341–354
Lee EK, Cho DY, Soh WY (2001) Enhanced production and germination of somatic embryos by temporary starvation in tissue cultures of *Daucus carota*. Plant Cell Rep 20: 408–415
Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2006) Simplified and improved somatic embryogenesis for clonal propagation of *Pinus pinaster* Ait. Plant Cell Rep 25: 767–776
Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2008) Clonal plant production from self- and cross-pollinated seed families of *Pinus sylvestris* (L.) through somatic embryogenesis. Plant Cell Tissue Organ Cult 92: 31–45
Lelu-Walter M-A, Thompson D, Harvengt L, Sanchez L, Toribio M, Pâques LE (2013) Somatic embryogenesis in forestry with a focus on Europe: state-of-the-art, benefits, challenges and future direction. Tree Genet Genomes 9: 883–899
Lepiniec L, Debeaujon I, Routaboul J-M, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol 57: 405–430
Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760
Lipavská H, Konrádová H (2004) Somatic embryogenesis in conifers: the role of carbohydrate metabolism. In Vitro Cell Dev Biol Plant 40: 23–30
Lipavská H, Svobodová H, Albrechtová J, Kumstýřová L, Vágner M, Vondráková Z (2000) Somatic embryogenesis in Norway spruce: carbohydrate status during embryo maturation and the effect of polyethylene glycol treatment. In Vitro Cell Dev Biol Plant 36: 260–267
Lippert D, Jun Z, Ralph S, Ellis DE, Gilbert M, Olafson R, Ritland K, Ellis B, Douglas CJ, Bohlmann J (2005) Proteome analysis of early somatic embryogenesis in *Picea glauca*. Proteomics 5: 461–473
Litvay JD, Verma DC, Johnson MA (1985) Influence of loblolly pine (*Pinus taeda* L.) culture medium and its components on growth and somatic embryogenesis of the wild carrot (*Daucus carota* L.). Plant Cell Rep 4: 325–328
Loreti E, Poggi A, Novi G, Alpi A, Perata P (2005) A genome-wide analysis of the effects of sucrose on gene expression in *Arabidopsis* seedlings under anoxia. Plant Physiol 137: 1130–1138
Mathieu M, Lelu-Walter M-A, Blervacq AS, David H, Hawkins S, Neutelings G (2006) Germin-like genes are expressed during somatic embryogenesis and early development of conifers. Plant Mol Biol 61: 615–627
Mithran M, Paparelli E, Novi G, Perata P, Loreti E (2014) Analysis of the role of the pyruvate decarboxylase gene family in *Arabidopsis thaliana* under low-oxygen conditions. Plant Biol 16: 28–34
Neutelings G, Domon JM, Membre N, Bernier F, Meyer Y, David A, David H (1998) Characterization of a germin-like protein gene expressed in somatic and zygotic embryos of pine (*Pinus caribaea* Morelet). Plant Mol Biol 38: 1179–1190
Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 27: 29–34
Pan Z, Guan R, Zhu S, Deng X (2009) Proteomic analysis of somatic embryogenesis in Valencia sweet orange (*Citrus sinensis* Osbeck). Plant Cell Rep 28: 281–289
Pan Z, Zeng Y, An J, Ye JXQ, Deng X (2012) An integrative analysis of transcriptome and proteome provides new insights into carotenoid biosynthesis and regulation in sweet orange fruits. J Proteomics 75: 2670–2684
Park YS (2002) Implementation of conifer somatic embryogenesis in clonal forestry: technical requirements and deployment considerations. Ann For Sci 59: 651–656
Park YS, Lelu-Walter M-A, Harvengt L, Trontin JF, McEacheron I, Klimaszewska K, Bonga JM (2006) Initiation of somatic embryogenesis in *Pinus banksiana*, *P. strobus*, *P. pinaster*, and *P. sylvestris* at three laboratories in Canada and France. Plant Cell Tissue Organ Cult 86: 87–101
Pérez Rodríguez MJ, Suárez MF, Heredia R, Ávila C, Breton D, Trontin J-F, Filonova L, Bozhkov P, von Arnold S, Harvengt L, Cánovas FM (2006) Expression patterns of two glutamine synthetase genes in zygotic and somatic pine embryos support specific roles in nitrogen metabolism during embryogenesis. New Phytol 169: 35–44
Petřek J, Víteček J, Vlašínová H, Kizek R, Kramer KJ, Adam V, Klejdus B, Havel L (2005) Application of computer imaging, stripping voltammetry and mass spectrometry to study the effect of lead (Pb-EDTA) on the growth and viability of early somatic embryos of Norway spruce (*Picea abies* L. Karst.). Anal Bioanal Chem 383: 576–586
R Development Core Team (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna Available at http://www.R-project.org/
Ramarosandratana A, Harvengt L, Bouvet A, Calvayrac R, Pâques M (2001) Effects of carbohydrate source, polyethylene glycol and gellan gum concentration on embryonal-suspensor mass (ESM) proliferation and

maturation of maritime pine somatic embryos. In Vitro Cell Dev Biol Plant 37: 29–34
Roberts A, Trapnell C, Donaghey J, Rinn J, Pachter L (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol 12: R22
Rode C, Lindhorst K, Braun H-P, Winkelmann T (2012) From callus to embryo: a proteomic view on the development and maturation of somatic embryos in *Cyclamen persicum*. Planta 235: 995–1011
Sallandrouze A, Faurobert M, El Maataoui M, Espagnac H (1999) Two-dimensional electrophoretic analysis of proteins associated with somatic embryogenesis development in *Cupressus sempervirens* L. Electrophoresis 20: 1109–1119
Sheen J (1998) Mutational analysis of protein phosphatase 2C involved in abscisic acid signal transduction in higher plants. Proc Natl Acad Sci USA 95: 975–980
Spencer MWB, Casson SA, Lindsey K (2007) Transcriptional profiling of the Arabidopsis embryo. Plant Physiol 143: 924–940
Stasolla C, Kong L, Yeung EC, Thorpe T (2002) Maturation of somatic embryos in conifers: morphogenesis, physiology, biochemistry, and molecular biology. In Vitro Cell Dev Biol Plant 38: 93–105
Stasolla C, Bozhkov PV, Chu T-M, Van Zyl L, Egertsdotter U, Suarez MF, Craig D, Wolfinger RD, von Arnold S, Sederoff RR (2004) Variation in transcript abundance during somatic embryogenesis in gymnosperms. Tree Physiol 24: 1073–1085
Tereso S, Zoglauer K, Milhinhos A, Miguel C, Oliveira M (2007) Zygotic and somatic embryo morphogenesis in *Pinus pinaster*: comparative histological and histochemical study. Tree Physiol 27: 661–669
Teyssier C, Grondin C, Bonhomme L, Lomenech A-M, Vallance M, Morabito D, Label P, Lelu-Walter M-A (2011) Increased gelling agent concentration promotes somatic embryo maturation in hybrid larch (*Larix* × *eurolepis*): a 2-DE proteomic analysis. Physiol Plant 141: 152–165
Teyssier C, Maury S, Beaufour M, Grondin C, Delaunay A, Le Metté C, Ader K, Cadene M, Label P, Lelu-Walter M-A (2014) In search of markers for somatic embryo maturation in hybrid larch (*Larix* × *eurolepis*): global DNA methylation and proteomic analyses. Physiol Plant 150: 271–291
Thibaud-Nissen F, Shealy RT, Khanna A, Vodkin LO (2003) Clustering of microarray data reveals transcript patterns associated with somatic embryogenesis in soybean. Plant Physiol 132: 118–136
Trontin J-F, Harvengt L, Garin E, Vernaza ML, Arancio L, Hoebeke J, Canlet F, Paques M (2002) Towards genetic engineering of maritime pine (*Pinus pinaster* Ait.). Ann For Sci 59: 687–697
Umezawa T, Sugiyama N, Mizoguchi M, Hayashi S, Myouga F, Yamaguchi-Shinozaki K, Ishihama Y, Hirayama T, Shinozaki K (2009) Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in Arabidopsis. Proc Natl Acad Sci USA 106: 17588–17593
Vandenbussche F, Petrasek J, Zadniková P, Hoyerová K, Pesek B, Raz V, Swarup R, Bennett M, Zazimalova E, Benkova E, Van Der Straeten D (2010) The auxin influx carriers AUX1 and LAX3 are involved in auxin-ethylene interactions during apical hook development in *Arabidopsis thaliana* seedlings. Development 137: 597–606
Vernoud V, Horton AC, Yang Z, Nielsen E (2003) Analysis of the small GTPase gene superfamily of Arabidopsis. Plant Physiol 131: 1191–1208
Vestman D, Larsson E, Uddenberg D, Cairney J, Clapham D, Sundberg E, von Arnold S (2011) Important processes during differentiation and early development of somatic embryos of Norway spruce as revealed by changes in global gene expression. Tree Genet Genomes 7: 347–362
Vierstra RD (2003) The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. Trends Plant Sci 8: 135–142
Vondráková Z, Cvikrová M, Eliášová K, Martincová O, Vágner M (2010) Cryotolerance in Norway spruce and its association with growth rates, anatomical features and polyamines of embryogenic cultures. Tree Physiol 30: 1335–1348
Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol 5: 218–223
Wiweger M, Farbos I, Ingouff M, Lagercrantz U, von Arnold S (2003) Expression of *Chia4-Pa* chitinase genes during somatic and zygotic embryo development in Norway spruce (*Picea abies*): similarities and differences between gymnosperm and angiosperm class IV chitinases. J Exp Bot 54: 2691–2699
Yoshida T, Nishimura N, Kitahata N, Kuromori T, Ito T, Asami T, Shinozaki K, Hirayama T (2006) ABA-hypersensitive germination3 encodes a protein phosphatase 2C (AtPP2CA) that strongly regulates abscisic acid signaling during germination among Arabidopsis protein phosphatase 2Cs. Plant Physiol 140: 115–126
Zhang Y, Zhang S, Han S, Li X, Qi L (2012) Transcriptome profiling and in silico analysis of somatic embryos in Japanese larch (*Larix leptolepis*). Plant Cell Rep 31: 1637–1657
Zhen Y, Zhao Z-Z, Zheng R-H, Shi J (2012) Proteomic analysis of early seed development in *Pinus massoniana* L. Plant Physiol Biochem 54: 97–104

Supporting Information

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

Fig. S1. Functional annotation of assembled sequences (significant transcripts) based on gene ontology (GO) categorization.

Fig. S2. Representative 2D PAGE map for proteins from immature somatic embryos of maritime pine matured for 1 week on high gellan gum (9 g l^{-1}).

Fig. S3. Assignment of the putative proteins identified in immature somatic embryos of maritime pine matured for 1 week to functional categories, according to Blast2GO (http://www.blast2go.com/b2ghome).

Table S1. Identification of differentially expressed transcripts in immature somatic embryos of maritime pine matured for 1 week on either 4 or 9 g l^{-1} gellan gum.

Table S2. Identification of differentially expressed proteins in immature somatic embryos of maritime pine matured for 1 week on either 4 or 9 g l^{-1} gellan gum.

Edited by S. Pelaz

Appendix S1: Identification of differentially expressed transcripts in immature somatic embryos of maritime pine matured for 1 week on either 4 or 9 g L^{-1} gellan gum. The assignments (hit description) and accession numbers were retrieved from the GenoToul database (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster2/index.html).

eValue *: E-value of the best BLAST hit ;* ***Sim mean*** *: Mean similarity value for the BLAST results. This value is computed as the average length cutoff-similarity value for all the hits produced with a given sequence ;* ***Enzyme*** *: Enzyme Commission number (EC) associated with the sequence ;* ***Gellan gum*** *: relative abundance (9>4 : transcript over-expressed in EMs grown on high gellan gum, 4>9 : transcript over-expressed in EMs grown on low gellan gum) ;* ***MeanFpkm9 and MeanFpkm4****: Mean of Fragments Per Kilobase of transcript per Million mapped reads in EMs grown on either high (9) or low (4) gellan gum (value rounded up to the nearest integer);* ***PvalueLogFpkm*** *: P-values adjusted for multiple hypothesis testing by the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995) to control for the false discovery rate ;* ***ratio 9/4*** *= (meanFpkm9 / meanFpkm4) ;* ***ratio 4/9*** *= (meanFpkm4 / meanFpkm9).*

Accession number	Hit description	eValue	Sim mean	Enzyme	Gellan gum	Mean Fpkm 9	Mean Fpkm 4	Pvalue LogFpkm	ratio 9/4	ratio 4/9
SC_OPR3_ARATH.2.6	12-oxophytodienoate reductase 3	1,90E-21	71.8%	EC:1.3.1.42	9>4	7	3	0.00576	2.40	0.42
SC_00061885.1.1	12-oxophytodienoate reductase 3	8,60E-25	71.2%		9>4	8	3	0.02108	2.45	0.41
SC_OPR3_SOLLC.2.3	12-oxophytodienoate reductase 3	6,10E-35	58.8%		9>4	8	2	0.02528	3.52	0.28
SC_00147094.1.1	12-oxophytodienoate reductase 3	1,00E-02	61.0%		9>4	10	3	0.04267	2.83	0.35
SC_1433_LILLO.6.9	14-3-3-like protein GF14 nu	5,10E-21	65.25%		9>4	6	2	0.02791	2.89	0.35
SC_00140036.1.1	2,3-bisphosphoglycerate-independent phosphoglycerate mutase 1	1,60E-39	86.5%		4>9	2	4	0.03174	0.49	2.03
SC_00079041.1.1	3beta-hydroxysteroid-dehydrogenase/decarboxylase isoform 1	2,20E-79	60.64%	EC:1.1.1.170	4>9	2	7	0.00743	0.28	3.56
SC_HIBC8_ARATH.1.13	3-hydroxyisobutyryl-CoA hydrolase-like protein 5	1,20E-123	56.22%		9>4	4	2	0.01007	2.35	0.43
SC_00022766.1.1	3-ketoacyl-CoA synthase 11	1,80E-49	80.7%		9>4	8	2	0.03466	4.23	0.24
SC_OS07G0565000_ORYSJ.1.3	40S ribosomal protein S11-2	4,40E-45	78.0%		9>4	5	2	0.04176	2.50	0.40
SC_R15A1_ARATH.13.17	40S ribosomal protein S15a-1	3,30E-51	67.0%		9>4	8	3	0.01480	3.07	0.33
SC_RS17_SOLLC.7.31	40S ribosomal protein S17-3	1,90E-03	76.0%		4>9	1	2	0.00472	0.41	2.46
SC_RS212_ARATH.4.27	40S ribosomal protein S21-1	1,00E-42	88.0%		9>4	6	2	0.00214	2.34	0.43
SC_RS29_ARATH.11.47	40S ribosomal protein S29	4,60E-28	91.0%		4>9	2	4	0.00172	0.41	2.43
SC_RS29_ARATH.4.47	40S ribosomal protein S29	9,40E-18	63.5%		9>4	5	2	0.01843	2.41	0.42
SC_RSSA_VITVI.11.16	40S ribosomal protein Sa-2	1,60E-63	80.0%		4>9	3	9	0.02261	0.36	2.79
SC_RK12_SPIOL.1.6	50S ribosomal protein L12-1	3,10E-21	77.63%		9>4	3	1	0.03829	2.17	0.46
SC_00015460.1.1	5'-AMP-activated protein kinase-like protein	7,50E-21	52.0%		4>9	3	6	0.01312	0.46	2.15
SC_METE_SOLSC.4.72	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	4,50E-89	94.33%	EC:2.1.1.13; EC:1.3.1.0; EC:2.1.1.14	9>4	14	5	0.00190	2.87	0.35
SC_METE_SOLSC.2.72	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	7,70E-57	93.0%	EC:2.1.1.13; EC:2.1.1.14	9>4	10	3	0.04450	2.94	0.34
SC_METE_SOLSC.54.72	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	0,00E+00	93.0%	EC:2.1.1.13; EC:2.1.1.14	9>4	9	4	0.04594	2.13	0.47
SC_RL10_PINTA.30.65	60S ribosomal protein L10-1	8,00E-21	59.75%		9>4	5	1	0.04983	3.73	0.27
SC_RL141_ARATH.1.10	60S ribosomal protein L14-1	1,20E-63	89.0%		9>4	6	3	0.04289	2.08	0.48

SC_RL142_ARATH.5.13	60S ribosomal protein L14-2	9,70E-08	69.0%		9>4	5	2	0.02286	2.05	0.49
SC_RL15B_PICMA.19.45	60S ribosomal protein L15-2	5,40E-10	86.0%		9>4	58	23	0.03045	2.48	0.40
SC_LOC100303995_MAIZE.4.5	60S ribosomal protein L23	9,30E-49	62.8%		4>9	1	3	0.01573	0.46	2.16
SC_RL23_PONAB.1.1	60S ribosomal protein L23	2,70E-17	83.0%		4>9	3	7	0.03123	0.46	2.19
SC_R27A2_ARATH.6.8	60S ribosomal protein L27a-3	3,60E-26	65.33%		9>4	5	2	0.01600	2.11	0.47
SC_R27A2_ARATH.2.8	60S ribosomal protein L27a-3	2,60E-08	86.0%		9>4	7	3	0.04047	2.31	0.43
SC_PCO139614B_MAIZE.1.1	60S ribosomal protein L3-2	4,10E-37	89.67%	EC:3.1.3.16	4>9	1	3	0.00359	0.42	2.38
SC_00154561.1.1	60S ribosomal protein L3-2	1,20E-50	77.67%	EC:3.1.3.16	4>9	1	3	0.00557	0.42	2.36
SC_RL3_ORYSJ.19.43	60S ribosomal protein L3-2	5,60E-66	84.0%	EC:3.1.3.16	4>9	2	5	0.02161	0.49	2.04
SC_00066504.1.1	60S ribosomal protein L3-2	5,30E-45	91.33%	EC:3.1.3.16	4>9	2	5	0.04018	0.43	2.31
SC_RL341_ARATH.4.9	60S ribosomal protein L34-1	2,90E-31	79.5%		9>4	6	3	0.00433	2.16	0.46
SC_RL5_SOLME.3.3	60S ribosomal protein L5-1	5,70E-07	67.0%		4>9	2	5	0.02183	0.46	2.15
SC_00002753.1.1	60S ribosomal protein L7-2	9,90E-54	87.17%		9>4	6	3	0.02711	2.22	0.45
SC_RL81_ARATH.10.17	60S ribosomal protein L8-3	2,30E-46	84.75%		4>9	2	4	0.04988	0.40	2.47
SC_K6PF2_ARATH.4.14	6-phosphofructokinase 2	1,20E-60	66.33%	EC:2.7.1.11	4>9	2	8	0.00910	0.25	3.99
SC_00079527.1.1	ABC transporter B family member 18	3,50E-40	80.15%	EC:3.6.3.44	9>4	15	6	0.00788	2.37	0.42
SC_PDR7_ORYSJ.3.17	ABC transporter G family member 35	5,40E-25	64.3%		9>4	6	2	0.01378	2.64	0.38
SC_AB35G_ARATH.2.3	ABC transporter G family member 35	4,40E-50	63.15%		4>9	3	6	0.03171	0.47	2.11
SC_STAR1_ORYSJ.1.1	ABC transporter I family member 17	1,20E-106	55.25%		9>4	4	2	0.04010	2.23	0.45
SC_ABAH1_ARATH.2.2	abscisic acid 8'-hydroxylase 1	0,00E+00	58.5%	EC:1.14.13.93	4>9	6	13	0.02110	0.45	2.20
SC_00008626.1.1	ACT domain repeat 1	1,90E-02	45.0%		9>4	5	2	0.02897	2.50	0.40
SC_00014901.1.1	ACT domain-containing protein	5,40E-09	48.86%		9>4	28	11	0.01370	2.67	0.37
SC_OS07G0647500_ORYSJ.15.24	Acytochrome-c oxidase/ electron carrier	2,30E-25	83.0%		9>4	127	59	0.01410	2.14	0.47
SC_00086806.1.1	ADP-ribosylation factor A1B	1,40E-44	71.89%		4>9	1	2	0.04103	0.47	2.13
SC_00028033.1.1	ADP-ribosylation factor A1E	8,40E-24	69.57%		9>4	8	3	0.02855	2.52	0.40
SC_MAD50_ORYSJ.2.15	agamous-like MADS-box protein AGL14	7,10E-27	89.45%		9>4	66	31	0.03141	2.10	0.48
SC_ADH_MALDO.13.21	alcohol dehydrogenase class-P	2,00E-112	62.65%	EC:1.1.1.284; EC:1.1.1.1	4>9	1	7	0.00568	0.18	5.58
SC_ADH3_SOLTU.6.15	alcohol dehydrogenase class-P	7,60E-53	58.54%	EC:1.1.1.1	4>9	4	11	0.00885	0.40	2.52
SC_ADH_FRAAN.21.30	alcohol dehydrogenase class-P	0,00E+00	61.2%	EC:1.1.1.284; EC:1.1.1.1	4>9	5	12	0.02285	0.39	2.55
SC_ADH_FRAAN.8.30	alcohol dehydrogenase class-P	1,80E-113	64.77%	EC:1.1.1.1	4>9	1	4	0.02537	0.34	2.96
SC_ADH3_SOLTU.10.15	alcohol dehydrogenase class-P	2,00E-51	60.08%	EC:1.1.1.1	4>9	5	12	0.03315	0.40	2.51
SC_ADH_FRAAN.13.30	alcohol dehydrogenase class-P	3,40E-36	63.91%	EC:1.1.1.1	4>9	4	10	0.04164	0.39	2.59
SC_ADH_FRAAN.23.30	alcohol dehydrogenase class-P	2,80E-135	65.92%	EC:1.1.1.1	4>9	1	2	0.04316	0.49	2.04
SC_00037875.1.1	aldehyde dehydrogenase 2B4	6,00E-12	75.0%		4>9	1	3	0.00618	0.40	2.53
SC_00005273.1.1	aldose 1-epimerase	3,10E-165	71.0%		9>4	5	2	0.00265	2.53	0.39
SC_OS09G0108600_ORYSJ.1.1	aldose 1-epimerase	2,00E-23	77.0%		9>4	62	18	0.03837	3.35	0.30
SC_OS03G0790500_ORYSJ.1.3	alpha/beta hydrolase fold-containing protein	1,00E-48	53.85%		9>4	15	2	0.00278	8.61	0.12
SC_AGAL_ORYSJ.16.30	alpha-galactosidase 1	3,80E-70	65.71%		4>9	4	10	0.04443	0.44	2.28

SC_00008291.1.1	aluminum induced protein with YGL and LRDR motif	1,40E-41	72.43%		9>4	5	2	0.02198	2.36	0.42
SC_ACCO_ACTDE.1.1	aminocyclopropanecarboxylate oxidase	1,20E-106	57.05%		4>9	1	5	0.02389	0.24	4.10
SC_00020973.1.1	ankyrin repeat domain-containing protein 2	8,80E-39	69.67%		4>9	2	5	0.03671	0.45	2.22
SC_00021038.1.1	ankyrin repeats-containing protein	7,50E-04	52.83%		9>4	16	3	0.00637	6.07	0.16
SC_OS03G0281000_ORYSJ.2.2	ankyrin repeats-containing protein	4,40E-37	69.14%		9>4	6	3	0.01282	2.09	0.48
SC_00039195.1.1	APO protein 4	6,90E-28	60.83%		4>9	2	3	0.04491	0.47	2.11
SC_00001064.1.1	aquaporin TIP1-1	1,10E-110	64.8%		9>4	9	4	0.00006	2.12	0.47
SC_00002754.1.1	aquaporin TIP1-1	1,30E-97	64.8%		9>4	4	2	0.00753	2.37	0.42
SC_TIP11_ARATH.2.14	aquaporin TIP1-1	2,80E-39	61.15%		9>4	6	2	0.01093	2.94	0.34
SC_00015040.1.1	aquaporin TIP1-1	5,70E-58	64.0%		9>4	4	1	0.01351	3.14	0.32
SC_00009860.1.1	Arabinanase/levansucrase/invertase	2,60E-09	49.0%		9>4	18	6	0.01312	2.76	0.36
SC_SPE1_PEA.5.11	arginine decarboxylase 2	1,40E-37	73.0%		9>4	6	2	0.00319	2.68	0.37
SC_00024700.1.1	armadillo repeat-containing protein	1,20E-102	59.25%		9>4	1	1	0.01911	2.05	0.49
SC_00106270.1.1	armadillo/beta-catenin-like repeats-containing protein	9,40E-01	61.0%		4>9	4	14	0.01026	0.28	3.54
SC_AAT1_ARATH.2.11	aspartate aminotransferase	9,30E-57	69.86%	EC:2.6.1.1	4>9	4	8	0.00803	0.47	2.13
SC_OS02G0236000_ORYSJ.2.5	aspartate aminotransferase	6,20E-37	64.0%	EC:2.6.1.0	4>9	2	4	0.03550	0.46	2.19
SC_00027439.1.1	aspartyl protease family protein	4,80E-10	54.89%		9>4	16	8	0.03614	2.01	0.50
SC_00025099.1.1	aspartyl protease-like protein	1,50E-68	40.2%		9>4	2	1	0.04393	2.23	0.45
SC_00028565.1.1	aspartyl protease-like protein	1,40E-47	46.95%		9>4	8	4	0.04451	2.19	0.46
SC_ATPG3_IPOBA.10.39	ATP synthase subunit gamma	3,30E-50	68.33%		9>4	6	3	0.01379	2.05	0.49
SC_00008346.1.1	ATPase, F0/V0 complex, subunit C protein	9,60E-56	78.33%		9>4	7	3	0.01319	2.12	0.47
SC_00000790.1.1	ATPase, V0 complex, subunit E	7,30E-30	85.5%		9>4	10	4	0.01308	2.63	0.38
SC_00113252.1.1	ATP-dependent DNA helicase 2 subunit KU80	2,20E-180	73.0%		4>9	15	32	0.04882	0.49	2.06
SC_LAX4_MEDTR.2.5	auxin transporter protein 1	0,00E+00	52.6%		9>4	3	1	0.04074	3.60	0.28
SC_00030864.1.1	axi 1 protein-like protein	6,80E-12	60.33%		9>4	5	3	0.01219	2.05	0.49
SC_BI1_ORYSJ.11.15	BAX inhibitor 1	3,60E-28	68.83%		9>4	4	1	0.00639	2.55	0.39
SC_00117887.1.1	beta glucosidase 40	1,10E-08	70.64%		9>4	21	8	0.04705	2.68	0.37
SC_00131967.1.1	beta-1,3-glucanase-like protein	7,80E-54	65.85%		4>9	2	6	0.04704	0.37	2.67
SC_00072606.1.1	beta-amylase	6,70E-01	53.0%		9>4	54	19	0.04320	2.83	0.35
SC_00010652.1.1	beta-galactosidase 3	1,70E-132	51.9%		4>9	1	2	0.00126	0.34	2.94
SC_BGAL6_ORYSJ.2.6	beta-galactosidase 3	6,70E-94	53.65%		9>4	11	5	0.04229	2.10	0.48
SC_00109824.1.1	beta-galactosidase 8	6,90E-06	51.89%		4>9	1	1	0.03656	0.44	2.27
SC_00053179.1.1	bifunctional enolase 2/transcriptional activator	3,50E-136	75.33%	EC:4.2.1.11	4>9	4	14	0.00519	0.28	3.51
SC_ENO1_HEVBR.16.31	bifunctional enolase 2/transcriptional activator	7,10E-80	78.67%	EC:4.2.1.11	4>9	4	12	0.00999	0.33	3.02
SC_ENO1_HEVBR.7.31	bifunctional enolase 2/transcriptional activator	0,00E+00	76.33%	EC:4.2.1.11	4>9	11	29	0.01329	0.39	2.59
SC_ENO1_HEVBR.14.31	bifunctional enolase 2/transcriptional activator	0,00E+00	82.33%	EC:4.2.1.11	4>9	6	18	0.03026	0.34	2.95
SC_00157718.1.1	bifunctional enolase 2/transcriptional activator	3,60E-68	81.33%	EC:4.2.1.11	4>9	1	5	0.03257	0.30	3.37

SC_00160291.1.1	bifunctional enolase 2/transcriptional activator	1,90E-97	74.0%	EC:4.2.1.11	4>9	2	6	0.03263	0.27	3.66
SC_ENO_SOLLC.9.19	bifunctional enolase 2/transcriptional activator	1,60E-68	85.67%	EC:4.2.1.11	4>9	15	32	0.03991	0.46	2.15
SC_ENO1_HEVBR.1.31	bifunctional enolase 2/transcriptional activator	1,10E-77	74.0%	EC:4.2.1.11	4>9	8	23	0.04141	0.36	2.76
SC_ENO_ORYSJ.4.15	bifunctional enolase 2/transcriptional activator	3,00E-77	86.33%	EC:4.2.1.11	4>9	3	9	0.04939	0.28	3.57
SC_00065615.1.1	brassinosteroid-responsive RING-H2	2,20E-23	55.05%		9>4	9	4	0.00037	2.36	0.42
SC_00018125.1.1	C2H2-like zinc finger protein	7,70E-65	60.88%		9>4	5	2	0.00399	3.15	0.32
SC_00009738.1.1	C2H2-like zinc finger protein	5,60E-62	66.63%		9>4	40	15	0.04100	2.68	0.37
SC_00076703.1.1	C3H4 type zinc finger protein	6,70E-16	63.67%		4>9	6	12	0.04106	0.49	2.02
SC_00071518.1.1	Ca(2+)-dependent nuclease-like protein	3,10E-36	70.0%		4>9	1	4	0.02824	0.35	2.85
SC_AGD11_ARATH.6.11	calcium-dependent lipid-binding domain	5,00E-49	64.55%		9>4	12	4	0.01681	2.88	0.35
SC_00144989.1.1	Calcium-dependent lipid-binding domain-containing protein	1,80E-10	48.3%		9>4	10	3	0.02653	4.10	0.24
SC_CAM1_MAIZE.1.1	calmodulin 5	1,30E-11	63.13%		4>9	2	6	0.00201	0.41	2.44
SC_00043177.1.1	calmodulin 7	1,60E-09	69.67%		9>4	8	4	0.04134	2.14	0.47
SC_00004739.1.1	calmodulin-domain protein kinase cdpk isoform 2	4,30E-64	66.65%		9>4	5	2	0.00360	2.37	0.42
SC_00024580.1.1	calmodulin-like protein 6	7,30E-44	61.35%		9>4	4	2	0.04594	2.08	0.48
SC_00029539.1.1	carbon catabolite repressor protein 4-like 4	4,00E-48	53.79%		4>9	2	4	0.01734	0.37	2.70
SC_00021135.1.1	carbonic anhydrase	1,10E-82	62.3%	EC:3.6.3.0	9>4	30	9	0.02501	3.17	0.32
SC_00011647.1.1	carboxyesterase 17	1,00E-36	56.6%		9>4	26	7	0.00412	3.78	0.26
SC_CYP306A1_BOMMO.843.3286	carboxypeptidase D	2,20E-01	46.0%		9>4	13	3	0.00216	4.01	0.25
SC_UMC2614_MAIZE.1.1	cation-chloride co-transporter 1	8,70E-15	59.5%		9>4	2	1	0.00470	2.33	0.43
SC_CIPKC_ARATH.4.4	CBL-interacting serine/threonine-protein kinase 12	1,80E-44	61.75%		9>4	4	2	0.01124	2.35	0.43
SC_00005309.1.1	CBS / octicosapeptide/Phox/Bemp1 domain-containing protein	1,70E-165	64.38%		4>9	4	8	0.01662	0.49	2.02
SC_00065520.1.1	CCHC-type zinc knuckle protein	5,00E-01	46.0%		4>9	3	8	0.00160	0.44	2.30
SC_00073192.1.1	CCHC-type zinc knuckle protein	4,90E-09	52.5%		4>9	2	7	0.03787	0.27	3.76
SC_00055879.1.1	CDPK-related kinase	1,70E-59	62.1%	EC:2.7.11.17	4>9	3	6	0.04787	0.47	2.12
SC_CDC48_SOYBN.26.34	cell division control protein 48-e	3,70E-102	57.7%		4>9	2	3	0.00987	0.47	2.13
SC_CDC48_SOYBN.28.34	cell division control protein 48-e	2,10E-128	83.0%	EC:3.6.1.3	4>9	1	3	0.01187	0.50	2.01
SC_00012817.1.1	cell growth defect factor -2	7,70E-28	76.0%		9>4	5	2	0.04087	2.44	0.41
SC_00106481.1.1	chaperone DnaJ-domain containing protein	1,40E-28	70.83%		4>9	4	8	0.04359	0.43	2.30
SC_00060745.1.1	chaperone protein dnaJ 11	2,40E-21	61.35%		4>9	3	15	0.02973	0.22	4.55
SC_00015744.1.1	chaperone protein dnaJ 11	1,00E-16	60.05%		4>9	3	7	0.04071	0.38	2.64
SC_DNAJ8_ARATH.1.8	chaperone protein dnaJ 8	1,20E-18	58.2%		9>4	4	2	0.00326	2.45	0.41
SC_00025909.1.1	chitinase	2,40E-73	58.0%		9>4	6	2	0.02160	3.67	0.27
SC_00008805.1.1	chitinase-like protein	2,30E-83	57.85%		9>4	25	12	0.02076	2.10	0.48
SC_CISY4_ARATH.6.8	citrate synthase 4	8,00E-16	68.0%	EC:2.3.3.0	4>9	2	4	0.04551	0.45	2.20
SC_OS01G0138800_ORYSJ.1.1	class I glutamine amidotransferase domain-containing protein	0,00E+00	72.0%		9>4	21	10	0.03746	2.17	0.46
SC_CHI8_ORYSJ.2.2	class II chitinase-like protein	7,90E-42	52.77%		9>4	25	7	0.01646	3.49	0.29
SC_CHI1_ORYSJ.4.9	class II chitinase-like protein	1,10E-82	53.54%		9>4	18	6	0.01811	2.92	0.34
SC_00107099.1.1	COP9 signalosome complex subunit 3	2,90E-20	65.0%		4>9	3	8	0.01169	0.34	2.98
SC_CYT3_ARATH.14.22	cysteine proteinase inhibitor 3	3,10E-20	66.88%		9>4	4	1	0.02590	3.27	0.31

SC_00049222.1.1	cysteine-rich receptor-like protein kinase 8	1,50E-43	60.75%		4>9	2	8	0.00106	0.24	4.16
SC_00064845.1.1	cysteine-rich receptor-like protein kinase 8	6,60E-20	55.33%		4>9	12	33	0.00189	0.38	2.63
SC_00148446.1.1	cysteine-rich receptor-like protein kinase 8	6,40E-24	49.75%		4>9	8	27	0.00451	0.29	3.42
SC_00057819.1.1	cysteine-rich receptor-like protein kinase 8	3,10E-56	61.0%		4>9	19	63	0.01302	0.31	3.25
SC_00075184.1.1	cysteine-rich receptor-like protein kinase 8	8,80E-52	61.75%		4>9	2	5	0.01653	0.42	2.40
SC_00136739.1.1	cysteine-rich receptor-like protein kinase 8	4,80E-16	58.0%		4>9	17	54	0.01701	0.32	3.12
SC_00093502.1.1	cysteine-rich receptor-like protein kinase 8	8,90E-53	62.0%		4>9	2	7	0.02264	0.37	2.69
SC_00134391.1.1	cysteine-rich receptor-like protein kinase 8	2,70E-21	58.0%		4>9	2	4	0.03919	0.41	2.41
SC_00034864.1.1	cysteine-rich receptor-like protein kinase 8	4,60E-88	63.0%		4>9	17	35	0.04419	0.49	2.05
SC_00064111.1.1	cysteine-rich receptor-like protein kinase 8	4,40E-23	57.33%		4>9	4	8	0.04630	0.44	2.29
SC_POLX_TOBAC.204.242	cysteine-rich receptor-like protein kinase 8	1,40E-55	55.5%		4>9	48	101	0.04911	0.48	2.10
SC_C75A1_PINTA.20.102	cytochrome P450 71A21	9,70E-87	59.8%		9>4	7	4	0.00298	2.04	0.49
SC_00042685.1.1	cytochrome P450 76C1	1,10E-40	69.35%		9>4	5	2	0.00335	3.38	0.30
SC_C76C3_ARATH.1.1	cytochrome P450 76C1	3,20E-110	60.25%		9>4	4	2	0.00447	2.07	0.48
SC_00029032.1.1	cytochrome P450, family 77, subfamily A, polypeptide 5 pseudogene	1,80E-166	61.0%		4>9	3	7	0.02087	0.47	2.14
SC_C78A4_PINRA.9.41	cytochrome P450, family 78, subfamily A, polypeptide 5	3,60E-48	59.1%	EC:1.14.13.0	9>4	4	2	0.00932	2.16	0.46
SC_00051079.1.1	cytochrome P450, family 87, subfamily A, polypeptide 6	6,30E-88	65.6%		9>4	15	5	0.00732	3.17	0.32
SC_00091528.1.1	cytokinin dehydrogenase 1	5,30E-01	64.0%		4>9	2	7	0.01476	0.36	2.81
SC_00044313.1.1	cytokinin dehydrogenase 7	8,60E-55	64.25%		4>9	2	5	0.03632	0.35	2.83
SC_00042033.1.1	D-3-phosphoglycerate dehydrogenase	4,60E-14	75.67%	EC:1.1.1.95	4>9	15	35	0.04295	0.42	2.36
SC_00012876.1.1	D-arabinono-1,4-lactone oxidase family protein	2,30E-58	54.71%		9>4	35	6	0.00455	5.40	0.19
SC_OS02G0124600_ORYSJ.2.2	D-arabinono-1,4-lactone oxidase family protein	2,70E-27	68.57%		9>4	9	2	0.01288	5.42	0.18
SC_DEF1_PINSY.27.41	defensin-like protein 7	1,90E-06	53.89%		4>9	1	3	0.03486	0.41	2.43
SC_00000544.1.1	dehydration-induced protein ERD15	1,40E-18	55.5%		9>4	7	3	0.03434	2.10	0.48
SC_DRE2C_ARATH.1.1	dehydration-responsive element-binding protein 2A	9,90E-21	78.75%		4>9	1	4	0.02142	0.36	2.81
SC_00001231.1.1	dehydrin COR47	4,50E-01	43.0%		9>4	40	17	0.03246	2.35	0.43
SC_00025556.1.1	DEK domain-containing chromatin associated protein	6,30E-31	64.0%		9>4	9	4	0.04745	2.05	0.49
SC_00162891.1.1	dicarboxylate transport 2.1	6,90E-13	75.5%		9>4	68	26	0.03747	2.62	0.38
SC_DFRA_MALDO.1.17	dihydroflavonol-4-reductase	1,30E-159	61.2%	EC:1.1.1.219	9>4	5	2	0.02743	2.86	0.35
SC_00009355.1.1	Dihydroneopterin aldolase	2,80E-42	74.25%		4>9	1	3	0.03983	0.46	2.16
SC_00009289.1.1	diphosphomevalonate decarboxylase	1,50E-20	63.0%		4>9	1	3	0.02918	0.40	2.48
SC_00105899.1.1	disease resistance-responsive, dirigent domain-containing protein	1,10E-52	61.75%		9>4	3	1	0.00564	2.25	0.45
SC_00006247.1.1	disease resistance-responsive, dirigent domain-containing protein	2,50E-56	61.05%		9>4	4	1	0.00664	2.88	0.35
SC_00060581.1.1	DNA binding protein GT-1	1,20E-125	55.06%		4>9	1	2	0.04248	0.45	2.24
SC_00075722.1.1	DNA-directed RNA polymerase III subunit C19	1,20E-12	62.0%		9>4	13	6	0.00298	2.25	0.44
SC_OS03G0656201_ORYSJ.1.1	DnaJ/Hsp40 cysteine-rich domain-containing protein	1,50E-22	54.5%		9>4	4	2	0.00788	2.34	0.43

SC_00018728.1.1	dormancy/auxin associated protein	6,50E-13	64.6%		9>4	14	6	0.00101	2.30	0.43
SC_UMC2505_MAIZE.3.5	dormancy/auxin associated protein	1,00E-13	63.4%		9>4	29	14	0.00345	2.06	0.49
SC_00001380.1.1	double-stranded DNA-binding-like protein	4,60E-37	68.5%		9>4	3	1	0.01570	2.04	0.49
SC_00040688.1.1	dual specificity protein phosphatase family protein	5,40E-50	71.0%	EC:3.1.3.16	4>9	1	3	0.02818	0.39	2.54
SC_RHF2A_ARATH.4.7	E3 ubiquitin-protein ligase RHF1A	1,10E-12	69.67%		4>9	2	5	0.02367	0.48	2.08
SC_00024848.1.1	early nodulin-related protein	1,80E-13	77.0%		4>9	4	21	0.01054	0.19	5.16
SC_00006107.1.1	early nodulin-related protein	6,10E-13	77.0%		4>9	1	4	0.01824	0.26	3.82
SC_NO93_SOYBN.2.4	early nodulin-related protein	1,30E-15	79.0%		4>9	9	29	0.03329	0.30	3.36
SC_OS02G0613000_ORYSJ.1.1	EF-hand, calcium binding motif-containing protein	1,20E-57	53.7%		4>9	1	4	0.00846	0.26	3.91
SC_00000762.1.1	Elongation factor 1-alpha	1,20E-16	89.67%		9>4	9	4	0.03566	2.01	0.50
SC_00149373.1.1	Elongation factor 1-alpha	4,60E-65	75.86%		9>4	4	2	0.04981	2.52	0.40
SC_00027980.1.1	Elongation factor 1-delta 2	8,50E-33	87.0%		9>4	4	2	0.00787	2.48	0.40
SC_EF1D_PIMBR.17.70	Elongation factor 1-delta 2	2,30E-36	87.4%		9>4	5	2	0.03322	2.65	0.38
SC_00049595.1.1	elongator complex protein 4	1,70E-23	73.0%		4>9	1	3	0.02571	0.35	2.87
SC_LE193_HORVU.1.1	Em-like protein GEA6	4,30E-15	72.5%		9>4	50	12	0.02053	4.27	0.23
SC_00146691.1.1	endonuclease 4	7,90E-28	71.6%		9>4	8	3	0.03049	2.25	0.44
SC_00002981.1.1	endonuclease 4	5,00E-109	72.2%		9>4	7	3	0.04358	2.48	0.40
SC_00009443.1.1	ENTH/VHS/GAT family protein	4,80E-03	58.0%		4>9	6	13	0.03411	0.47	2.13
SC_OS06G0687800_ORYSJ.6.8	EPS15 homology domain 1 protein	1,00E-30	79.67%		4>9	2	4	0.04639	0.47	2.12
SC_OS04G0376700_ORYSJ.1.1	ER lumen protein retaining receptor	2,40E-70	67.57%		9>4	3	1	0.04087	2.46	0.41
SC_00129785.1.1	ERD1 protein	1,10E-22	79.0%		9>4	13	6	0.02479	2.28	0.44
SC_EF101_ARATH.1.1	ethylene-responsive transcription factor 2	2,60E-24	78.6%		9>4	6	2	0.02182	2.52	0.40
SC_00157909.1.1	ethylene-responsive transcription factor 8	1,00E-02	68.91%		9>4	4	2	0.00794	2.29	0.44
SC_00148594.1.1	exostosin-like protein	5,80E-12	52.3%		9>4	11	5	0.03692	2.17	0.46
SC_00008084.1.1	Fatty acid/sphingolipid desaturase	1,10E-49	66.0%		9>4	4	2	0.04595	2.32	0.43
SC_00019830.1.1	F-box only protein 6	1,20E-18	40.0%		4>9	2	6	0.01029	0.41	2.42
SC_00010566.1.1	F-box protein	4,00E-98	64.0%		9>4	2	1	0.03497	2.54	0.39
SC_00012195.1.1	F-box protein	5,70E-17	51.13%		9>4	9	4	0.04468	2.22	0.45
SC_FBK46_ARATH.1.2	F-box/kelch-repeat protein	5,80E-37	43.2%		4>9	4	33	0.00435	0.13	7.81
SC_00021518.1.1	F-box/kelch-repeat protein	3,60E-10	53.85%		9>4	3	1	0.02351	2.37	0.42
SC_00099800.1.1	F-box/kelch-repeat protein	4,10E-02	48.5%		4>9	1	2	0.03061	0.48	2.09
SC_FBK30_ARATH.2.2	F-box/kelch-repeat protein	1,50E-29	49.35%		9>4	8	3	0.03536	3.09	0.32
SC_00034885.1.1	F-box/kelch-repeat protein SKIP20	1,70E-04	47.5%		9>4	37	8	0.03158	4.46	0.22
SC_OS09G0454300_ORYSJ.1.2	F-box/RNI-like protein	1,10E-11	47.0%		4>9	13	29	0.02057	0.45	2.23
SC_FRI3_SOYBN.11.17	ferritin heavy chain	2,80E-31	68.0%		9>4	31	12	0.01296	2.60	0.38
SC_FLS_CITUN.2.7	flavonol synthase/flavanone 3-hydroxylase	2,10E-120	59.05%	EC:1.14.11.0	9>4	8	3	0.03847	2.68	0.37
SC_F16P2_SPIOL.1.5	fructose-1,6-bisphosphatase, cytosolic	0,00E+00	66.25%	EC:3.1.3.11	9>4	9	4	0.00761	2.64	0.38
SC_ALF_SPIOL.3.20	fructose-bisphosphate aldolase, class I	5,10E-53	81.62%		4>9	2	4	0.04247	0.50	2.02
SC_00146640.1.1	Gag-Pol-related retrotransposon family protein	6,00E-03	46.67%		4>9	2	6	0.01022	0.33	3.05
SC_00161992.1.1	Gag-Pol-related retrotransposon family protein	2,20E-03	53.0%		4>9	3	6	0.03482	0.47	2.13
SC_OS08G0125300_ORYSJ.3.3	Gag-Pol-related retrotransposon family protein	1,60E-02	41.75%		4>9	4	9	0.03862	0.42	2.37

SC_00009357.1.1	galactinol synthase 4	4,90E-165	59.47%	EC:2.4.1.123	9>4	6	2	0.03594	3.24	0.31
SC_00105779.1.1	gamma-glutamyl transpeptidase 4	9,60E-53	74.0%	EC:2.3.2.2; EC:2.6.1.0; EC:2.3.2.15	4>9	1	4	0.00708	0.38	2.60
SC_00130256.1.1	GATA type zinc finger transcription factor-like protein	1,40E-98	77.88%		9>4	12	6	0.00033	2.11	0.47
SC_00024287.1.1	GCN5-related N-acetyltransferase (GNAT) family protein	6,00E-47	63.67%		9>4	6	2	0.01250	3.00	0.33
SC_00034295.1.1	GCN5-related N-acetyltransferase-like protein	2,50E-50	50.0%		9>4	9	3	0.02766	3.17	0.32
SC_00010373.1.1	germin-like protein subfamily 1 member 1	8,00E-57	59.0%		9>4	69	30	0.00348	2.28	0.44
SC_GL11_ARATH.1.2	germin-like protein subfamily 1 member 1	5,00E-24	62.8%		9>4	94	39	0.00837	2.37	0.42
SC_00006936.1.1	germin-like protein subfamily 2 member 2	2,60E-56	62.25%		9>4	5	1	0.01430	3.15	0.32
SC_00033457.1.1	Gibberellin-regulated family protein	2,20E-20	59.82%		9>4	6	3	0.02906	2.30	0.44
SC_00145769.1.1	glucan endo-1,3-beta-glucosidase 3	1,20E-67	57.3%		9>4	3	1	0.04556	2.44	0.41
SC_00104524.1.1	glucose-1-phosphate adenylyltransferase large subunit 2	2,00E-69	70.67%	EC:2.7.7.27	9>4	4	2	0.01774	2.01	0.50
SC_00149171.1.1	glucose-6-phosphate 1-epimerase	6,60E-54	65.43%		9>4	11	5	0.02076	2.12	0.47
SC_GLRX_RICCO.2.33	glutaredoxin-C2	8,20E-50	60.2%		9>4	9	4	0.04520	2.33	0.43
SC_00026575.1.1	glutathione S-transferase (class phi) 14	2,60E-01	41.0%		9>4	14	7	0.02354	2.16	0.46
SC_G3PC_PINSY.4.66	glyceraldehyde 3-phosphate dehydrogenase	1,10E-165	74.2%	EC:1.2.1.9; EC:1.2.1.12	4>9	1	3	0.01545	0.44	2.26
SC_G3PC_PINSY.36.66	glyceraldehyde-3-phosphate dehydrogenase, cytosolic	1,10E-65	79.2%	EC:1.2.1.9; EC:1.2.1.12	4>9	1	2	0.00714	0.48	2.08
SC_G3PC_PINSY.48.66	glyceraldehyde-3-phosphate dehydrogenase, cytosolic	2,10E-134	74.3%	EC:1.2.1.9; EC:1.2.1.12	4>9	1	3	0.04317	0.44	2.25
SC_00032109.1.1	glycerol kinase	9,40E-33	56.5%	EC:2.7.1.0	9>4	5	2	0.01861	2.29	0.44
SC_GCSP_PEA.4.14	glycine dehydrogenase	7,50E-51	72.33%		9>4	4	2	0.00012	2.42	0.41
SC_00033901.1.1	Glycosyltransferase family 61 protein	1,90E-103	55.78%		9>4	21	9	0.01937	2.40	0.42
SC_00040015.1.1	glyoxalase II 3	7,20E-65	66.14%		9>4	7	3	0.00492	2.13	0.47
SC_00018441.1.1	growth-regulating factor 5	2,00E-43	64.08%		4>9	1	2	0.02313	0.36	2.80
SC_00083548.1.1	GTP1/OBG family protein	1,50E-06	68.0%		9>4	4	2	0.01160	2.34	0.43
SC_00107620.1.1	H/ACA ribonucleoprotein complex subunit 4	4,30E-54	69.5%	EC:5.4.99.12	4>9	2	6	0.02641	0.43	2.33
SC_00047245.1.1	HAD superfamily, subfamily IIIB acid phosphatase	2,70E-32	63.64%		9>4	5	2	0.00265	2.43	0.41
SC_PPA1_SOLLC.4.14	HAD superfamily, subfamily IIIB acid phosphatase	1,10E-30	62.58%		9>4	2	1	0.01771	2.10	0.48
SC_00079315.1.1	HAD superfamily, subfamily IIIB acid phosphatase	8,10E-42	59.38%		9>4	7	2	0.02046	3.66	0.27
SC_00148235.1.1	HAD superfamily, subfamily IIIB acid phosphatase	1,10E-42	59.82%		9>4	11	4	0.02211	2.87	0.35
SC_00067581.1.1	HAD superfamily, subfamily IIIB acid phosphatase	3,30E-63	58.67%		9>4	29	2	0.03790	12.20	0.08
SC_PPA1_SOLLC.8.14	HAD superfamily, subfamily IIIB acid phosphatase	9,10E-07	61.0%		9>4	12	5	0.03965	2.65	0.38

SC_00032802.1.1	haloacid dehalogenase-like hydrolase domain-containing protein	4,60E-34	55.5%		4>9	1	2	0.00728	0.46	2.20
SC_00022638.1.1	haloacid dehalogenase-like hydrolase domain-containing protein	2,80E-40	59.4%		9>4	10	4	0.01375	2.55	0.39
SC_00150326.1.1	hAT dimerization domain-containing protein	8,50E-07	50.8%		4>9	2	4	0.03253	0.45	2.22
SC_HSP72_SOLLC.6.19	heat shock 70kDa protein 1/8	1,00E-97	73.11%		4>9	2	5	0.00266	0.41	2.45
SC_HSP7C_PETHY.30.45	heat shock 70kDa protein 1/8	4,00E-66	71.17%		4>9	5	10	0.04853	0.48	2.07
SC_HS181_ORYSJ.5.10	heat shock protein 18.2	2,00E-26	73.31%		9>4	8	2	0.02668	3.55	0.28
SC_00151185.1.1	heat shock protein 18.2	5,00E-26	74.47%		9>4	10	3	0.03033	3.62	0.28
SC_HS181_ORYSJ.3.10	heat shock protein 18.2	3,00E-22	73.33%		9>4	34	8	0.04366	4.45	0.22
SC_00056772.1.1	heat shock protein 21	9,80E-04	63.2%		4>9	7	14	0.03143	0.48	2.09
SC_00007286.1.1	heat shock protein 70-4	6,80E-79	72.0%		4>9	9	18	0.03832	0.48	2.08
SC_00000112.1.1	high mobility group B2 protein	4,20E-30	62.65%		4>9	3	6	0.01387	0.50	2.00
SC_00152437.1.1	high mobility group B3 protein	8,70E-20	68.06%		4>9	2	5	0.00198	0.49	2.02
SC_00008087.1.1	high mobility group B3 protein	5,60E-06	67.0%		4>9	1	3	0.04306	0.45	2.22
SC_00041622.1.1	high response to osmotic stress 10	1,90E-13	70.95%		9>4	62	27	0.00084	2.32	0.43
SC_00012695.1.1	histidine kinase 4	2,70E-81	65.78%	EC:2.7.13.3; EC:3.1.3.16	4>9	0	1	0.04013	0.42	2.40
SC_00014854.1.1	histone deacetylase complex subunit SAP18	8,00E-35	95.0%		4>9	2	4	0.02907	0.48	2.07
SC_H2A6_ORYSJ.1.9	histone H2A 13	1,90E-43	83.31%		9>4	3	2	0.01546	2.14	0.47
SC_H33_ARATH.29.70	histone H3	8,30E-82	92.67%		9>4	15	6	0.00360	2.42	0.41
SC_H4_OLILU.1.5	histone H4	2,70E-29	91.0%		9>4	8	4	0.02342	2.07	0.48
SC_00079928.1.1	histone H4	4,80E-28	91.0%		9>4	6	2	0.03794	3.04	0.33
SC_00160008.1.1	homolog of anti-oxidant 1	2,30E-27	69.05%	EC:6.6.1.1; EC:2.1.1.43	9>4	18	5	0.02919	3.35	0.30
SC_00003722.1.1	homolog of anti-oxidant 1	1,00E-33	67.55%	EC:6.6.1.1; EC:2.1.1.43	9>4	7	4	0.03027	2.03	0.49
SC_00003952.1.1	homolog of anti-oxidant 1	5,30E-34	66.75%	EC:6.6.1.1; EC:2.1.1.43	9>4	11	5	0.04668	2.36	0.42
SC_LOC100191955_MAIZE.1.1	HXXXD-type acyl-transferase-like protein	1,60E-24	55.8%		9>4	9	4	0.02439	2.10	0.48
SC_00061026.1.1	HXXXD-type acyl-transferase-like protein	5,80E-13	48.21%	EC:2.3.1.0	9>4	13	6	0.02495	2.33	0.43
SC_00154817.1.1	hydroxyproline-rich glycoprotein-like protein	2,20E-57	75.0%		9>4	7	4	0.01099	2.08	0.48
SC_00073636.1.1	hydroxysteroid dehydrogenase 1	4,30E-14	54.33%		9>4	18	7	0.03442	2.65	0.38
SC_POLX_TOBAC.128.242	hypothetical protein ArthMp028	2,30E-19	53.0%		4>9	2	5	0.01396	0.38	2.63
SC_00159201.1.1	hypothetical protein ArthMp028	8,00E-21	52.33%		4>9	5	12	0.04335	0.38	2.62
SC_POLX_TOBAC.225.242	hypothetical protein ArthMp028	1,10E-20	57.0%		4>9	2	4	0.04972	0.40	2.48
SC_00059893.1.1	hypothetical protein ArthMp059	5,20E-14	54.0%		4>9	0	1	0.04463	0.28	3.53
SC_00061435.1.1	hypothetical protein ArthMp063	4,80E-10	55.0%		4>9	8	17	0.01336	0.44	2.27
SC_00106634.1.1	hypothetical protein ArthMp067	6,90E-08	71.0%		4>9	2	4	0.03046	0.44	2.29
SC_00053538.1.1	hypothetical protein ArthMp070	1,60E-21	56.0%		4>9	2	6	0.01579	0.37	2.73
SC_00147409.1.1	hypothetical protein ArthMp071	4,50E-03	47.5%		4>9	3	10	0.04172	0.30	3.37
SC_00077143.1.1	hypothetical protein ArthMp074	5,60E-01	66.0%		4>9	1	2	0.01756	0.46	2.15
SC_00069808.1.1	hypothetical protein ArthMp075	1,80E-29	59.0%		4>9	4	13	0.00267	0.28	3.57
SC_00030482.1.1	hypothetical protein ArthMp075	2,90E-37	69.0%		4>9	2	4	0.01688	0.48	2.07
SC_00129109.1.1	hypothetical protein ArthMp075	1,70E-13	57.0%		4>9	4	12	0.03441	0.33	3.06
SC_00048728.1.1	hypothetical protein ArthMp075	4,60E-22	64.0%		4>9	3	11	0.04221	0.31	3.21
SC_00083162.1.1	IBR domain containing protein	1,60E-02	47.0%		4>9	4	33	0.02022	0.11	9.03

SC_MIOX_ORYSJ.3.8	inositol oxygenase	1,50E-53	82.5%	EC:1.13.99.1	9>4	155	48	0.03322	3.25	0.31
SC_00014095.1.1	inositol oxygenase 1	3,50E-71	85.0%	EC:1.13.99.1	9>4	24	8	0.00436	2.87	0.35
SC_MIOX_ORYSJ.1.8	inositol oxygenase 1	1,30E-151	76.0%	EC:1.13.99.1	9>4	59	25	0.01591	2.36	0.42
SC_00026140.1.1	inositol oxygenase 1	1,10E-43	84.75%	EC:1.13.99.1	9>4	11	5	0.02911	2.29	0.44
SC_00040126.1.1	inositol oxygenase 1	2,30E-56	69.0%		9>4	34	16	0.03139	2.17	0.46
SC_00045575.1.1	inositol oxygenase 1	1,70E-40	87.0%	EC:1.13.99.1	9>4	14	6	0.04980	2.46	0.41
SC_00040249.1.1	Intron maturase, type II family protein	9,40E-125	78.5%		4>9	2	5	0.04593	0.47	2.14
SC_00150045.1.1	invertase/pectin methylesterase inhibitor-like protein	8,70E-26	65.9%		9>4	40	19	0.04511	2.08	0.48
SC_LAR_DESUN.3.7	Isoflavone reductase-P3	7,70E-34	58.25%		9>4	8	4	0.00595	2.13	0.47
SC_00062410.1.1	Isoflavone reductase-P3	1,80E-23	63.71%		4>9	1	3	0.02480	0.42	2.36
SC_00016186.1.1	lactoylglutathione lyase / glyoxalase I-like protein	1,30E-68	60.63%		9>4	192	79	0.00631	2.43	0.41
SC_CYP306A1_BOMMO.301.3286	Lamin-like protein	8,90E-01	56.0%		9>4	5	2	0.01689	2.26	0.44
SC_00015352.1.1	Lamin-like protein	3,70E-09	54.5%		9>4	12	6	0.03896	2.26	0.44
SC_OS08G0404300_ORYSJ.2.2	large subunit ribosomal protein L27	2,20E-59	74.67%		4>9	1	3	0.02593	0.47	2.13
SC_CYP306A1_BOMMO.407.3286	late embryogenesis abundant domain-containing protein	7,70E-01	59.0%		9>4	22	7	0.02833	3.25	0.31
SC_LEA11_HELAN.2.5	Late embryogenesis abundant protein, group 1 protein	7,60E-07	69.0%		9>4	32	9	0.03604	3.41	0.29
SC_00025014.1.1	Late embryogenesis abundant protein, group 6	1,30E-03	65.0%		9>4	7	2	0.00634	3.53	0.28
SC_00013261.1.1	late embryogenesis abundant protein-like protein	8,40E-76	59.67%		9>4	7	3	0.00268	2.44	0.41
SC_00033538.1.1	late embryogenesis abundant protein-like protein	3,10E-59	64.67%		9>4	44	16	0.03822	2.82	0.35
SC_OS07G0109600_ORYSJ.1.2	leo1-like family protein	4,20E-11	64.0%		9>4	3	1	0.04510	2.22	0.45
SC_00044107.1.1	Leucine carboxyl methyltransferase	1,00E-65	66.0%		4>9	10	22	0.00918	0.44	2.29
SC_00017300.1.1	leucine-rich receptor-like protein kinase	1,20E-76	50.85%		9>4	2	1	0.02135	2.95	0.34
SC_00052279.1.1	leucine-rich receptor-like protein kinase	9,60E-47	51.0%		9>4	3	1	0.03859	2.53	0.40
SC_00100016.1.1	leucyl-tRNA synthetase	2,10E-88	63.0%		4>9	2	4	0.03115	0.48	2.08
SC_00011966.1.1	light-regulated zinc finger protein 1	4,60E-35	61.4%		9>4	11	5	0.02149	2.17	0.46
SC_LOX2_SOYBN.1.4	lipoxygenase 5	2,20E-104	66.33%		9>4	14	2	0.00968	6.34	0.16
SC_LBD15_ARATH.2.6	LOB domain-containing protein 15	1,20E-37	73.2%		9>4	2	1	0.02524	2.85	0.35
SC_00016048.1.1	LOB domain-containing protein 40	1,60E-02	54.0%		4>9	1	2	0.02597	0.46	2.18
SC_00114075.1.1	lung seven transmembrane receptor-like protein	3,20E-55	77.17%		4>9	4	9	0.04109	0.49	2.06
SC_CYP306A1_BOMMO.3042.3286	LYR family of Fe/S cluster biogenesis protein	8,50E-08	55.0%		9>4	17	8	0.04343	2.14	0.47
SC_00061533.1.1	MAC/Perforin domain-containing protein	3,20E-44	62.0%		4>9	3	6	0.04150	0.45	2.22
SC_00019345.1.1	macrophage migration inhibitory factor family protein	4,80E-27	62.33%		9>4	3	2	0.03653	2.10	0.48
SC_00010646.1.1	MAK10-like protein	6,20E-165	60.67%		4>9	7	14	0.03740	0.49	2.04
SC_MAOX_VITVI.4.14	malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	3,50E-83	74.86%	EC:1.1.1.40	4>9	2	4	0.03051	0.43	2.31
SC_MDHM_CITLA.5.5	malate dehydrogenase 1	2,10E-81	76.9%	EC:1.1.1.37	4>9	1	2	0.02197	0.36	2.74
SC_00027266.1.1	Metallo-beta-lactamase domain-containing protein	4,70E-85	42.5%		4>9	1	2	0.03618	0.47	2.12

SC_YSL16_ORYSJ.1.1	metal-nicotianamine transporter YSL2	2,70E-79	62.67%		9>4	6	3	0.02627	2.26	0.44
SC_OS12G0623900_ORYSJ.8.15	methionine synthase 2	6,40E-53	92.67%	EC:2.1.1.13; EC:1.3.1.0; EC:2.1.1.14	9>4	29	9	0.01334	3.29	0.30
SC_METE_MESCR.27.30	methionine synthase 2	8,80E-70	90.33%	EC:2.1.1.13	9>4	5	2	0.01803	2.17	0.46
SC_00001884.1.1	methyl esterase 1	1,10E-61	54.7%		9>4	47	19	0.03624	2.47	0.41
SC_00070671.1.1	mevalonate kinase	1,00E-09	73.0%	EC:2.7.1.0	9>4	29	10	0.00315	2.87	0.35
SC_00003972.1.1	mitochondrial outer membrane protein porin 2	2,40E-20	77.2%		4>9	2	4	0.02286	0.48	2.08
SC_00067062.1.1	Mitochondrial substrate carrier family protein	8,20E-22	50.57%		4>9	1	2	0.04436	0.41	2.41
SC_MMK1_MEDSA.2.2	mitogen-activated protein kinase 6	3,70E-54	83.3%	EC:2.7.11.24	4>9	2	5	0.01426	0.40	2.48
SC_MZT1_PICSI.1.3	mitotic-spindle organizing protein 1B	3,80E-23	82.0%		9>4	6	3	0.00049	2.08	0.48
SC_LOC100193989_MAIZE.1.1	monothiol glutaredoxin-S16	3,40E-49	61.17%		4>9	1	3	0.00084	0.31	3.20
SC_GRS12_ORYSJ.1.2	monothiol glutaredoxin-S16	9,30E-70	53.11%		4>9	5	10	0.01313	0.50	2.01
SC_00148168.1.1	mRNA splicing factor, thioredoxin-like U5 snRNP	4,50E-92	76.0%		9>4	7	2	0.02360	3.42	0.29
SC_00007535.1.1	myb domain protein 117	7,00E-66	77.5%		4>9	15	31	0.02402	0.50	2.01
SC_00098050.1.1	myb domain protein 56	2,00E-16	75.15%		4>9	2	4	0.03411	0.47	2.13
SC_00027363.1.1	NAC domain containing protein 32	1,60E-68	70.2%		9>4	3	1	0.00731	2.39	0.42
SC_00010295.1.1	NAC domain containing protein 36	5,20E-52	65.1%		4>9	5	16	0.00188	0.34	2.94
SC_00004156.1.1	NAC domain containing protein 47	2,00E-77	72.45%		9>4	4	2	0.01488	2.37	0.42
SC_OS05G0509200_ORYSJ.1.1	NADH dehydrogenase (ubiquinone) flavoprotein 2	1,90E-93	64.0%		9>4	5	2	0.00263	2.43	0.41
SC_FL3H_MALDO.19.23	Naringenin,2-oxoglutarate 3-dioxygenase	3,70E-91	58.95%	EC:1.14.11.9	9>4	13	3	0.00038	4.69	0.21
SC_FL3H_PETHY.5.8	Naringenin,2-oxoglutarate 3-dioxygenase	3,10E-122	58.7%	EC:1.14.11.9	9>4	8	2	0.00328	3.42	0.29
SC_FL3H_MALDO.12.23	Naringenin,2-oxoglutarate 3-dioxygenase	3,80E-81	52.15%	EC:1.14.11.9	9>4	5	1	0.01116	3.18	0.31
SC_FL3H_ARATH.5.7	Naringenin,2-oxoglutarate 3-dioxygenase	1,50E-34	61.58%	EC:1.14.11.9	9>4	7	2	0.01472	3.75	0.27
SC_FL3H_MALDO.2.23	Naringenin,2-oxoglutarate 3-dioxygenase	3,60E-131	60.45%	EC:1.14.11.9	9>4	8	2	0.03186	4.50	0.22
SC_NACA_PINTA.12.36	Nascent polypeptide-associated complex subunit alpha-like protein 1	4,30E-15	87.8%		4>9	1	2	0.03239	0.42	2.38
SC_RUB1_ARATH.2.6	NEDD8-like protein RUB1	2,00E-54	73.4%		9>4	9	4	0.03252	2.03	0.49
SC_RUB1_ARATH.3.6	NEDD8-like protein RUB1	1,90E-93	82.55%		9>4	9	4	0.04661	2.33	0.43
SC_00014161.1.1	NmrA-like negative transcriptional regulator family protein	1,50E-04	50.0%		9>4	9	3	0.01758	2.79	0.36
SC_00005301.1.1	N-MYC downregulated-like 1 protein	7,10E-129	82.75%		9>4	17	6	0.03378	3.05	0.33
SC_OS05G0493800_ORYSJ.1.2	nodulin MtN21 /EamA-like transporter family protein	1,30E-105	62.0%		9>4	5	1	0.01619	3.69	0.27
SC_00016623.1.1	nodulin MtN3-like protein	1,10E-76	61.82%		9>4	16	5	0.00807	3.23	0.31
SC_NLTP2_LENCU.9.30	non-specific lipid-transfer protein 12	6,90E-16	53.4%		9>4	13	4	0.02045	3.36	0.30
SC_NLTP_PINTA.219.283	non-specific lipid-transfer protein 3	5,70E-07	47.29%		9>4	12	5	0.02873	2.31	0.43
SC_NLTP_PINTA.5.283	non-specific lipid-transfer protein 5	4,70E-14	51.2%		9>4	9	4	0.02424	2.38	0.42
SC_00034320.1.1	non-specific phospholipase C6	5,70E-10	55.4%		9>4	11	3	0.01583	4.45	0.22
SC_00001690.1.1	nuclear poly(a) polymerase	2,80E-01	39.0%		9>4	27	13	0.00464	2.15	0.46
SC_00039591.1.1	nuclear RNA polymerase D1B	2,80E-01	57.5%		4>9	1	2	0.03355	0.37	2.72
SC_00079070.1.1	nucleotide exchange factor SIL1	1,30E-09	60.0%		4>9	1	3	0.00565	0.41	2.44

SC_00124126.1.1	Nucleotide-diphospho-sugar transferases superfamily protein	4,40E-73	71.0%		4>9	5	12	0.03751	0.44	2.30
SC_UGAL2_ARATH.2.7	Nucleotide-sugar transporter family protein	9,60E-20	75.57%		4>9	1	2	0.03041	0.44	2.25
SC_00054546.1.1	OB-fold nucleic acid binding domain-containing protein	9,30E-42	77.33%		9>4	8	2	0.02469	3.35	0.30
SC_00004702.1.1	octicosapeptide/Phox/Bem1p domain-containing protein	2,50E-32	67.1%		9>4	8	3	0.00218	2.40	0.42
SC_00001587.1.1	Oleosin-like protein	4,50E-18	65.13%		9>4	36	11	0.02440	3.17	0.32
SC_00025127.1.1	Oleosin-like protein	3,80E-19	63.44%		9>4	12	5	0.03591	2.51	0.40
SC_LOC100191494_MAIZE.2.2	oligopeptide transporter 7	1,30E-62	63.13%		4>9	9	18	0.04672	0.50	2.00
SC_00014965.1.1	omega-hydroxypalmitate O-feruloyl transferase	1,40E-35	46.75%	EC:2.3.1.0	9>4	6	2	0.03207	3.34	0.30
SC_00094725.1.1	OPC-8:0 CoA ligase1	1,80E-24	71.79%	EC:2.3.1.86; EC:6.2.1.12	9>4	5	2	0.02426	2.16	0.46
SC_00031918.1.1	outer membrane lipoprotein Blc	3,20E-54	62.5%		9>4	8	3	0.03679	2.24	0.45
SC_00027924.1.1	paired amphipathic helix protein Sin3-like 2	7,10E-47	64.45%		4>9	2	5	0.04275	0.49	2.06
SC_ZEAM_MAIZE.2.4	pathogenesis-related thaumatin-like protein	4,30E-31	61.55%		9>4	4	2	0.01339	2.59	0.39
SC_00009487.1.1	pathogenesis-related thaumatin-like protein	6,60E-64	59.25%		9>4	6	2	0.04254	2.36	0.42
SC_PP354_ARATH.1.2	pentatricopeptide repeat-containing protein	3,60E-49	68.1%		4>9	1	2	0.00075	0.43	2.33
SC_00035932.1.1	pentatricopeptide repeat-containing protein	1,10E-125	59.55%		4>9	1	2	0.00844	0.45	2.23
SC_00054770.1.1	pentatricopeptide repeat-containing protein	1,20E-75	47.0%		4>9	3	7	0.00901	0.48	2.10
SC_00127608.1.1	pentatricopeptide repeat-containing protein	1,70E-01	56.0%		4>9	1	2	0.01439	0.50	2.01
SC_PPR53_ARATH.4.9	pentatricopeptide repeat-containing protein	2,80E-39	56.6%		4>9	1	2	0.01986	0.48	2.07
SC_00068057.1.1	pentatricopeptide repeat-containing protein	3,70E-27	60.25%		4>9	1	3	0.03572	0.40	2.52
SC_00058851.1.1	pentatricopeptide repeat-containing protein	2,60E-19	52.44%		4>9	2	4	0.03888	0.49	2.03
SC_00111338.1.1	pentatricopeptide repeat-containing protein	2,70E-01	57.0%		4>9	2	4	0.04257	0.48	2.08
SC_00096575.1.1	pentatricopeptide repeat-containing protein	5,70E-85	72.2%		4>9	3	5	0.04474	0.50	2.01
SC_PP258_ARATH.1.9	pentatricopeptide repeat-containing protein	7,70E-35	67.0%		4>9	1	2	0.04692	0.47	2.14
SC_00035126.1.1	peptide alpha-N-acetyltransferase	2,80E-08	65.0%		9>4	8	3	0.02733	2.21	0.45
SC_MSRB1_ORYSJ.2.5	peptide methionine sulfoxide reductase B1	4,00E-20	61.77%	EC:1.8.4.12	9>4	4	2	0.02060	2.02	0.49
SC_AT1G26550_ARATH.1.1	peptidyl-prolyl cis-trans isomerase NIMA-interacting 4	2,70E-26	64.67%		9>4	7	3	0.02099	2.48	0.40
SC_PER12_ARATH.2.14	peroxidase 12	8,60E-77	58.1%		9>4	10	3	0.02309	3.42	0.29
SC_00055627.1.1	peroxidase 41	3,10E-63	67.7%		9>4	15	5	0.00043	2.85	0.35
SC_00077646.1.1	peroxidase 57	1,70E-21	66.35%		9>4	19	3	0.02723	5.45	0.18
SC_00024915.1.1	peroxin 14	3,40E-06	86.0%		4>9	3	6	0.01073	0.43	2.32
SC_00003923.1.1	peroxiredoxin (PRX)-like 2 domain-containing protein	9,60E-98	65.5%		9>4	3	1	0.01387	2.19	0.46
SC_PRX2C_ORYSJ.2.15	peroxiredoxin-2C	2,20E-29	76.33%		9>4	3	1	0.04615	2.62	0.38
SC_00008655.1.1	phosphate starvation-induced protein 2	3,60E-91	72.5%	EC:3.1.3.0; EC:3.6.1.1	9>4	5	1	0.04025	4.92	0.20
SC_00111912.1.1	phosphatidylinositol glycan, class V	2,10E-33	60.0%		4>9	2	5	0.02446	0.44	2.27
SC_00012584.1.1	phosphoglycerate kinase	3,40E-37	49.5%		4>9	158	331	0.00787	0.48	2.09
SC_00076737.1.1	phosphoglycerate kinase	3,20E-21	64.5%		4>9	7	25	0.00804	0.29	3.47
SC_00013921.1.1	phosphoglycerate kinase	2,30E-07	56.2%		4>9	10	31	0.01206	0.32	3.11
SC_00031404.1.1	phosphoglycerate kinase	7,80E-23	57.0%		4>9	51	127	0.01906	0.40	2.52

SC_00024181.1.1	phosphoglycerate kinase	2,10E-09	60.25%		4>9	11	41	0.02110	0.26	3.91
SC_00160794.1.1	phosphoglycerate kinase	5,00E-19	47.5%		4>9	10	34	0.03623	0.28	3.60
SC_00056329.1.1	phosphoglycerate kinase	9,80E-23	57.0%		4>9	8	34	0.04116	0.25	4.08
SC_00002182.1.1	phosphoglycerate kinase 1	5,00E-06	43.0%		4>9	81	169	0.01493	0.48	2.09
SC_00056376.1.1	phosphoglycerate kinase 1	1,10E-07	44.5%		4>9	34	77	0.01923	0.44	2.27
SC_LOC100284361_MAIZE.1.1	Phospholipase A2-like protein	6,60E-33	59.8%		9>4	18	9	0.03077	2.04	0.49
SC_OS03G0852800_ORYSJ.1.1	phospholipase C	1,40E-25	58.0%		9>4	8	2	0.00310	3.35	0.30
SC_00122309.1.1	phospholipase C	1,90E-67	65.0%		9>4	12	4	0.01523	3.34	0.30
SC_00122755.1.1	photosystem II 47 kDa protein	1,20E-02	43.0%		4>9	3	7	0.04657	0.38	2.63
SC_CB21_PINTH.29.69	photosystem II light harvesting complex protein 2.3	6,80E-87	70.5%		9>4	9	4	0.03940	2.35	0.43
SC_00040154.1.1	phragmoplast-associated kinesin-related protein 2	1,50E-19	56.0%		4>9	1	3	0.00889	0.44	2.26
SC_00081116.1.1	phragmoplastin interacting protein 1	5,10E-01	51.0%		4>9	1	4	0.01795	0.29	3.50
SC_00054390.1.1	phragmoplastin interacting protein 1	2,80E-02	55.0%		4>9	6	17	0.04226	0.33	3.07
SC_MCAC1_ARATH.2.2	PLAC8 domain-containing protein	2,80E-49	77.5%		4>9	1	4	0.02363	0.34	2.91
SC_CNR8_MAIZE.1.2	PLAC8 domain-containing protein	1,10E-56	59.0%		9>4	22	10	0.02734	2.18	0.46
SC_00018385.1.1	PLAC8 domain-containing protein	8,10E-49	76.0%		9>4	43	20	0.02769	2.13	0.47
SC_PCR6_ARATH.2.12	PLAC8 family protein	3,00E-34	63.15%		9>4	35	14	0.03478	2.47	0.41
SC_00137006.1.1	plant invertase/pectin methylesterase inhibitor domain-containing protein	2,90E-09	63.69%		9>4	20	4	0.02463	5.02	0.20
SC_00084238.1.1	plant invertase/pectin methylesterase inhibitor domain-containing protein	2,80E-18	55.6%		9>4	10	4	0.02610	2.38	0.42
SC_LOC100502337_MAIZE.1.1	plant invertase/pectin methylesterase inhibitor domain-containing protein	1,60E-02	66.0%		9>4	46	20	0.03395	2.35	0.43
SC_00054318.1.1	Plant Tudor-like RNA-binding protein	6,70E-05	60.0%		9>4	7	3	0.02286	2.26	0.44
SC_OLEE1_BETVE.1.1	pollen Ole e 1 allergen and extensin family protein	3,00E-08	47.0%		9>4	230	97	0.00478	2.37	0.42
SC_00039056.1.1	polypyrimidine tract-binding protein 3	5,90E-35	62.5%		4>9	1	3	0.02107	0.46	2.15
SC_00041668.1.1	POZ/BTB containin G-protein 1	3,30E-22	55.25%		4>9	4	9	0.01927	0.44	2.27
SC_PCO141631_MAIZE.1.2	PPPDE putative thiol peptidase family protein	3,90E-61	51.8%		9>4	11	3	0.01150	3.36	0.30
SC_00030197.1.1	Pre-mRNA-processing factor 19	8,60E-39	55.9%		4>9	2	4	0.04328	0.49	2.05
SC_00014809.1.1	pre-mRNA-splicing factor 38A	2,40E-24	70.0%		9>4	3	1	0.03096	2.03	0.49
SC_00061443.1.1	Pre-mRNA-splicing factor SF2	2,30E-44	64.22%		9>4	3	1	0.02756	2.43	0.41
SC_00014618.1.1	prephenate dehydrogenase family protein	7,60E-133	66.33%		9>4	5	3	0.00511	2.06	0.49
SC_00029298.1.1	Presequence protease 1	1,20E-53	85.0%		4>9	3	6	0.00525	0.49	2.06
SC_00015260.1.1	proline transporter 2	1,40E-132	51.55%		9>4	18	6	0.00664	3.33	0.30
SC_PSA2A_ARATH.2.2	proteasome subunit alpha type-2-A	7,20E-50	64.71%		4>9	6	12	0.00572	0.49	2.03
SC_PSA2B_ARATH.5.5	proteasome subunit alpha type-2-B	1,70E-74	58.2%		4>9	3	7	0.03206	0.44	2.29
SC_APRL2_ORYSJ.2.2	protein alfin-like 5	2,10E-08	55.0%		4>9	1	2	0.01673	0.43	2.31
SC_00041319.1.1	protein BCCIP-like protein	4,70E-20	67.5%		4>9	1	2	0.03027	0.44	2.29
SC_BRK1_ARATH.1.2	protein BRICK 1	4,30E-31	83.0%		9>4	86	43	0.01948	2.03	0.49
SC_00063786.1.1	protein BRICK 1	3,40E-31	91.0%		9>4	6	2	0.03319	3.74	0.27
SC_00006425.1.1	protein embryo defective 2752	1,60E-32	81.0%		4>9	2	5	0.02372	0.50	2.01
SC_00037132.1.1	protein EXORDIUM like 2	7,80E-61	65.63%		9>4	6	1	0.04734	4.22	0.24

SC_00020031.1.1	protein FAB1D	1,40E-01	45.0%		9>4	7	3	0.00759	2.24	0.45
SC_UMC1467_MAIZE.1.1	protein inositol requiring 1-1	3,60E-02	60.75%		9>4	20	9	0.02947	2.13	0.47
SC_00024418.1.1	protein kinase-like protein	4,10E-114	71.15%	EC:2.7.11.0	4>9	1	1	0.02194	0.40	2.49
SC_00161926.1.1	protein little nuclei1	6,00E-01	63.0%		4>9	2	6	0.04817	0.41	2.46
SC_00008510.1.1	protein maternal effect embryo arrest 59	1,30E-25	56.0%		4>9	2	6	0.00828	0.31	3.23
SC_P2C08_ORYSJ.7.9	protein phosphatase 2C 16	7,30E-20	62.45%		9>4	14	5	0.01297	3.02	0.33
SC_P2C08_ORYSJ.2.9	protein phosphatase 2C 16	4,20E-23	60.35%		9>4	43	19	0.01486	2.22	0.45
SC_00155910.1.1	protein PHYTOALEXIN DEFICIENT 4	1,40E-21	58.0%		4>9	2	5	0.00655	0.38	2.62
SC_00001008.1.1	protein PHYTOALEXIN DEFICIENT 4	1,60E-62	46.6%		9>4	2	1	0.04062	2.35	0.43
SC_00011745.1.1	protein SAH7	8,60E-20	48.33%		9>4	125	59	0.03205	2.13	0.47
SC_CHIB_VITVI.2.3	protein SCAR4	7,10E-01	52.0%		9>4	122	41	0.03992	2.96	0.34
SC_00003107.1.1	protein SPIRAL1-like2	1,30E-10	85.0%		9>4	25	10	0.01781	2.51	0.40
SC_00134731.1.1	protein SPIRAL1-like2	1,60E-11	81.5%		9>4	33	11	0.02063	3.14	0.32
SC_OS03G0417800_ORYSJ.1.6	protein SPIRAL1-like2	9,00E-10	85.0%		9>4	38	19	0.02173	2.03	0.49
SC_00150816.1.1	protein SPIRAL1-like2	1,70E-10	85.0%		9>4	61	25	0.03303	2.40	0.42
SC_00089066.1.1	protein SPIRAL1-like2	1,10E-10	85.0%		9>4	46	20	0.03803	2.31	0.43
SC_00000089.1.1	protein SPIRAL1-like2	1,40E-09	85.0%		9>4	85	39	0.03806	2.15	0.47
SC_OS03G0417800_ORYSJ.2.6	protein SPIRAL1-like2	9,70E-16	89.5%		9>4	9	4	0.04819	2.11	0.47
SC_00012175.1.1	protein SUPPRESSOR OF GENE SILENCING 3	3,00E-121	53.33%		4>9	1	1	0.02327	0.45	2.23
SC_00039577.1.1	protein trichome birefringence-like 36	1,40E-85	60.4%		9>4	40	15	0.00099	2.60	0.38
SC_PP344_ARATH.3.4	proton gradient regulation 3	5,30E-50	59.65%		4>9	3	9	0.03709	0.35	2.83
SC_PPA27_ARATH.1.2	purple acid phosphatase 27	0,00E+00	48.9%		9>4	8	4	0.00558	2.19	0.46
SC_00003596.1.1	putative adenine nucleotide alpha hydrolase domain-containing universal stress protein	1,80E-40	54.3%		9>4	16	6	0.01298	2.60	0.38
SC_00010664.1.1	putative AT-hook DNA-binding protein	5,80E-66	73.3%		9>4	30	11	0.01431	2.64	0.38
SC_00026249.1.1	putative ATP synthase subunit	2,80E-24	65.0%		4>9	1	3	0.00432	0.47	2.12
SC_POLC2_JUNOX.18.31	putative calcium-binding protein CML25	3,60E-36	60.55%		9>4	14	5	0.00132	2.92	0.34
SC_CML10_ORYSJ.2.26	putative calcium-binding protein CML25	4,80E-30	62.7%		9>4	12	5	0.01351	2.38	0.42
SC_POLC2_JUNOX.2.31	putative calcium-binding protein CML25	5,70E-40	58.1%		9>4	45	20	0.02673	2.30	0.44
SC_CHI4_ORYSJ.1.7	putative chitinase	3,60E-47	50.24%		9>4	25	4	0.02240	6.60	0.15
SC_CYSP_HEMSP.3.4	putative cysteine proteinase	4,00E-79	70.35%		9>4	5	2	0.00662	2.05	0.49
SC_RD19A_ARATH.8.18	putative cysteine proteinase A494	2,70E-98	58.0%		9>4	7	4	0.01066	2.02	0.49
SC_00001272.1.1	putative E3 ubiquitin-protein ligase BAH1-like protein	6,00E-97	52.42%		9>4	22	9	0.04858	2.30	0.43
SC_00027351.1.1	putative epoxide hydrolase	1,00E-09	51.83%		9>4	4	2	0.03593	2.71	0.37
SC_RFS1_ARATH.5.8	putative galactinol--sucrose galactosyltransferase 1	1,80E-19	59.63%	EC:2.4.1.82; EC:3.2.1.22	9>4	5	2	0.00563	2.43	0.41
SC_RFS2_ARATH.16.16	putative galactinol--sucrose galactosyltransferase 2	1,20E-40	51.67%		9>4	3	1	0.01114	2.59	0.39
SC_GL91_ORYSJ.4.4	putative germin-like protein subfamily 1 member 9	1,80E-10	61.1%		9>4	10	2	0.00354	4.19	0.24
SC_00053117.1.1	putative gibberellin-regulated protein	1,10E-32	60.12%		9>4	112	43	0.00320	2.58	0.39
SC_AT3G10185_ARATH.3.3	putative gibberellin-regulated protein	4,40E-26	66.71%		9>4	57	19	0.00692	2.95	0.34
SC_GASA2_ARATH.5.10	putative gibberellin-regulated protein	6,20E-23	59.29%		4>9	10	32	0.00727	0.30	3.33

SC_00030584.1.1	putative gibberellin-regulated protein	8,20E-18	56.12%		9>4	23	11	0.01639	2.09	0.48
SC_00151200.1.1	putative gibberellin-regulated protein	3,10E-30	62.59%		9>4	21	8	0.04296	2.65	0.38
SC_PPA1_ARATH.2.2	putative inactive purple acid phosphatase 1	2,70E-62	73.0%		9>4	6	3	0.04397	2.36	0.42
SC_00133290.1.1	putative LRR receptor-like serine/threonine-protein kinase	1,30E-21	55.95%		4>9	2	6	0.04887	0.35	2.88
SC_00133959.1.1	putative methyltransferase PMT17	1,30E-70	62.95%		4>9	2	4	0.03154	0.47	2.14
SC_PMT8_ARATH.1.4	putative methyltransferase PMT9	9,40E-23	52.8%		4>9	1	2	0.02943	0.47	2.15
SC_00017143.1.1	putative NADPH quinone oxidoreductase	6,30E-32	62.8%		9>4	8	3	0.01191	2.37	0.42
SC_PEL8_ARATH.5.6	putative pectate lyase 8	1,00E-94	74.6%		4>9	4	10	0.04897	0.46	2.19
SC_00037025.1.1	putative pectinesterase 68	1,70E-53	56.0%		9>4	4	2	0.01518	2.31	0.43
SC_00145385.1.1	putative peptide/nitrate transporter	2,20E-31	65.45%		9>4	15	6	0.01035	2.34	0.43
SC_00030820.1.1	putative phosphatidylethanolamine-binding protein	2,80E-33	63.67%		9>4	11	5	0.04133	2.44	0.41
SC_00030895.1.1	putative polyamine oxidase 5	4,60E-21	61.75%		9>4	62	31	0.01349	2.01	0.50
SC_OS04G0678900_ORYSJ.1.1	putative polyol transporter 4	4,00E-13	68.35%		4>9	1	3	0.04320	0.45	2.22
SC_00042371.1.1	putative protein Pop3	2,70E-25	60.8%		9>4	56	27	0.02437	2.11	0.47
SC_00125177.1.1	putative receptor protein kinase	1,20E-87	63.95%		4>9	4	9	0.04508	0.48	2.07
SC_00078278.1.1	putative receptor serine/threonine kinase	3,30E-01	48.0%		4>9	2	3	0.01273	0.45	2.22
SC_00003633.1.1	putative RNA helicase	2,80E-02	68.0%		9>4	6	3	0.01357	2.09	0.48
SC_PTI1_SOLLC.4.6	putative serine/threonine protein kinase	3,30E-114	72.85%		4>9	1	3	0.02474	0.45	2.24
SC_00012505.1.1	putative serine/threonine-protein kinase	4,10E-01	56.0%		9>4	3	1	0.00919	2.57	0.39
SC_00005583.1.1	putative S-ribonuclease binding protein SBP1	1,80E-37	57.6%		9>4	12	5	0.02952	2.41	0.41
SC_00009446.1.1	putative TIR-NBS-LRR class disease resistance protein	1,20E-14	48.3%		9>4	10	3	0.03743	3.68	0.27
SC_00098360.1.1	putative TIR-NBS-LRR class disease resistance protein	5,70E-12	50.45%		9>4	123	46	0.04290	2.66	0.38
SC_00150529.1.1	putative type 1 membrane protein	1,30E-15	71.0%		9>4	3	1	0.03214	2.07	0.48
SC_00090072.1.1	putative type 1 membrane protein	5,70E-04	73.0%		9>4	8	2	0.04981	3.81	0.26
SC_00040622.1.1	putative ubiquitin-conjugating enzyme E2 18	3,60E-35	65.63%	EC:6.3.2.19	9>4	4	2	0.01132	2.36	0.42
SC_00022801.1.1	putative wall-associated receptor kinase-like 13	2,90E-01	43.0%		9>4	7	2	0.04273	2.79	0.36
SC_00015833.1.1	putative WRKY transcription factor 42	1,70E-36	67.7%		9>4	2	1	0.02362	2.61	0.38
SC_00009805.1.1	Pyridoxal phosphate -dependent transferase-like protein	8,20E-01	48.0%		9>4	2	1	0.04052	2.04	0.49
SC_00162587.1.1	pyrophosphate-energized membrane proton pump 2	8,00E-02	60.0%		4>9	30	89	0.02436	0.34	2.93
SC_00025912.1.1	Pyrophosphate-energized vacuolar membrane proton pump 1	2,20E-88	76.0%		4>9	2	4	0.01079	0.50	2.01
SC_AVP1_ARATH.11.38	Pyrophosphate-energized vacuolar membrane proton pump 1	4,70E-62	76.0%		4>9	5	12	0.01557	0.45	2.21
SC_PDC2_PEA.10.21	pyruvate decarboxylase	1,20E-132	76.83%	EC:4.1.1.0	4>9	2	14	0.00029	0.14	7.37
SC_PDC2_ORYSI.17.19	pyruvate decarboxylase	1,30E-123	74.5%		4>9	1	8	0.00534	0.17	5.98
SC_PDC1_ORYSI.1.2	pyruvate decarboxylase	6,00E-59	81.5%	EC:4.1.1.0	4>9	9	22	0.00700	0.42	2.37
SC_00159905.1.1	pyruvate decarboxylase	5,90E-41	75.0%		4>9	13	30	0.02256	0.43	2.35

SC_PDC2_ORYSI.16.19	pyruvate decarboxylase	2,00E-49	90.25%	EC:4.1.1.0	4>9	2	4	0.04191	0.44	2.27
SC_PDC2_PEA.15.21	pyruvate decarboxylase	6,00E-90	76.0%	EC:4.1.1.0	4>9	5	11	0.04192	0.45	2.21
SC_PDC2_ORYSI.4.19	pyruvate decarboxylase	2,70E-133	71.4%		4>9	7	16	0.04573	0.45	2.22
SC_PDC2_ORYSI.1.19	pyruvate decarboxylase-2	0,00E+00	66.33%		4>9	6	14	0.00616	0.44	2.29
SC_PDC2_PEA.5.21	pyruvate decarboxylase-2	2,90E-129	63.83%		4>9	4	12	0.00909	0.33	3.07
SC_PDC2_PEA.20.21	pyruvate decarboxylase-2	5,40E-84	76.67%		4>9	2	5	0.03506	0.44	2.27
SC_PDC2_ORYSI.10.19	pyruvate decarboxylase-2	1,40E-42	73.4%	EC:4.1.1.0	4>9	1	5	0.04117	0.24	4.16
SC_OS11G0148500_ORYSJ.1.4	pyruvate kinase	0,00E+00	61.27%		4>9	1	4	0.02809	0.30	3.28
SC_00049136.1.1	pyruvate kinase-like protein	2,00E-50	67.8%		4>9	8	17	0.00217	0.47	2.12
SC_00091862.1.1	RabGAP/TBC domain-containing protein	2,20E-64	65.45%		4>9	1	2	0.04400	0.49	2.04
SC_00144612.1.1	RabGAP/TBC domain-containing protein	1,30E-97	58.85%		9>4	4	2	0.04918	2.32	0.43
SC_00015729.1.1	receptor lectin kinase	5,70E-05	46.75%		9>4	5	2	0.01201	2.98	0.34
SC_00073038.1.1	receptor like protein 15	5,30E-01	53.0%		4>9	2	5	0.04690	0.36	2.79
SC_HSL1_ARATH.14.17	receptor-like protein kinase HSL1	3,20E-40	68.8%		4>9	1	3	0.04613	0.48	2.09
SC_00061865.1.1	regulator of chromosome condensation repeat-containing protein	1,10E-22	79.0%		9>4	25	12	0.01343	2.07	0.48
SC_OS01G0835500_ORYSJ.1.2	respiratory burst oxidase	0,00E+00	56.1%	EC:1.6.3.0	4>9	1	1	0.03006	0.47	2.15
SC_UPTG_MAIZE.10.12	reversibly glycosylated polypeptide 1	6,10E-83	72.33%	EC:5.4.99.30	9>4	3	1	0.01864	2.09	0.48
SC_UPTG2_SOLTU.36.38	reversibly glycosylated polypeptide 1	1,60E-103	89.0%	EC:5.4.99.30; EC:2.4.1.12	9>4	4	2	0.04588	2.22	0.45
SC_UPTG_MAIZE.2.12	reversibly glycosylated polypeptide 2	3,80E-25	58.4%		9>4	3	1	0.01283	2.59	0.39
SC_00112388.1.1	reversibly glycosylated polypeptide 2	4,30E-29	73.2%	EC:5.4.99.30	9>4	14	6	0.03572	2.53	0.39
SC_00008208.1.1	ribokinase	5,40E-03	43.0%		9>4	8	4	0.02995	2.05	0.49
SC_RIR2_TOBAC.2.13	Ribonucleoside-diphosphate reductase small chain C	1,70E-144	84.5%	EC:1.17.4.1	4>9	13	27	0.01240	0.48	2.06
SC_00079546.1.1	ribophorin II (RPN2) family protein	2,90E-22	63.0%		4>9	2	5	0.01141	0.43	2.33
SC_RR7_PINTH.4.5	ribosomal protein S7	1,30E-39	67.44%		9>4	128	58	0.04741	2.19	0.46
SC_00067000.1.1	ring finger and CHY zinc finger domain-containing protein 1	3,10E-02	60.0%		9>4	26	7	0.00415	3.52	0.28
SC_00155753.1.1	ring finger and CHY zinc finger domain-containing protein 1	8,90E-02	56.0%		9>4	9	4	0.02683	2.36	0.42
SC_00048061.1.1	RING/U-box domain-containing protein	4,10E-04	52.0%		4>9	1	2	0.00069	0.43	2.33
SC_00032163.1.1	RING/U-box domain-containing protein	1,30E-29	56.9%		9>4	4	2	0.04234	2.79	0.36
SC_00012143.1.1	RING/U-box protein with C6HC-type zinc finger	2,50E-01	56.0%		4>9	4	9	0.01250	0.48	2.09
SC_ATL52_ARATH.1.1	RING-H2 finger protein ATL52	7,20E-31	63.0%		9>4	3	1	0.01782	2.29	0.44
SC_00011399.1.1	RmlC-like cupin	1,80E-33	52.22%		9>4	3	1	0.00224	2.27	0.44
SC_00008432.1.1	RmlC-like cupin	2,70E-50	56.4%		9>4	16	6	0.00695	2.54	0.39
SC_00121599.1.1	RmlC-like cupin	4,80E-15	62.0%		9>4	6	3	0.00914	2.29	0.44
SC_00037227.1.1	RNA polymerase I-associated factor PAF67	5,50E-60	69.0%		4>9	4	8	0.02023	0.48	2.08
SC_00146154.1.1	RNA recognition motif and CCHC-type zinc finger domain-containing protein	1,10E-10	58.33%		9>4	4	2	0.00554	2.34	0.43
SC_00033353.1.1	RNA-directed DNA polymerase (reverse transcriptase)-related family protein	7,30E-07	42.29%		9>4	4	2	0.00898	2.45	0.41

SC_00087183.1.1	RNase H domain-containing protein	6,00E-11	47.0%		4>9	1	5	0.00876	0.26	3.80
SC_00034067.1.1	Rossmann-fold NAD(P)-binding domain-containing protein	3,50E-33	84.0%		4>9	3	6	0.01320	0.47	2.14
SC_00074040.1.1	Rossmann-fold NAD(P)-binding domain-containing protein	1,70E-22	52.12%		4>9	1	2	0.02605	0.42	2.38
SC_DFRA_CALCH.3.13	Rossmann-fold NAD(P)-binding domain-containing protein	1,20E-34	64.8%		4>9	3	7	0.04087	0.48	2.07
SC_00057838.1.1	Rossmann-fold NAD(P)-binding domain-containing protein	1,50E-73	59.1%		4>9	14	27	0.04612	0.50	2.01
SC_00010473.1.1	rubredoxin-like protein	4,00E-27	80.0%		9>4	8	3	0.01808	2.66	0.38
SC_00007904.1.1	S-adenosyl-L-methionine-dependent methyltransferase-like protein	5,90E-89	56.43%		9>4	88	44	0.00247	2.00	0.50
SC_LOC100282379_MAIZE.1.1	SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein	7,60E-64	64.33%		9>4	2	1	0.02439	2.32	0.43
SC_00095086.1.1	sec24-like transport protein	5,60E-57	61.83%		4>9	2	5	0.02998	0.49	2.03
SC_00082390.1.1	SecY protein transport family protein	5,50E-69	85.0%		9>4	4	2	0.02147	2.28	0.44
SC_00011034.1.1	Selenoprotein, Rdx type	5,50E-23	80.33%		4>9	2	5	0.02094	0.46	2.18
SC_00000675.1.1	senescence-related protein	1,60E-58	61.2%		9>4	4	1	0.01457	3.28	0.31
SC_PCO140607_MAIZE.1.1	sequence-specific DNA binding transcription factor	4,30E-33	65.6%		9>4	44	21	0.01376	2.06	0.48
SC_CYP306A1_BOMMO.1281.3286	sequence-specific DNA binding transcription factor	4,20E-04	61.33%		9>4	29	11	0.01665	2.53	0.40
SC_00006288.1.1	sequence-specific DNA binding transcription factor	4,30E-40	62.31%		9>4	3	1	0.02043	2.07	0.48
SC_00010551.1.1	sequence-specific DNA binding transcription factor	1,00E-43	68.0%		4>9	3	8	0.02127	0.42	2.38
SC_SCP45_ARATH.2.9	serine carboxypeptidase-like 45	3,50E-98	62.75%		9>4	4	1	0.01682	3.05	0.33
SC_00013598.1.1	serine/threonine kinase	5,30E-54	69.75%		9>4	16	8	0.02293	2.13	0.47
SC_KPK1_ARATH.1.1	serine/threonine protein kinase 2	2,30E-130	58.85%	EC:2.7.11.0	9>4	8	3	0.00019	2.22	0.45
SC_PCO100533A_MAIZE.1.1	serine/threonine protein kinase 3	4,90E-56	48.8%		4>9	4	9	0.03246	0.47	2.14
SC_SRK2E_ARATH.3.3	serine/threonine-protein kinase SRK2E	2,30E-51	76.05%	EC:2.7.11.0	9>4	29	5	0.02215	5.42	0.18
SC_00012518.1.1	serine/threonine-protein kinase SRK2E	2,40E-72	59.15%		9>4	59	11	0.03121	5.50	0.18
SC_SAPK1_ORYSJ.1.1	serine/threonine-protein kinase SRK2E	2,20E-91	77.3%		9>4	8	2	0.04936	4.46	0.22
SC_SRK2E_ARATH.1.3	serine/threonine-protein kinase SRK2H	5,60E-137	75.45%	EC:2.7.11.0	9>4	55	12	0.04754	4.46	0.22
SC_PP2A_TOBAC.4.4	serine/threonine-protein phosphatase PP2A-3 catalytic subunit	3,80E-66	77.4%		9>4	11	5	0.01049	2.31	0.43
SC_00020664.1.1	serine-rich protein-like protein	7,50E-06	54.0%		9>4	6	3	0.02120	2.16	0.46
SC_CYP306A1_BOMMO.1306.3286	serine-type endopeptidase inhibitor	5,80E-04	54.0%		9>4	9	2	0.01234	3.45	0.29
SC_00021442.1.1	SET domain-containing protein	3,90E-119	48.08%		4>9	0	2	0.01566	0.27	3.72
SC_00151339.1.1	S-formylglutathione hydrolase	5,00E-24	80.0%	EC:3.1.2.12	4>9	1	3	0.02998	0.46	2.18
SC_00158728.1.1	S-formylglutathione hydrolase	8,40E-52	82.0%	EC:3.1.2.12	4>9	1	3	0.03525	0.37	2.73
SC_CYP306A1_BOMMO.2595.3286	SKU5 similar 6	3,10E-03	52.0%		4>9	2	3	0.01687	0.45	2.23
SC_00011238.1.1	SKU5 similar 8 protein	6,00E-17	68.94%		9>4	176	68	0.00393	2.60	0.38

SC_00004873.1.1	small nuclear ribonucleoprotein associated protein B	3,10E-34	68.13%		9>4	10	4	0.03289	2.63	0.38
SC_00031832.1.1	small RNA degrading nuclease 3	9,30E-01	55.0%		4>9	6	15	0.01663	0.39	2.54
SC_SPA3_ARATH.2.2	SPA1-related 4 protein	3,30E-43	56.6%		4>9	1	3	0.01229	0.46	2.19
SC_00134500.1.1	SPX domain-containing protein 4	5,50E-05	57.0%		9>4	6	2	0.01016	2.76	0.36
SC_00016812.1.1	SPX domain-containing protein 4	1,40E-09	69.33%		9>4	4	1	0.01317	3.04	0.33
SC_00029068.1.1	strictosidine synthase-like 4 protein	1,70E-01	63.0%		9>4	20	6	0.00074	3.32	0.30
SC_SUS2_PEA.13.43	sucrose synthase 3	0,00E+00	57.06%	EC:2.4.1.13	4>9	3	10	0.00210	0.33	3.05
SC_STP7_ARATH.4.14	sugar transport protein 7	0,00E+00	65.75%		9>4	17	4	0.01275	3.84	0.26
SC_00046691.1.1	Sulfite exporter TauE/SafE family protein	2,70E-53	65.67%		4>9	2	5	0.03738	0.43	2.31
SC_00093956.1.1	sulfite oxidase	1,30E-95	65.8%	EC:1.8.3.1	4>9	1	4	0.03785	0.35	2.86
SC_UBC8_ARATH.17.24	SUMO-conjugating enzyme UBC9	1,30E-99	82.15%	EC:6.3.2.19	9>4	48	19	0.00379	2.50	0.40
SC_UBC8_ARATH.1.24	SUMO-conjugating enzyme UBC9	1,70E-95	81.85%	EC:6.3.2.19	9>4	24	12	0.01136	2.03	0.49
SC_UBC8_ARATH.4.24	SUMO-conjugating enzyme UBC9	3,10E-103	82.55%	EC:6.3.2.19	9>4	33	16	0.01603	2.11	0.47
SC_00000189.1.1	SUMO-conjugating enzyme UBC9	4,80E-97	82.15%	EC:6.3.2.19	9>4	19	8	0.03139	2.33	0.43
SC_00011128.1.1	Surfeit locus protein 6	8,10E-12	69.0%		9>4	3	1	0.04628	2.61	0.38
SC_00090605.1.1	synaptobrevin-related family protein	3,40E-07	59.43%		4>9	1	4	0.01732	0.42	2.40
SC_SY132_ARATH.1.3	syntaxin-132	1,20E-09	74.4%		9>4	6	3	0.00610	2.28	0.44
SC_00161130.1.1	syntaxin-52	2,60E-21	71.67%		9>4	8	3	0.04139	2.47	0.41
SC_00000804.1.1	TBP-associated factor 15B	1,50E-47	48.65%		9>4	1	0	0.00318	2.77	0.36
SC_TCPE_ARATH.2.12	T-complex protein 1 subunit epsilon	1,60E-91	53.33%		4>9	26	53	0.00267	0.50	2.01
SC_TCPE_ARATH.3.12	T-complex protein 1 subunit epsilon	4,00E-88	54.09%		4>9	15	35	0.00533	0.44	2.25
SC_PCO099347_MAIZE.1.1	T-complex protein 1 subunit epsilon	4,30E-54	64.83%		4>9	29	68	0.02662	0.43	2.32
SC_00013455.1.1	T-complex protein 1 subunit epsilon	2,20E-29	74.13%		4>9	8	16	0.03980	0.49	2.03
SC_00098212.1.1	T-complex protein 11	2,20E-13	58.0%		9>4	15	7	0.00158	2.11	0.47
SC_00014534.1.1	TCP-1/cpn60 chaperonin family protein	3,10E-37	58.6%		9>4	5	2	0.02568	2.27	0.44
SC_00070045.1.1	Telomerase activating protein Est1	2,10E-02	57.5%		4>9	1	2	0.00715	0.37	2.73
SC_00144986.1.1	thioesterase-like protein	1,90E-14	75.5%		9>4	6	3	0.01045	2.05	0.49
SC_TRL12_ORYSJ.7.7	thioredoxin-like 1-1	3,90E-44	63.0%	EC:1.8.4.0	4>9	2	5	0.00153	0.30	3.38
SC_OS12G0608600_ORYSJ.1.2	thiosulfate sulfurtransferase	1,50E-128	83.8%	EC:2.8.1.1; EC:2.8.1.2	9>4	2	1	0.03980	2.00	0.50
SC_DRTS2_ARATH.3.5	thymidylate synthase	9,60E-99	73.0%		4>9	2	4	0.04918	0.49	2.05
SC_00126646.1.1	TIR-NBS class of disease resistance protein	5,90E-05	59.0%		4>9	3	11	0.01871	0.29	3.48
SC_LRC40_DANRE.1.2	TIR-NBS-LRR class disease resistance protein	7,20E-18	52.35%		9>4	3	1	0.02214	2.22	0.45
SC_00058122.1.1	TIR-NBS-LRR class disease resistance protein	1,20E-04	51.0%		9>4	3	1	0.04070	3.14	0.32
SC_TCMO_PEA.1.10	trans-cinnamate 4-monooxygenase	0,00E+00	49.4%	EC:1.14.13.11	9>4	4	2	0.02293	2.05	0.49
SC_00038410.1.1	transcription factor bHLH66	2,50E-27	82.8%		9>4	4	2	0.04520	2.04	0.49
SC_00020954.1.1	Transcription factor jumonji (jmjC) domain-containing protein	0,00E+00	53.4%		4>9	3	10	0.03959	0.30	3.28
SC_MYB08_ANTMA.4.12	transcription factor TRY	2,30E-16	77.9%		9>4	3	1	0.01162	2.31	0.43
SC_00098729.1.1	transcriptional coactivator/pterin dehydratase	5,20E-29	54.0%		9>4	3	2	0.03121	2.16	0.46
SC_00007365.1.1	transducin/WD40 domain-containing protein	2,50E-02	40.0%		9>4	66	26	0.01951	2.52	0.40
SC_IF2B_WHEAT.6.7	translation initiation factor eIF-2 beta subunit	1,20E-51	68.17%		9>4	4	2	0.00110	2.24	0.45
SC_EIF3D_ARATH.3.5	translation initiation factor eIF-3 subunit 7	8,80E-11	60.5%		4>9	6	12	0.01103	0.47	2.12

SC_00056484.1.1	translation initiation factor eIF-3 subunit 7	6,10E-90	84.5%		4>9	3	5	0.04384	0.46	2.15
SC_IF5Z_ARATH.3.6	translation initiation factor eIF-5	2,20E-16	61.33%		9>4	18	8	0.01537	2.29	0.44
SC_IF4A1_ORYSJ.11.11	translational initiation factor 4A-1	5,20E-100	60.95%	EC:3.6.1.15	4>9	2	4	0.02501	0.46	2.20
SC_00019542.1.1	transmembrane CLPTM1 family protein	1,60E-06	68.0%		9>4	7	3	0.01617	2.25	0.44
SC_00077418.1.1	trehalase 1	4,50E-17	49.0%		4>9	1	4	0.03123	0.39	2.57
SC_00025981.1.1	tRNA pseudouridine synthase A	3,00E-86	47.63%		4>9	2	4	0.04412	0.44	2.28
SC_00134932.1.1	tubulin beta-5 chain	7,50E-01	55.0%		4>9	17	73	0.02891	0.24	4.21
SC_00029167.1.1	two-component response regulator ARR5	1,10E-04	66.0%		4>9	2	4	0.02799	0.41	2.43
SC_00155605.1.1	ubiquitin 11	2,50E-84	79.95%		4>9	1	3	0.03842	0.39	2.58
SC_UBL5_ARATH.27.27	ubiquitin family protein	3,20E-45	62.0%		9>4	22	11	0.00275	2.11	0.47
SC_UBL5_ARATH.26.27	ubiquitin family protein	1,10E-44	82.0%		9>4	12	6	0.00959	2.00	0.50
SC_UBL5_ARATH.8.27	ubiquitin family protein	7,60E-44	84.0%		9>4	103	51	0.02509	2.02	0.50
SC_00155240.1.1	ubiquitin-conjugating enzyme E2 27	1,30E-22	61.7%	EC:6.3.2.19	9>4	4	1	0.00370	4.01	0.25
SC_00016734.1.1	ubiquitin-conjugating enzyme E2 28	6,10E-86	80.1%	EC:6.3.2.19	9>4	4	2	0.01460	2.48	0.40
SC_UBC30_ARATH.3.10	ubiquitin-conjugating enzyme E2 28	4,80E-66	79.35%	EC:6.3.2.19	9>4	87	36	0.01720	2.41	0.42
SC_UBC30_ARATH.10.10	ubiquitin-conjugating enzyme E2 30	4,00E-71	76.6%	EC:6.3.2.19	9>4	9	4	0.02425	2.36	0.42
SC_00034141.1.1	ubiquitin-conjugating enzyme E2 30	1,10E-64	76.6%	EC:6.3.2.19	9>4	49	15	0.02739	3.21	0.31
SC_UBC35_ARATH.9.14	ubiquitin-conjugating enzyme E2 35	1,10E-21	89.4%	EC:6.3.2.19	9>4	4	2	0.01756	2.15	0.47
SC_UBC9_ARATH.14.24	ubiquitin-conjugating enzyme E2 8	1,80E-57	82.2%	EC:6.3.2.19	9>4	8	4	0.00099	2.23	0.45
SC_UBC8_ARATH.3.24	ubiquitin-conjugating enzyme E2 8	3,40E-25	79.45%	EC:6.3.2.19	9>4	31	11	0.00114	2.94	0.34
SC_UBC8_ARATH.12.24	ubiquitin-conjugating enzyme E2 8	3,40E-48	71.85%	EC:6.3.2.19	9>4	15	6	0.00626	2.49	0.40
SC_UBC8_ARATH.6.24	ubiquitin-conjugating enzyme E2 8	6,20E-42	82.1%	EC:6.3.2.19	9>4	26	13	0.03091	2.00	0.50
SC_UBC8_ARATH.10.24	ubiquitin-conjugating enzyme E2 8	7,70E-62	75.75%	EC:6.3.2.19	9>4	14	7	0.03554	2.14	0.47
SC_UBC10_ARATH.2.4	ubiquitin-conjugating enzyme E2 8	1,80E-65	83.2%	EC:6.3.2.19	9>4	41	17	0.04183	2.36	0.42
SC_00013883.1.1	ubiquitin-conjugating enzyme E2 A	9,50E-64	56.75%	EC:6.3.2.19	9>4	3	1	0.04090	2.00	0.50
SC_00005776.1.1	U-box domain-containing protein 25	1,40E-54	56.7%		9>4	7	3	0.02828	2.21	0.45
SC_PUB35_ARATH.1.1	U-box domain-containing protein 35	1,50E-121	73.85%		9>4	1	1	0.04394	2.58	0.39
SC_LOC100285647_MAIZE.1.2	uclacyanin 1	1,70E-26	56.25%		9>4	10	2	0.01010	6.77	0.15
SC_00155485.1.1	UDP-glycosyltransferase family protein	6,60E-01	50.0%		9>4	68	34	0.01325	2.04	0.49
SC_00006900.1.1	UDP-glycosyltransferase-like protein	6,70E-57	48.45%		9>4	69	27	0.01316	2.60	0.38
SC_00146082.1.1	UDP-glycosyltransferase-like protein	6,70E-01	48.0%		4>9	5	18	0.02884	0.30	3.28
SC_00022565.1.1	uncharacterized protein	4,20E-67	71.5%		9>4	12	6	0.00068	2.06	0.49
SC_CYP306A1_BOMMO.930.3286	uncharacterized protein	2,80E-05	64.0%		9>4	163	67	0.00072	2.43	0.41
SC_00156207.1.1	uncharacterized protein	1,70E-03	54.0%		9>4	7	3	0.00109	2.56	0.39
SC_00003991.1.1	uncharacterized protein	2,90E-41	63.5%		9>4	3	1	0.00109	2.41	0.41
SC_00026715.1.1	uncharacterized protein	1,60E-33	58.8%		4>9	1	3	0.00234	0.43	2.33
SC_00015028.1.1	uncharacterized protein	1,10E-65	77.17%		9>4	7	3	0.00241	2.15	0.46
SC_00006572.1.1	uncharacterized protein	2,00E-01	50.5%		9>4	5	2	0.00306	2.14	0.47
SC_00036880.1.1	uncharacterized protein	6,80E-52	56.4%		4>9	1	1	0.00399	0.48	2.10
SC_00083346.1.1	uncharacterized protein	6,60E-14	79.67%		4>9	3	6	0.00565	0.48	2.08
SC_00009912.1.1	uncharacterized protein	2,30E-31	83.0%		9>4	7	3	0.00619	2.76	0.36
SC_00011016.1.1	uncharacterized protein	3,20E-16	59.5%		9>4	83	17	0.00640	4.92	0.20

SC_00041792.1.1	uncharacterized protein	5,20E-22	77.5%		9>4	5	2	0.00720	2.40	0.42
SC_00028827.1.1	uncharacterized protein	6,20E-35	81.0%		9>4	11	4	0.00721	2.49	0.40
SC_00000551.1.1	uncharacterized protein	1,90E-05	51.0%		9>4	10	4	0.00833	2.40	0.42
SC_CYP306A1_BOMMO.1499.3286	uncharacterized protein	2,90E-02	58.0%		4>9	15	34	0.00909	0.45	2.24
SC_00115042.1.1	uncharacterized protein	2,70E-24	61.25%		9>4	10	5	0.01038	2.04	0.49
SC_00148050.1.1	uncharacterized protein	9,00E-43	64.0%		9>4	12	6	0.01056	2.07	0.48
SC_00052196.1.1	uncharacterized protein	6,10E-07	57.0%		9>4	35	17	0.01063	2.06	0.49
SC_00007372.1.1	uncharacterized protein	8,60E-08	59.0%		9>4	5	2	0.01126	2.58	0.39
SC_00077734.1.1	uncharacterized protein	5,60E-03	75.0%		9>4	8	4	0.01315	2.31	0.43
SC_00147980.1.1	uncharacterized protein	1,80E-31	71.0%		9>4	6	3	0.01325	2.45	0.41
SC_00004453.1.1	uncharacterized protein	3,90E-02	53.0%		9>4	2	1	0.01457	2.94	0.34
SC_U4974_POPTR.1.1	uncharacterized protein	2,20E-05	46.6%		9>4	5	2	0.01527	2.24	0.45
SC_00147695.1.1	uncharacterized protein	3,70E-02	54.5%		4>9	6	17	0.01614	0.36	2.80
SC_00138719.1.1	uncharacterized protein	7,90E-04	53.33%		4>9	21	51	0.02002	0.42	2.40
SC_00144919.1.1	uncharacterized protein	7,80E-13	70.0%		9>4	9	4	0.02005	2.28	0.44
SC_00147158.1.1	uncharacterized protein	4,80E-58	75.25%		9>4	9	3	0.02009	2.88	0.35
SC_CYP306A1_BOMMO.2175.3286	uncharacterized protein	2,00E-02	70.0%		9>4	9	3	0.02103	2.54	0.39
SC_00037467.1.1	uncharacterized protein	1,20E-27	58.5%		9>4	7	2	0.02132	3.51	0.29
SC_CL4730_1_MAIZE.1.1	uncharacterized protein	2,70E-56	63.25%		9>4	8	2	0.02310	3.06	0.33
SC_00017211.1.1	uncharacterized protein	8,00E-13	84.0%		9>4	5	2	0.02370	2.12	0.47
SC_00005017.1.1	uncharacterized protein	4,70E-19	70.93%		9>4	4	2	0.02392	2.05	0.49
SC_00147752.1.1	uncharacterized protein	7,60E-03	43.0%		4>9	2	5	0.02547	0.43	2.34
SC_00022315.1.1	uncharacterized protein	2,30E-13	57.0%		9>4	10	2	0.02567	4.73	0.21
SC_00147646.1.1	uncharacterized protein	6,40E-17	70.0%		9>4	7	3	0.02575	2.11	0.47
SC_OS11G0528300_ORYSJ.1.1	uncharacterized protein	1,70E-37	82.0%		9>4	9	4	0.02580	2.42	0.41
SC_00016213.1.1	uncharacterized protein	7,10E-05	45.0%		4>9	1	2	0.02726	0.46	2.16
SC_00157607.1.1	uncharacterized protein	7,30E-20	81.0%		9>4	5	3	0.02734	2.04	0.49
SC_00017283.1.1	uncharacterized protein	2,20E-01	42.0%		9>4	83	23	0.02740	3.59	0.28
SC_00079583.1.1	uncharacterized protein	1,70E-04	55.0%		4>9	10	27	0.02793	0.38	2.66
SC_00006509.1.1	uncharacterized protein	1,50E-21	80.5%		9>4	2	1	0.02868	2.17	0.46
SC_00073419.1.1	uncharacterized protein	1,50E-35	71.5%		9>4	5	2	0.02870	2.26	0.44
SC_00025146.1.1	uncharacterized protein	3,00E-13	54.5%		4>9	2	5	0.02994	0.45	2.21
SC_00001566.1.1	uncharacterized protein	3,10E-03	66.0%		9>4	10	4	0.03018	2.22	0.45
SC_00113311.1.1	uncharacterized protein	4,30E-07	53.2%		4>9	1	3	0.03092	0.47	2.11
SC_00059496.1.1	uncharacterized protein	3,60E-23	66.0%		9>4	3	2	0.03096	2.01	0.50
SC_00005416.1.1	uncharacterized protein	1,70E-29	92.0%		9>4	16	7	0.03109	2.40	0.42
SC_00002193.1.1	uncharacterized protein	2,10E-13	78.0%		9>4	6	3	0.03153	2.06	0.49
SC_00021284.1.1	uncharacterized protein	1,90E-37	69.0%		9>4	19	9	0.03284	2.03	0.49
SC_00003223.1.1	uncharacterized protein	9,00E-20	59.4%		9>4	6	3	0.03651	2.09	0.48
SC_00113827.1.1	uncharacterized protein	4,40E-01	41.0%		4>9	2	7	0.03703	0.29	3.46
SC_00054237.1.1	uncharacterized protein	5,80E-02	45.0%		4>9	3	9	0.03808	0.37	2.70
SC_00077018.1.1	uncharacterized protein	6,40E-13	43.0%		9>4	11	5	0.03827	2.29	0.44
SC_00049891.1.1	uncharacterized protein	1,90E-15	70.0%		4>9	2	5	0.03860	0.42	2.39
SC_00008424.1.1	uncharacterized protein	3,60E-04	55.0%		9>4	3	2	0.03925	2.01	0.50

SC_CYP306A1_BOMMO.196.3286	uncharacterized protein	1,40E-02	61.0%		9>4	42	20	0.03960	2.12	0.47
SC_00019317.1.1	uncharacterized protein	3,10E-10	57.67%		9>4	4	2	0.04025	2.64	0.38
SC_FGLH1_TAKRU.762.762	uncharacterized protein	1,40E-01	50.5%		4>9	6	12	0.04087	0.49	2.03
SC_00145605.1.1	uncharacterized protein	3,90E-15	68.5%		9>4	136	63	0.04152	2.17	0.46
SC_00010162.1.1	uncharacterized protein	1,00E-124	51.67%		4>9	11	22	0.04191	0.48	2.08
SC_00067630.1.1	uncharacterized protein	1,60E-64	72.0%		4>9	3	7	0.04255	0.40	2.52
SC_00040109.1.1	uncharacterized protein	2,10E-13	64.0%		9>4	6	2	0.04311	3.00	0.33
SC_00012193.1.1	uncharacterized protein	2,60E-10	57.2%		9>4	41	13	0.04337	3.11	0.32
SC_00037896.1.1	uncharacterized protein	5,00E-35	53.0%		4>9	2	4	0.04359	0.40	2.50
SC_00010091.1.1	uncharacterized protein	5,50E-18	47.5%		9>4	24	12	0.04537	2.00	0.50
SC_AT5G51040_ARATH.3.6	uncharacterized protein	7,40E-49	82.0%		9>4	11	4	0.04807	3.01	0.33
SC_00072816.1.1	uncharacterized protein	2,90E-02	56.0%		9>4	10	4	0.04812	2.77	0.36
SC_00016386.1.1	uncharacterized protein	7,30E-15	46.33%		9>4	7	3	0.04827	2.03	0.49
SC_00010940.1.1	uncharacterized protein	4,40E-04	54.0%		9>4	3	1	0.04963	2.11	0.47
SC_00074029.1.1	uncharacterized protein	2,20E-02	51.0%		9>4	5	2	0.04966	2.26	0.44
SC_OS07G0673400_ORYSJ.1.1	universal stress protein (USP) family protein	2,00E-08	56.8%		9>4	15	7	0.00292	2.12	0.47
SC_OS10G0130700_ORYSJ.11.11	universal stress protein (USP) family protein	4,50E-11	66.06%		4>9	6	18	0.01658	0.32	3.08
SC_00071147.1.1	urease accessory protein D	6,50E-66	84.0%		9>4	6	3	0.01716	2.25	0.44
SC_00056184.1.1	uridylate kinase	2,00E-32	62.0%		4>9	1	3	0.04183	0.44	2.28
SC_CAX1A_ORYSJ.1.1	vacuolar cation/proton exchanger 3	5,00E-35	70.83%		9>4	11	4	0.03582	2.74	0.36
SC_00119624.1.1	vacuolar iron transporter-like protein	1,70E-55	64.0%		4>9	1	2	0.02310	0.48	2.10
SC_VTI11_ARATH.1.1	vesicle transport v-SNARE 11	2,10E-62	77.25%		9>4	20	10	0.01485	2.00	0.50
SC_00020356.1.1	VQ motif-containing protein	7,90E-08	56.75%		9>4	11	2	0.01142	5.14	0.19
SC_FGLH1_TAKRU.23.762	VQ motif-containing protein	1,20E-03	56.0%		9>4	26	6	0.04233	4.04	0.25
SC_00157340.1.1	V-type proton ATPase subunit D	9,90E-28	74.0%		4>9	1	5	0.01153	0.32	3.09
SC_00044261.1.1	V-type proton ATPase subunit d1	9,90E-74	92.0%		9>4	3	2	0.04602	2.02	0.50
SC_00075273.1.1	WcaG domain-containing protein	2,10E-13	54.0%		9>4	9	3	0.00011	3.08	0.32
SC_00008218.1.1	WD40/YVTN repeat and Bromo-WDR9-I-like domain-containing protein	7,00E-73	56.0%		9>4	1	1	0.04291	2.09	0.48
SC_PR4_PHAVU.9.12	wound-responsive-related protein	1,70E-15	61.67%		9>4	39	15	0.04627	2.61	0.38
SC_00011070.1.1	WUSCHEL related homeobox 11	2,00E-30	69.72%		9>4	8	4	0.01416	2.18	0.46
SC_00005814.1.1	WUSCHEL-related homeobox 2	1,70E-24	65.7%		9>4	32	16	0.00525	2.03	0.49
SC_XTHA_PHAAN.3.10	xyloglucan:xyloglucosyl transferase	5,50E-99	69.2%	EC:2.4.1.207	9>4	13	4	0.01040	3.25	0.31
SC_SAP6_ARATH.3.3	zinc finger A20 and AN1 domain-containing stress-associated protein 4	1,70E-17	69.09%	EC:6.3.2.19	9>4	8	3	0.00254	2.51	0.40
SC_SAP6_ARATH.1.3	zinc finger A20 and AN1 domain-containing stress-associated protein 4	1,20E-31	67.1%	EC:6.3.2.19	9>4	3	2	0.03070	2.04	0.49
SC_00001257.1.1	zinc finger CCCH domain-containing protein 52	4,10E-46	54.55%		9>4	7	3	0.03954	2.16	0.46
SC_COL5_ARATH.1.3	zinc finger protein CONSTANS-LIKE 5	3,60E-20	76.05%		9>4	9	4	0.02190	2.30	0.44
SC_00112637.1.1	zinc ion binding / nucleic acid binding	1,30E-16	39.0%		4>9	5	26	0.00179	0.19	5.20
SC_00041493.1.1	zinc ion binding protein	2,10E-07	69.0%		4>9	1	3	0.00460	0.38	2.65
SC_OS09G0551200_ORYSJ.3.7	zinc knuckle (CCHC-type) family protein	4,10E-07	45.5%		4>9	23	93	0.01836	0.25	4.07
SC_ZIP10_ORYSJ.1.5	zinc transporter 4 precursor	2,60E-81	67.53%		9>4	8	3	0.02027	2.42	0.41
SC_00144857.1.1	no matched	-	-	-	9>4	13	3	0.00004	3.95	0.25

SC_CYP306A1_BOMMO.2686.3286	no matched	-	-	-	4>9	5	12	0.00005	0.38	2.63
SC_00021019.1.1	no matched	-	-	-	9>4	2	1	0.00012	2.54	0.39
SC_00021039.1.1	no matched	-	-	-	9>4	9	3	0.00051	3.07	0.33
SC_CYP306A1_BOMMO.513.3286	no matched	-	-	-	9>4	50	22	0.00076	2.30	0.44
SC_00026448.1.1	no matched	-	-	-	4>9	21	49	0.00078	0.44	2.29
SC_00096677.1.1	no matched	-	-	-	4>9	1	5	0.00078	0.30	3.30
SC_00108716.1.1	no matched	-	-	-	9>4	9	3	0.00080	2.54	0.39
SC_00025582.1.1	no matched	-	-	-	4>9	20	45	0.00089	0.44	2.29
SC_00023001.1.1	no matched	-	-	-	9>4	33	15	0.00094	2.24	0.45
SC_00156197.1.1	no matched	-	-	-	9>4	9	4	0.00109	2.23	0.45
SC_00134671.1.1	no matched	-	-	-	4>9	3	14	0.00110	0.19	5.22
SC_00150103.1.1	no matched	-	-	-	4>9	9	28	0.00126	0.33	3.00
SC_00044812.1.1	no matched	-	-	-	9>4	5	2	0.00155	2.91	0.34
SC_00071446.1.1	no matched	-	-	-	4>9	10	25	0.00174	0.41	2.44
SC_00007760.1.1	no matched	-	-	-	4>9	13	61	0.00184	0.22	4.54
SC_00017685.1.1	no matched	-	-	-	9>4	17	7	0.00194	2.36	0.42
SC_00021353.1.1	no matched	-	-	-	4>9	4	8	0.00195	0.49	2.02
SC_CYP306A1_BOMMO.1085.3286	no matched	-	-	-	9>4	21	8	0.00214	2.77	0.36
SC_00053711.1.1	no matched	-	-	-	4>9	1	3	0.00219	0.45	2.20
SC_00038712.1.1	no matched	-	-	-	9>4	10	3	0.00230	3.69	0.27
SC_00071022.1.1	no matched	-	-	-	4>9	3	14	0.00230	0.21	4.67
SC_00012284.1.1	no matched	-	-	-	9>4	5	1	0.00230	3.95	0.25
SC_00132252.1.1	no matched	-	-	-	4>9	7	26	0.00235	0.27	3.64
SC_00122682.1.1	no matched	-	-	-	4>9	1	3	0.00283	0.43	2.31
SC_00029234.1.1	no matched	-	-	-	9>4	16	5	0.00293	3.40	0.29
SC_00001624.1.1	no matched	-	-	-	9>4	14	6	0.00323	2.30	0.44
SC_00039338.1.1	no matched	-	-	-	9>4	12	5	0.00341	2.46	0.41
SC_00005400.1.1	no matched	-	-	-	9>4	7	4	0.00349	2.03	0.49
SC_00002334.1.1	no matched	-	-	-	9>4	7	2	0.00359	2.79	0.36
SC_CYP306A1_BOMMO.640.3286	no matched	-	-	-	9>4	54	27	0.00371	2.02	0.50
SC_CYP306A1_BOMMO.99.3286	no matched	-	-	-	9>4	68	31	0.00377	2.17	0.46
SC_EXTN1_ARATH.1.4	no matched	-	-	-	9>4	10	1	0.00388	6.79	0.15
SC_00066119.1.1	no matched	-	-	-	9>4	14	5	0.00395	2.77	0.36
SC_00046490.1.1	no matched	-	-	-	9>4	20	8	0.00395	2.43	0.41
SC_00161030.1.1	no matched	-	-	-	4>9	3	8	0.00424	0.39	2.59
SC_CYP306A1_BOMMO.1413.3286	no matched	-	-	-	4>9	5	21	0.00478	0.24	4.12
SC_00105623.1.1	no matched	-	-	-	4>9	1	2	0.00478	0.34	2.91
SC_00028636.1.1	no matched	-	-	-	9>4	5	2	0.00492	2.13	0.47
SC_00135582.1.1	no matched	-	-	-	4>9	8	17	0.00503	0.44	2.25
SC_00010601.1.1	no matched	-	-	-	9>4	23	9	0.00534	2.63	0.38
SC_00112501.1.1	no matched	-	-	-	4>9	5	10	0.00544	0.49	2.03
SC_00033142.1.1	no matched	-	-	-	9>4	4	2	0.00557	2.43	0.41
SC_CYP306A1_BOMMO.414.3286	no matched	-	-	-	9>4	151	44	0.00594	3.40	0.29
SC_00014951.1.1	no matched	-	-	-	9>4	41	9	0.00602	4.48	0.22
SC_00064240.1.1	no matched	-	-	-	9>4	8	3	0.00615	2.40	0.42

SC_00017426.1.1	no matched	-	-	-	9>4	13	6	0.00628	2.23	0.45
SC_00031703.1.1	no matched	-	-	-	9>4	3	1	0.00633	3.10	0.32
SC_00018304.1.1	no matched	-	-	-	4>9	7	17	0.00654	0.43	2.33
SC_00006437.1.1	no matched	-	-	-	9>4	22	10	0.00677	2.22	0.45
SC_00120610.1.1	no matched	-	-	-	4>9	1	2	0.00689	0.46	2.19
SC_CYP306A1_BOMMO.1088.3286	no matched	-	-	-	4>9	5	12	0.00694	0.43	2.31
SC_00131523.1.1	no matched	-	-	-	4>9	1	9	0.00718	0.15	6.73
SC_00002621.1.1	no matched	-	-	-	9>4	123	55	0.00739	2.22	0.45
SC_00079694.1.1	no matched	-	-	-	9>4	93	39	0.00746	2.38	0.42
SC_00145708.1.1	no matched	-	-	-	4>9	4	8	0.00759	0.44	2.27
SC_00114576.1.1	no matched	-	-	-	9>4	33	15	0.00788	2.23	0.45
SC_00099137.1.1	no matched	-	-	-	4>9	6	28	0.00838	0.21	4.74
SC_00081475.1.1	no matched	-	-	-	9>4	71	24	0.00845	2.94	0.34
SC_00015836.1.1	no matched	-	-	-	9>4	5	1	0.00865	4.35	0.23
SC_00158685.1.1	no matched	-	-	-	9>4	4	2	0.00885	2.29	0.44
SC_00052481.1.1	no matched	-	-	-	4>9	2	10	0.00894	0.19	5.21
SC_00128577.1.1	no matched	-	-	-	4>9	2	7	0.00896	0.33	3.06
SC_00100960.1.1	no matched	-	-	-	4>9	4	12	0.00897	0.35	2.82
SC_00073335.1.1	no matched	-	-	-	9>4	7	2	0.00942	3.73	0.27
SC_00094008.1.1	no matched	-	-	-	4>9	2	5	0.00953	0.42	2.38
SC_00051862.1.1	no matched	-	-	-	9>4	7	3	0.00981	2.60	0.38
SC_ASR2_SOLLC.2.31	no matched	-	-	-	9>4	19	3	0.00990	6.90	0.15
SC_00061637.1.1	no matched	-	-	-	4>9	11	53	0.01017	0.21	4.77
SC_00108776.1.1	no matched	-	-	-	4>9	10	24	0.01039	0.42	2.39
SC_00028469.1.1	no matched	-	-	-	9>4	4	2	0.01043	2.34	0.43
SC_00061043.1.1	no matched	-	-	-	4>9	5	20	0.01050	0.24	4.16
SC_FGLH1_TAKRU.326.762	no matched	-	-	-	4>9	3	13	0.01053	0.22	4.58
SC_CYP306A1_BOMMO.2954.3286	no matched	-	-	-	9>4	20	3	0.01057	5.70	0.18
SC_00075042.1.1	no matched	-	-	-	4>9	3	9	0.01059	0.35	2.83
SC_00049815.1.1	no matched	-	-	-	9>4	8	4	0.01068	2.15	0.47
SC_00054127.1.1	no matched	-	-	-	4>9	22	57	0.01071	0.39	2.56
SC_00070279.1.1	no matched	-	-	-	4>9	2	14	0.01085	0.15	6.47
SC_00034480.1.1	no matched	-	-	-	9>4	5	2	0.01104	2.92	0.34
SC_00019745.1.1	no matched	-	-	-	9>4	7	2	0.01109	3.23	0.31
SC_00047677.1.1	no matched	-	-	-	9>4	45	18	0.01109	2.53	0.40
SC_CYP306A1_BOMMO.818.3286	no matched	-	-	-	4>9	4	11	0.01111	0.42	2.40
SC_FGLH1_TAKRU.645.762	no matched	-	-	-	4>9	2	9	0.01116	0.21	4.69
SC_00047031.1.1	no matched	-	-	-	9>4	4	2	0.01119	2.40	0.42
SC_POLX_TOBAC.25.242	no matched	-	-	-	4>9	4	11	0.01130	0.38	2.63
SC_00041242.1.1	no matched	-	-	-	4>9	9	19	0.01132	0.49	2.06
SC_00044495.1.1	no matched	-	-	-	4>9	6	17	0.01152	0.33	3.01
SC_00146338.1.1	no matched	-	-	-	4>9	26	90	0.01182	0.28	3.52
SC_00049282.1.1	no matched	-	-	-	9>4	3	1	0.01185	2.50	0.40
SC_00151869.1.1	no matched	-	-	-	4>9	4	13	0.01188	0.35	2.87

SC_00035252.1.1	no matched	-	-	-	4>9	0	1	0.01197	0.30	3.37
SC_CYP306A1_BOMMO.1063.3286	no matched	-	-	-	9>4	32	7	0.01212	4.55	0.22
SC_00072813.1.1	no matched	-	-	-	4>9	2	5	0.01223	0.46	2.18
SC_00023032.1.1	no matched	-	-	-	4>9	1	3	0.01243	0.49	2.06
SC_00028306.1.1	no matched	-	-	-	9>4	40	17	0.01247	2.35	0.43
SC_OS09G0551200_ORYSJ.7.7	no matched	-	-	-	4>9	2	7	0.01250	0.29	3.41
SC_00134365.1.1	no matched	-	-	-	4>9	4	13	0.01264	0.27	3.74
SC_00074835.1.1	no matched	-	-	-	9>4	4	1	0.01290	3.18	0.31
SC_FGLH1_TAKRU.448.762	no matched	-	-	-	4>9	2	5	0.01291	0.37	2.73
SC_00061670.1.1	no matched	-	-	-	4>9	4	7	0.01291	0.49	2.06
SC_00086016.1.1	no matched	-	-	-	9>4	7	3	0.01308	2.58	0.39
SC_00046048.1.1	no matched	-	-	-	9>4	16	8	0.01334	2.05	0.49
SC_00029703.1.1	no matched	-	-	-	4>9	3	7	0.01362	0.49	2.06
SC_00025080.1.1	no matched	-	-	-	4>9	18	64	0.01376	0.29	3.47
SC_00154137.1.1	no matched	-	-	-	9>4	8	3	0.01383	2.46	0.41
SC_CYP306A1_BOMMO.926.3286	no matched	-	-	-	9>4	24	11	0.01390	2.23	0.45
SC_00038593.1.1	no matched	-	-	-	9>4	10	4	0.01414	2.70	0.37
SC_00022812.1.1	no matched	-	-	-	9>4	2	1	0.01416	2.02	0.49
SC_00058349.1.1	no matched	-	-	-	9>4	6	2	0.01419	2.68	0.37
SC_00048953.1.1	no matched	-	-	-	9>4	8	3	0.01434	2.37	0.42
SC_00017745.1.1	no matched	-	-	-	4>9	2	4	0.01448	0.44	2.25
SC_00109349.1.1	no matched	-	-	-	9>4	22	9	0.01454	2.60	0.39
SC_00055810.1.1	no matched	-	-	-	4>9	3	9	0.01465	0.31	3.21
SC_00145795.1.1	no matched	-	-	-	4>9	2	4	0.01507	0.44	2.27
SC_00099104.1.1	no matched	-	-	-	9>4	6	2	0.01508	2.36	0.42
SC_FGLH1_TAKRU.295.762	no matched	-	-	-	9>4	11	3	0.01513	3.68	0.27
SC_00041396.1.1	no matched	-	-	-	9>4	4	2	0.01521	2.27	0.44
SC_CYP306A1_BOMMO.2835.3286	no matched	-	-	-	4>9	6	21	0.01522	0.28	3.56
SC_CYP306A1_BOMMO.270.3286	no matched	-	-	-	4>9	2	16	0.01541	0.13	7.71
SC_CYP306A1_BOMMO.848.3286	no matched	-	-	-	9>4	33	14	0.01545	2.30	0.43
SC_00105701.1.1	no matched	-	-	-	9>4	17	4	0.01546	4.48	0.22
SC_00028891.1.1	no matched	-	-	-	9>4	109	52	0.01549	2.10	0.48
SC_00014449.1.1	no matched	-	-	-	4>9	3	11	0.01568	0.29	3.47
SC_00154199.1.1	no matched	-	-	-	4>9	3	10	0.01572	0.29	3.47
SC_00067128.1.1	no matched	-	-	-	4>9	3	11	0.01589	0.31	3.23
SC_CYP306A1_BOMMO.2773.3286	no matched	-	-	-	9>4	9	3	0.01597	2.86	0.35
SC_00024944.1.1	no matched	-	-	-	4>9	46	310	0.01604	0.15	6.75
SC_00115268.1.1	no matched	-	-	-	4>9	3	14	0.01619	0.20	4.94
SC_00073555.1.1	no matched	-	-	-	4>9	18	51	0.01637	0.36	2.78
SC_00034915.1.1	no matched	-	-	-	4>9	22	49	0.01654	0.46	2.20
SC_00043769.1.1	no matched	-	-	-	9>4	9	4	0.01659	2.48	0.40
SC_00038854.1.1	no matched	-	-	-	9>4	5	2	0.01661	2.32	0.43
SC_00133492.1.1	no matched	-	-	-	9>4	4	2	0.01663	2.13	0.47
SC_00091972.1.1	no matched	-	-	-	4>9	2	6	0.01667	0.28	3.51

SC_00151136.1.1	no matched	-	-	-	4>9	12	46	0.01673	0.25	3.93
SC_CYP306A1_BOMMO.679.3286	no matched	-	-	-	9>4	52	24	0.01673	2.14	0.47
SC_00155483.1.1	no matched	-	-	-	9>4	119	52	0.01680	2.29	0.44
SC_00087314.1.1	no matched	-	-	-	4>9	1	3	0.01696	0.45	2.23
SC_00121676.1.1	no matched	-	-	-	4>9	3	7	0.01698	0.42	2.39
SC_00145154.1.1	no matched	-	-	-	4>9	2	5	0.01707	0.42	2.40
SC_00117462.1.1	no matched	-	-	-	4>9	2	4	0.01711	0.41	2.43
SC_00157781.1.1	no matched	-	-	-	4>9	2	13	0.01724	0.18	5.52
SC_ASR2_SOLLC.14.31	no matched	-	-	-	9>4	73	26	0.01725	2.81	0.36
SC_00063076.1.1	no matched	-	-	-	9>4	20	9	0.01738	2.32	0.43
SC_00146121.1.1	no matched	-	-	-	9>4	4	2	0.01788	2.30	0.43
SC_00105696.1.1	no matched	-	-	-	9>4	7	3	0.01811	2.19	0.46
SC_00147185.1.1	no matched	-	-	-	4>9	15	45	0.01822	0.34	2.95
SC_00075604.1.1	no matched	-	-	-	4>9	1	5	0.01835	0.29	3.47
SC_00023325.1.1	no matched	-	-	-	9>4	5	2	0.01842	2.32	0.43
SC_CYP306A1_BOMMO.2784.3286	no matched	-	-	-	4>9	3	9	0.01853	0.35	2.83
SC_00012757.1.1	no matched	-	-	-	9>4	5	2	0.01892	2.34	0.43
SC_00126053.1.1	no matched	-	-	-	9>4	10	3	0.01902	2.96	0.34
SC_FGLH1_TAKRU.91.762	no matched	-	-	-	9>4	36	14	0.01914	2.47	0.40
SC_00062201.1.1	no matched	-	-	-	4>9	3	7	0.01917	0.45	2.23
SC_LOC100275739_MAIZE.55.160	no matched	-	-	-	9>4	3	1	0.01974	2.24	0.45
SC_00023334.1.1	no matched	-	-	-	9>4	29	4	0.01976	7.74	0.13
SC_00104751.1.1	no matched	-	-	-	9>4	4	2	0.01993	2.18	0.46
SC_00072993.1.1	no matched	-	-	-	9>4	19	7	0.02008	2.54	0.39
SC_00002392.1.1	no matched	-	-	-	9>4	2	1	0.02020	2.97	0.34
SC_00010324.1.1	no matched	-	-	-	4>9	24	56	0.02026	0.43	2.33
SC_00113113.1.1	no matched	-	-	-	4>9	1	2	0.02031	0.32	3.14
SC_00046097.1.1	no matched	-	-	-	4>9	1	3	0.02031	0.37	2.72
SC_00073385.1.1	no matched	-	-	-	9>4	7	2	0.02048	2.93	0.34
SC_00146207.1.1	no matched	-	-	-	4>9	10	24	0.02055	0.40	2.53
SC_00149830.1.1	no matched	-	-	-	4>9	2	7	0.02059	0.24	4.16
SC_CYP306A1_BOMMO.1255.3286	no matched	-	-	-	9>4	5	1	0.02066	3.74	0.27
SC_OS07G0588800_ORYSJ.1.1	no matched	-	-	-	4>9	4	9	0.02075	0.42	2.36
SC_00151972.1.1	no matched	-	-	-	9>4	9	3	0.02083	3.20	0.31
SC_00058817.1.1	no matched	-	-	-	4>9	2	4	0.02087	0.42	2.38
SC_00002597.1.1	no matched	-	-	-	9>4	5	2	0.02119	2.72	0.37
SC_00020423.1.1	no matched	-	-	-	9>4	9	3	0.02149	2.91	0.34
SC_00020573.1.1	no matched	-	-	-	9>4	90	37	0.02151	2.44	0.41
SC_00025611.1.1	no matched	-	-	-	9>4	3	1	0.02162	2.37	0.42
SC_00162092.1.1	no matched	-	-	-	4>9	3	12	0.02177	0.23	4.36
SC_00021982.1.1	no matched	-	-	-	9>4	27	8	0.02213	3.47	0.29
SC_00064124.1.1	no matched	-	-	-	4>9	3	11	0.02215	0.30	3.38
SC_00083660.1.1	no matched	-	-	-	4>9	15	46	0.02231	0.33	3.02
SC_00131627.1.1	no matched	-	-	-	9>4	14	4	0.02238	3.04	0.33

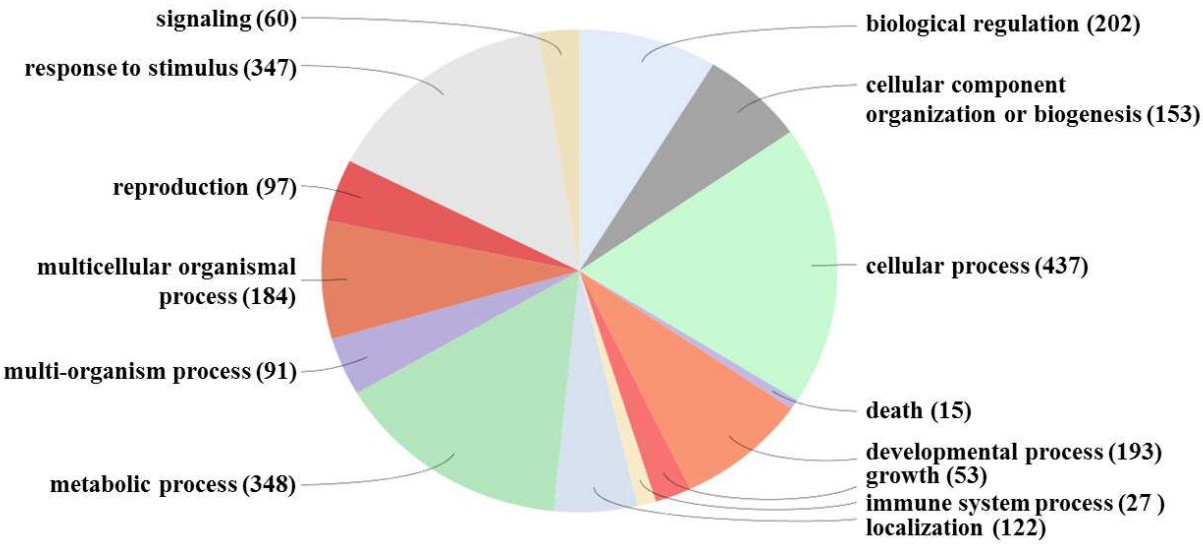
SC_00054566.1.1	no matched	-	-	-	9>4	10	3	0.02246	3.05	0.33
SC_00027502.1.1	no matched	-	-	-	9>4	12	5	0.02247	2.51	0.40
SC_00133558.1.1	no matched	-	-	-	4>9	3	11	0.02259	0.29	3.41
SC_LOC100275739_MAIZE.44.160	no matched	-	-	-	9>4	6	2	0.02293	2.28	0.44
SC_00019533.1.1	no matched	-	-	-	9>4	5	2	0.02309	2.20	0.45
SC_00064754.1.1	no matched	-	-	-	4>9	4	10	0.02345	0.34	2.95
SC_00125088.1.1	no matched	-	-	-	4>9	4	10	0.02349	0.41	2.47
SC_00093552.1.1	no matched	-	-	-	4>9	3	7	0.02354	0.39	2.57
SC_00012473.1.1	no matched	-	-	-	9>4	14	7	0.02366	2.01	0.50
SC_00136469.1.1	no matched	-	-	-	9>4	4	2	0.02399	2.77	0.36
SC_00146022.1.1	no matched	-	-	-	9>4	48	23	0.02415	2.04	0.49
SC_FGLH1_TAKRU.654.762	no matched	-	-	-	4>9	9	25	0.02416	0.36	2.81
SC_00088276.1.1	no matched	-	-	-	4>9	1	3	0.02419	0.48	2.07
SC_POLX_TOBAC.162.242	no matched	-	-	-	4>9	5	17	0.02436	0.29	3.43
SC_00086530.1.1	no matched	-	-	-	4>9	4	16	0.02452	0.25	4.02
SC_00044395.1.1	no matched	-	-	-	9>4	18	6	0.02464	2.92	0.34
SC_00076605.1.1	no matched	-	-	-	4>9	5	13	0.02484	0.40	2.51
SC_00011326.1.1	no matched	-	-	-	9>4	5	3	0.02492	2.06	0.49
SC_00031242.1.1	no matched	-	-	-	9>4	2	1	0.02513	2.23	0.45
SC_00043435.1.1	no matched	-	-	-	4>9	3	8	0.02542	0.33	3.04
SC_00101180.1.1	no matched	-	-	-	4>9	2	6	0.02554	0.43	2.35
SC_CYP306A1_BOMMO.308.3286	no matched	-	-	-	9>4	9	4	0.02562	2.17	0.46
SC_POLX_TOBAC.26.242	no matched	-	-	-	4>9	8	17	0.02567	0.45	2.23
SC_00013506.1.1	no matched	-	-	-	4>9	13	52	0.02570	0.26	3.90
SC_00006982.1.1	no matched	-	-	-	9>4	9	4	0.02588	2.33	0.43
SC_00090880.1.1	no matched	-	-	-	4>9	1	3	0.02636	0.40	2.49
SC_CYP306A1_BOMMO.2153.3286	no matched	-	-	-	4>9	4	8	0.02644	0.50	2.02
SC_00161982.1.1	no matched	-	-	-	4>9	2	6	0.02665	0.37	2.73
SC_00066739.1.1	no matched	-	-	-	4>9	4	8	0.02670	0.50	2.00
SC_00076111.1.1	no matched	-	-	-	4>9	1	2	0.02704	0.25	3.92
SC_00054884.1.1	no matched	-	-	-	4>9	11	30	0.02711	0.35	2.86
SC_00094136.1.1	no matched	-	-	-	9>4	18	8	0.02716	2.15	0.47
SC_00116760.1.1	no matched	-	-	-	4>9	4	8	0.02777	0.46	2.15
SC_00022851.1.1	no matched	-	-	-	4>9	1	2	0.02783	0.49	2.05
SC_00147506.1.1	no matched	-	-	-	4>9	6	15	0.02800	0.37	2.71
SC_00070410.1.1	no matched	-	-	-	4>9	4	12	0.02808	0.33	3.02
SC_00092243.1.1	no matched	-	-	-	4>9	2	3	0.02826	0.50	2.01
SC_00047112.1.1	no matched	-	-	-	9>4	7	3	0.02828	2.40	0.42
SC_00159539.1.1	no matched	-	-	-	9>4	6	3	0.02832	2.09	0.48
SC_00029002.1.1	no matched	-	-	-	9>4	14	5	0.02842	2.53	0.39
SC_00123836.1.1	no matched	-	-	-	4>9	2	5	0.02843	0.40	2.53
SC_00035054.1.1	no matched	-	-	-	4>9	1	1	0.02849	0.42	2.39
SC_00014202.1.1	no matched	-	-	-	4>9	4	19	0.02856	0.20	5.02
SC_CYP306A1_BOMMO.697.3286	no matched	-	-	-	4>9	15	50	0.02867	0.30	3.29

SC_00154951.1.1	no matched	-	-	-	4>9	2	6	0.02868	0.38	2.65
SC_00048811.1.1	no matched	-	-	-	4>9	2	9	0.02874	0.23	4.39
SC_LOC100275739_MAIZE.49.160	no matched	-	-	-	9>4	100	32	0.02927	3.09	0.32
SC_00035732.1.1	no matched	-	-	-	9>4	8	3	0.02944	2.39	0.42
SC_00108688.1.1	no matched	-	-	-	9>4	4	2	0.02991	2.18	0.46
SC_00031692.1.1	no matched	-	-	-	9>4	11	5	0.02994	2.06	0.48
SC_00083243.1.1	no matched	-	-	-	4>9	7	14	0.02995	0.50	2.02
SC_00046538.1.1	no matched	-	-	-	4>9	1	4	0.03019	0.27	3.70
SC_00053540.1.1	no matched	-	-	-	4>9	21	85	0.03027	0.24	4.14
SC_00161843.1.1	no matched	-	-	-	9>4	7	3	0.03046	2.87	0.35
SC_CYP306A1_BOMMO.85.3286	no matched	-	-	-	9>4	4	2	0.03047	2.25	0.44
SC_FGLH1_TAKRU.267.762	no matched	-	-	-	9>4	10	4	0.03065	2.42	0.41
SC_00010962.1.1	no matched	-	-	-	9>4	4	2	0.03066	2.22	0.45
SC_00018807.1.1	no matched	-	-	-	9>4	18	7	0.03066	2.61	0.38
SC_00138552.1.1	no matched	-	-	-	9>4	6	2	0.03070	3.05	0.33
SC_CYP306A1_BOMMO.457.3286	no matched	-	-	-	9>4	33	5	0.03088	6.26	0.16
SC_00022783.1.1	no matched	-	-	-	4>9	2	3	0.03094	0.49	2.04
SC_00081508.1.1	no matched	-	-	-	4>9	1	3	0.03094	0.40	2.48
SC_00147087.1.1	no matched	-	-	-	4>9	14	74	0.03116	0.19	5.16
SC_00115211.1.1	no matched	-	-	-	4>9	3	6	0.03127	0.41	2.42
SC_MT3_PICGL.68.527	no matched	-	-	-	9>4	8	4	0.03131	2.01	0.50
SC_00109204.1.1	no matched	-	-	-	4>9	2	5	0.03142	0.29	3.46
SC_00140529.1.1	no matched	-	-	-	9>4	26	5	0.03144	5.76	0.17
SC_00066448.1.1	no matched	-	-	-	9>4	8	4	0.03159	2.07	0.48
SC_00009984.1.1	no matched	-	-	-	9>4	6	2	0.03173	2.56	0.39
SC_00116751.1.1	no matched	-	-	-	9>4	8	3	0.03197	3.14	0.32
SC_00025027.1.1	no matched	-	-	-	9>4	4	1	0.03198	2.97	0.34
SC_00110914.1.1	no matched	-	-	-	4>9	1	3	0.03234	0.43	2.30
SC_00012406.1.1	no matched	-	-	-	9>4	61	17	0.03247	3.63	0.28
SC_00122783.1.1	no matched	-	-	-	9>4	5	2	0.03255	2.98	0.34
SC_00118330.1.1	no matched	-	-	-	9>4	8	4	0.03265	2.31	0.43
SC_00011064.1.1	no matched	-	-	-	9>4	7	3	0.03276	2.46	0.41
SC_00038666.1.1	no matched	-	-	-	4>9	14	44	0.03282	0.33	3.03
SC_00019914.1.1	no matched	-	-	-	9>4	76	25	0.03290	3.10	0.32
SC_00074162.1.1	no matched	-	-	-	9>4	27	6	0.03294	4.26	0.23
SC_00052679.1.1	no matched	-	-	-	4>9	1	4	0.03317	0.30	3.33
SC_00068810.1.1	no matched	-	-	-	9>4	3	1	0.03322	2.50	0.40
SC_00016973.1.1	no matched	-	-	-	9>4	7	2	0.03325	2.70	0.37
SC_00056034.1.1	no matched	-	-	-	4>9	2	3	0.03327	0.45	2.24
SC_00034404.1.1	no matched	-	-	-	9>4	7	3	0.03328	2.13	0.47
SC_00034088.1.1	no matched	-	-	-	4>9	3	13	0.03329	0.25	4.07
SC_COPIA_DROME.8.25	no matched	-	-	-	4>9	24	89	0.03331	0.27	3.75
SC_00063532.1.1	no matched	-	-	-	4>9	4	12	0.03335	0.35	2.87
SC_00126520.1.1	no matched	-	-	-	4>9	3	7	0.03432	0.44	2.28

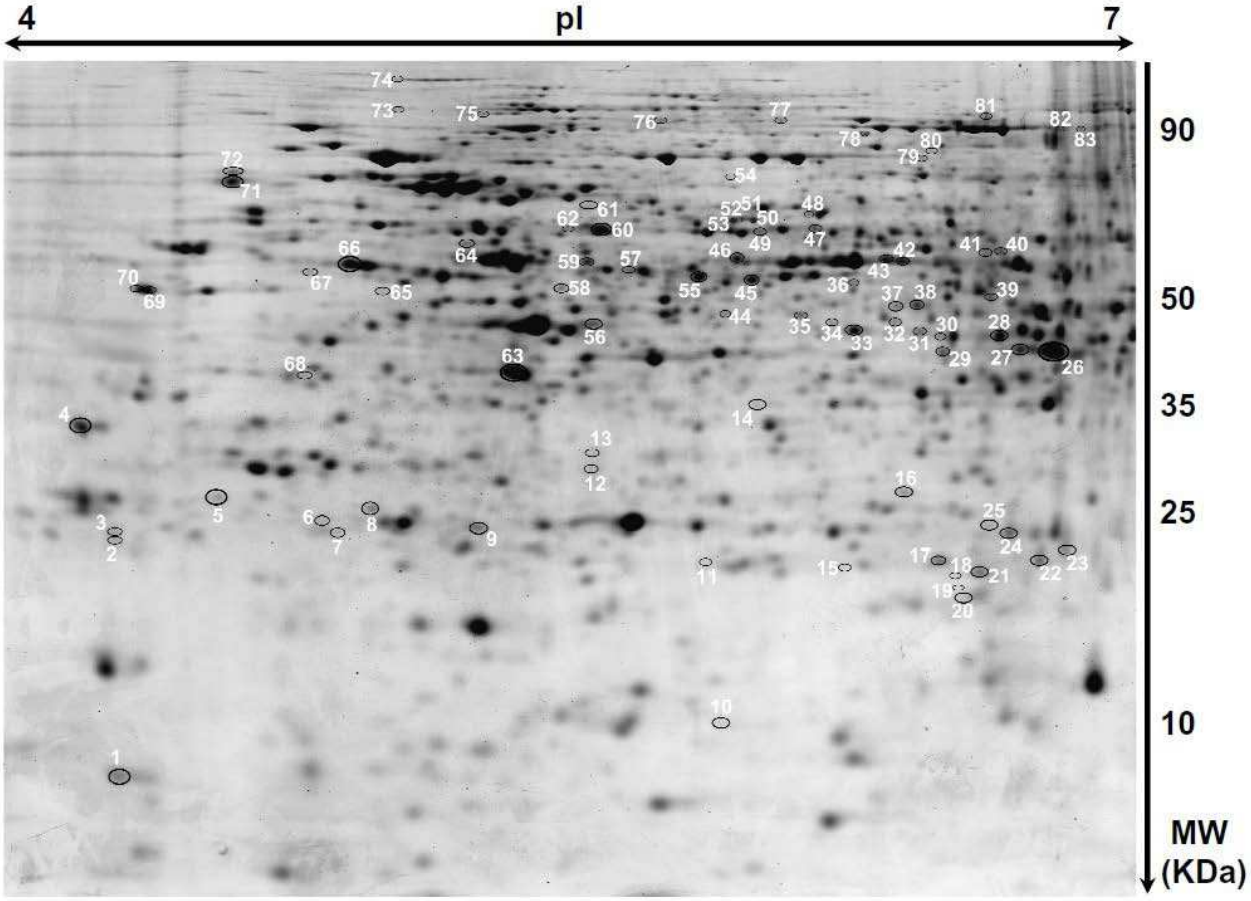
SC_CYP306A1_BOMMO.157.3286	no matched	-	-	-	9>4	14	4	0.03467	3.44	0.29
SC_CYP306A1_BOMMO.2975.3286	no matched	-	-	-	9>4	22	7	0.03489	3.35	0.30
SC_00036612.1.1	no matched	-	-	-	4>9	2	5	0.03500	0.40	2.48
SC_00023097.1.1	no matched	-	-	-	9>4	4	1	0.03529	2.95	0.34
SC_CYP306A1_BOMMO.772.3286	no matched	-	-	-	9>4	71	33	0.03561	2.19	0.46
SC_00145762.1.1	no matched	-	-	-	4>9	16	63	0.03565	0.26	3.81
SC_00161835.1.1	no matched	-	-	-	9>4	23	11	0.03606	2.07	0.48
SC_00062835.1.1	no matched	-	-	-	9>4	46	20	0.03608	2.29	0.44
SC_00049965.1.1	no matched	-	-	-	4>9	2	4	0.03609	0.40	2.50
SC_00133842.1.1	no matched	-	-	-	9>4	105	48	0.03612	2.18	0.46
SC_00154588.1.1	no matched	-	-	-	9>4	5	3	0.03647	2.01	0.50
SC_00130469.1.1	no matched	-	-	-	4>9	1	6	0.03655	0.16	6.09
SC_00064423.1.1	no matched	-	-	-	4>9	1	2	0.03724	0.44	2.29
SC_00162186.1.1	no matched	-	-	-	4>9	2	4	0.03731	0.36	2.75
SC_00022897.1.1	no matched	-	-	-	9>4	17	7	0.03749	2.54	0.39
SC_00146449.1.1	no matched	-	-	-	4>9	9	29	0.03759	0.30	3.28
SC_00058254.1.1	no matched	-	-	-	4>9	2	5	0.03783	0.36	2.80
SC_00109061.1.1	no matched	-	-	-	4>9	4	12	0.03784	0.39	2.60
SC_00093830.1.1	no matched	-	-	-	4>9	3	10	0.03791	0.29	3.43
SC_00124258.1.1	no matched	-	-	-	4>9	2	13	0.03796	0.19	5.14
SC_00084927.1.1	no matched	-	-	-	9>4	20	7	0.03814	2.78	0.36
SC_00052697.1.1	no matched	-	-	-	9>4	3	1	0.03825	2.17	0.46
SC_00022061.1.1	no matched	-	-	-	9>4	5	2	0.03901	2.87	0.35
SC_00121492.1.1	no matched	-	-	-	4>9	2	4	0.03922	0.46	2.16
SC_00044164.1.1	no matched	-	-	-	4>9	2	4	0.03941	0.38	2.60
SC_00002755.1.1	no matched	-	-	-	9>4	16	8	0.03968	2.02	0.50
SC_00068580.1.1	no matched	-	-	-	4>9	2	8	0.03976	0.19	5.29
SC_CYP306A1_BOMMO.167.3286	no matched	-	-	-	9>4	6	2	0.03981	2.87	0.35
SC_00151665.1.1	no matched	-	-	-	9>4	4	2	0.04009	2.01	0.50
SC_LOC100275739_MAIZE.95.160	no matched	-	-	-	9>4	11	4	0.04032	2.37	0.42
SC_00082769.1.1	no matched	-	-	-	4>9	1	3	0.04037	0.35	2.85
SC_00131895.1.1	no matched	-	-	-	4>9	1	4	0.04060	0.34	2.95
SC_00049186.1.1	no matched	-	-	-	9>4	3	1	0.04116	2.27	0.44
SC_CYP306A1_BOMMO.1364.3286	no matched	-	-	-	4>9	64	300	0.04122	0.21	4.69
SC_00159269.1.1	no matched	-	-	-	9>4	5	2	0.04131	2.73	0.37
SC_00021001.1.1	no matched	-	-	-	9>4	7	3	0.04137	2.66	0.38
SC_00079015.1.1	no matched	-	-	-	9>4	78	23	0.04138	3.40	0.29
SC_00043998.1.1	no matched	-	-	-	4>9	2	6	0.04162	0.41	2.41
SC_CYP306A1_BOMMO.2316.3286	no matched	-	-	-	4>9	11	28	0.04174	0.40	2.51
SC_FGLH1_TAKRU.531.762	no matched	-	-	-	4>9	9	32	0.04184	0.29	3.46
SC_00048476.1.1	no matched	-	-	-	4>9	27	118	0.04201	0.23	4.32
SC_00152392.1.1	no matched	-	-	-	4>9	7	28	0.04201	0.26	3.90
SC_CYP306A1_BOMMO.1394.3286	no matched	-	-	-	9>4	90	34	0.04222	2.65	0.38
SC_00146540.1.1	no matched	-	-	-	4>9	2	9	0.04227	0.26	3.87

SC_00056906.1.1	no matched	-	-	-	4>9	1	3	0.04233	0.47	2.14
SC_OS10G0130700_ORYSJ.8.11	no matched	-	-	-	4>9	5	17	0.04245	0.28	3.51
SC_FGLH1_TAKRU.122.762	no matched	-	-	-	4>9	2	11	0.04262	0.20	4.89
SC_00126526.1.1	no matched	-	-	-	4>9	3	7	0.04270	0.48	2.08
SC_00011564.1.1	no matched	-	-	-	9>4	6	2	0.04272	2.70	0.37
SC_00097193.1.1	no matched	-	-	-	4>9	32	162	0.04272	0.19	5.14
SC_00153798.1.1	no matched	-	-	-	4>9	5	10	0.04280	0.44	2.26
SC_00099923.1.1	no matched	-	-	-	9>4	3	1	0.04286	2.70	0.37
SC_CYP306A1_BOMMO.992.3286	no matched	-	-	-	9>4	179	60	0.04296	3.01	0.33
SC_00012293.1.1	no matched	-	-	-	9>4	54	21	0.04317	2.53	0.40
SC_00058532.1.1	no matched	-	-	-	4>9	1	3	0.04328	0.49	2.05
SC_POLX_TOBAC.150.242	no matched	-	-	-	4>9	1	3	0.04338	0.33	3.03
SC_00076478.1.1	no matched	-	-	-	9>4	15	6	0.04349	2.41	0.41
SC_00075428.1.1	no matched	-	-	-	4>9	2	5	0.04357	0.45	2.23
SC_00135681.1.1	no matched	-	-	-	9>4	71	31	0.04381	2.30	0.44
SC_00009126.1.1	no matched	-	-	-	9>4	3	2	0.04388	2.01	0.50
SC_00161477.1.1	no matched	-	-	-	4>9	19	55	0.04411	0.34	2.98
SC_00024105.1.1	no matched	-	-	-	4>9	2	3	0.04412	0.48	2.09
SC_LOC100275739_MAIZE.63.160	no matched	-	-	-	9>4	7	3	0.04500	2.47	0.41
SC_00156821.1.1	no matched	-	-	-	9>4	3	1	0.04502	2.73	0.37
SC_00118636.1.1	no matched	-	-	-	4>9	12	38	0.04511	0.33	3.05
SC_00097862.1.1	no matched	-	-	-	4>9	1	2	0.04513	0.42	2.36
SC_00046478.1.1	no matched	-	-	-	9>4	13	5	0.04516	2.79	0.36
SC_00148144.1.1	no matched	-	-	-	9>4	10	5	0.04520	2.10	0.48
SC_00148385.1.1	no matched	-	-	-	9>4	8	3	0.04527	2.15	0.46
SC_POLX_TOBAC.118.242	no matched	-	-	-	4>9	1	2	0.04545	0.32	3.09
SC_00130255.1.1	no matched	-	-	-	4>9	18	64	0.04568	0.28	3.55
SC_00058413.1.1	no matched	-	-	-	4>9	3	6	0.04605	0.49	2.06
SC_00067217.1.1	no matched	-	-	-	4>9	4	10	0.04625	0.45	2.20
SC_FGLH1_TAKRU.650.762	no matched	-	-	-	9>4	19	7	0.04651	2.58	0.39
SC_00037167.1.1	no matched	-	-	-	4>9	1	2	0.04651	0.48	2.08
SC_00106485.1.1	no matched	-	-	-	4>9	5	11	0.04658	0.48	2.07
SC_FGLH1_TAKRU.317.762	no matched	-	-	-	9>4	49	17	0.04683	2.88	0.35
SC_00061695.1.1	no matched	-	-	-	9>4	4	2	0.04706	2.40	0.42
SC_00017068.1.1	no matched	-	-	-	9>4	7	2	0.04707	3.02	0.33
SC_00071167.1.1	no matched	-	-	-	9>4	3	2	0.04717	2.10	0.48
SC_00139706.1.1	no matched	-	-	-	4>9	1	3	0.04803	0.42	2.39
SC_00008127.1.1	no matched	-	-	-	9>4	20	6	0.04807	3.42	0.29
SC_00073482.1.1	no matched	-	-	-	9>4	32	11	0.04823	2.86	0.35
SC_00018632.1.1	no matched	-	-	-	9>4	240	67	0.04857	3.60	0.28
SC_00039005.1.1	no matched	-	-	-	9>4	36	10	0.04865	3.65	0.27
SC_00132110.1.1	no matched	-	-	-	4>9	10	21	0.04876	0.47	2.13
SC_00111852.1.1	no matched	-	-	-	4>9	3	7	0.04888	0.42	2.36
SC_00029533.1.1	no matched	-	-	-	4>9	93	338	0.04902	0.28	3.63

SC_00003945.1.1	no matched	-	-	-	9>4	49	19	0.04905	2.58	0.39
SC_00080584.1.1	no matched	-	-	-	4>9	14	46	0.04926	0.31	3.22
SC_00137345.1.1	no matched	-	-	-	9>4	12	6	0.04940	2.15	0.46
SC_00012793.1.1	no matched	-	-	-	9>4	8	3	0.04946	2.37	0.42



Appendix S2. Functional annotation of assembled sequences (significant transcripts) based on gene ontology (GO) categorization. GO analysis was performed at level 2 for the Biological Process categories; the numbers in parentheses represent the number of sequences associated with each category.



Appendix S3. Representative 2-D PAGE map for proteins from immature somatic embryos of maritime pine matured for 1 week on high gellan gum (9 g L^{-1}). Marked spots (1-83) displayed significant differences in intensity (P<0.05) from those of the corresponding spots obtained for somatic embryos matured on low gellan gum (4 g L^{-1}).

Appendix S4. Identification of differentially expressed proteins in immature somatic embryos of maritime pine matured for 1 week on either 4 or 9 g L^{-1} gellan gum. Excised spot number refers to the spot marked in Fig. S2. The assignments and accession numbers were retrieved from the GenoToul database (http://*genotoul*-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster2/index.html) by the PEAKS identification engine. The most homologous protein in the UniProt database (http://www.uniprot.org/) is given.

p-Value *: Significance of difference in mean normalized volume of spot calculated from both conditions (t-test) ;* ***MoyVol 4G*** *: mean normalised volume (5 samples) of protein spot from EMs grown on low gellan gum (4G= 4 g* L^{-1}*) ;* ***MoyVol 9G*** *: mean normalised volume (5 samples) of protein spot from EMs grown on high gellan gum (9G=9 g* L^{-1}*);* ***Gellan gum*** *: relative protein abundance in EMs matured on high (9G) vs. low (4G) gellan gum. (9>4 : protein over-expressed on 9G, 4>9 : protein over-expressed on 4G) ;* ***ratio*** *= (MoyVol4G / MoyVol 9G) ;* ***score (%)*** *: PEAKS confidence score ;* ***PEAKS p-Value*** *: value for identified protein ;* ***Unique Peptides*** *: Number of peptide sequences unique to a protein group.*

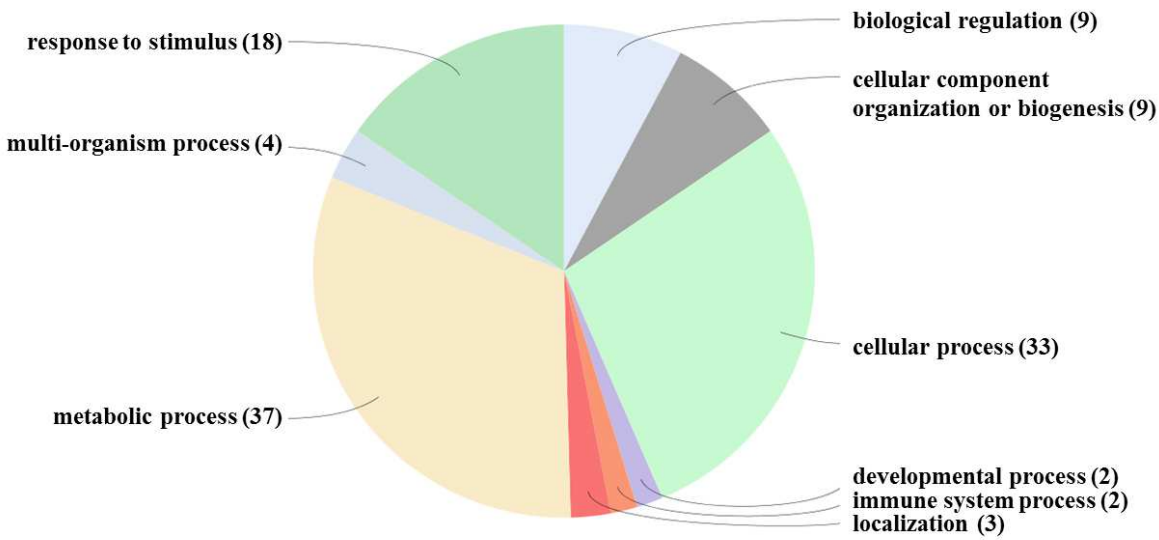
Spot number	Accession number	p-Value	MoyVol 4G	MoyVol 9G	Gellan gum	ratio	UniProt Accession	Protein name	score (%)	PEAKS p-Value	Unique Peptides
5	SC_RP12A_ARATH.1.6	0.041	9700371	8264037	4>9	1.17	Q0GEA1	26S proteasome regulatory particle non-ATPase subunit 12	99	2.04E-313	6
56	SC_STAD_RICCO.1.4	0.004	2600298	1730505	4>9	1.50	A9NVD4	Acyl-[acyl-carrier-protein] desaturase (EC 1.14.19.2)	99.1	4.9E-324	7
36	SC_00001024.1.1	0.035	4241994	3569472	4>9	1.19	Q84UX4	Alanine aminotransferase	99.1	4.9E-324	11
34	SC_ADH_FRAAN.16.30	0.026	2659434	1841416	4>9	1.44	Q43025	Alcohol dehydrogenase (EC 1.1.1.1) (Fragment)	99.1	4.9E-324	16
30	SC_ADH3_SOLTU.1.15	0,002	4339533	3444106	4>9	1.26	Q43025	Alcohol dehydrogenase (EC 1.1.1.1) (Fragment)	99.1	4.9E-324	10
32	SC_00002407.1.1	0.008	2986883	2294163	4>9	1.30	Q9LVW0	At5g54100 (Putative uncharacterized protein At5g54100/MJP23_8) (SPFH/Band 7/PHB domain-containing membrane-associated protein) (Similarity to stomatin like protein)	99.1	4.9E-324	13
59	SC_00001003.1.1	0.039	10397012	9065974	4>9	1.15	D8T9Y4	BAHD family acyltransferase, clade V (EC 2.3.1.-)	99.1	4.9E-324	13
10	SC_RRAA3_ARATH.2.3	0.006	5122693	3771480	4>9	1.36	A3TJS8	Cell division protein FtsQ	99.1	1E-323	10

55	SC_CH62_CUCMA.1.9	0.006	17660376	14891450	4>9	1.19	Q05046	Chaperonin CPN60-2, mitochondrial (HSP60-2)	99.1	4.9E-324	15
62	SC_CH62_CUCMA.2.9	0.003	4462426	3702996	4>9	1.21	Q05046	Chaperonin CPN60-2, mitochondrial (HSP60-2)	99.1	5.81E-252	7
60	SC_CH62_CUCMA.2.9	0.008	25511495	23199527	4>9	1.10	Q05046	Chaperonin CPN60-2, mitochondrial (HSP60-2)	99.1	4.9E-324	21
38	SC_00001092.1.1	0.004	10350683	8755308	4>9	1.18	Q5VNM3	DEAD-box ATP-dependent RNA helicase 2 (EC 3.6.4.13)	99.1	4.9E-324	18
78	SC_MCM31_MAIZE.1.5	0.013	4051517	3259108	4>9	1.24	Q9SX03	DNA replication licensing factor MCM3 homolog 3 (EC 3.6.4.12) (Replication origin activator 3) (ROA-3)	99.1	4.9E-324	13
33	SC_EFTM_ARATH.1.6	0,014	27416086	24017977	4>9	1.14	Q9ZT91	Elongation factor Tu, mitochondrial	99.1	4.9E-324	16
45	SC_ENO1_HEVBR.7.31	0.005	20729471	16586034	4>9	1.25	Q9LEJ0	Enolase 1	99.1	4.9E-324	20
58	SC_GLGS_ARATH.1.1	0,012	5670130	4431002	4>9	1.28	B8LNV7	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)	99.1	4.9E-324	15
37	SC_MDARS_CUCSA.1.10	0.047	15121628	13341808	4>9	1.13	A5JPK7	Monodehydroascorbate reductase	99.1	4.9E-324	13
29	SC_LOC100274026_MAIZE.1.3	0.030	6178574	5151285	4>9	1.20	Q6ZI55	NAD-dependent isocitrate dehydrogenase c;2 (Os02g0595500 protein) (Putative NAD-dependent isocitrate dehydrogenase)	99.1	4.9E-324	6
28	SC_PGKH1_ARATH.1.6	0.017	21775721	18129305	4>9	1.20	F4MKM1	Phosphoglycerate kinase (EC 2.7.2.3)	99.0	1.25E-223	5
31	SC_PGKH1_ARATH.1.6	0.047	15663900	13674117	4>9	1.15	F4MKM1	Phosphoglycerate kinase (EC 2.7.2.3)	99.0	1.13E-223	5
75	SC_PREP1_ARATH.1.3	0,045	2633810	2256228	4>9	1.17	Q9LJL3	Presequence protease 1, chloroplastic/mitochondrial (AtPreP1) (PreP 1) (EC 3.4.24.-) (Zinc metalloprotease 1) (AtZnMP1)	99.1	1.5E-259	10
50	SC_PDC2_ORYSI.1.19	0.011	4060820	3304815	4>9	1.23	A2XFI3	Pyruvate decarboxylase isozyme 2 (PDC) (EC 4.1.1.1)	98.9	9.75E-214	4
51	SC_PDC2_ORYSI.1.19	0.040	5721329	4845971	4>9	1.18	A2XFI3	Pyruvate decarboxylase isozyme 2 (PDC) (EC 4.1.1.1)	99.1	4.9E-324	11

52	SC_PDC2_ORYSI.3.19	0.007	6328814	5130352	4>9	1.23	Q10MW3	Pyruvate decarboxylase isozyme 2 (PDC) (EC 4.1.1.1)	99	4.9E-324	7
53	SC_PDC2_ORYSI.3.19	0.019	1822563	1382152	4>9	1.32	Q10MW3	Pyruvate decarboxylase isozyme 2 (PDC) (EC 4.1.1.1)	99	1.9E-322	8
44	SC_METK1_PICSI.34.55	0.003	3420898	2615036	4>9	1.31	A9NUH8	S-adenosylmethionine synthase 1 (AdoMet synthase 1) (EC 2.5.1.6) (Methionine adenosyltransferase 1) (MAT 1)	99	4.9E-324	6
46	SC_00154823.1.1	0.045	16788470	14777160	4>9	1.14	Q6RK07	UDP-glucose 6-dehydrogenase (EC 1.1.1.22)	99.1	4.9E-324	10
76	SC_DPE2_ARATH.1.1	0.025	3144515	3556831	9>4	1.13	Q69Q02	4-alpha-glucanotransferase DPE2 (EC 2.4.1.25) (Amylomaltase) (Disproportionating enzyme) (D-enzyme) (Protein DISPROPORTIONATING ENZYME 2)	-	-	-
43	SC_00000649.1.1	0.024	13865879	15890116	9>4	1.15	B9RVA7	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)	99.1	4.9E-324	24
40	SC_00000578.1.1	0.006	1976653	2539816	9>4	1.28	O23960	Acetyl-CoA carboxylase (EC 6.4.1.2) (Biotin carboxylase) (EC 6.3.4.14) (Uncharacterized protein)	99.1	4.86E-263	8
42	SC_SAHH_NICSY.1.26	0.022	10267479	12978095	9>4	1.26	B8LLL7	Adenosylhomocysteinase (EC 3.3.1.1)	99.1	4.9E-324	25
35	SC_ADH1_PETHY.12.16	0.030	6826221	8498742	9>4	1.25	Q43300	Alcohol dehydrogenase (EC 1.1.1.1) (Fragment)	99.1	4.9E-324	11
6	SC_00003988.1.1	0.038	4303160	5448471	9>4	1.27	Q9LUG8	AT3g23600/MDB19_9 (Contains similarity to endo-1,3-1,4-beta-D-glucanase) (Dienelactone hydrolase family protein) (Similarity to endo-1)	98.5	2.1E-114	4
22	SC_00002431.1.1	0.050	12132476	13642890	9>4	1.12	A3F7Q3	Benzoquinone reductase	98.9	1.9E-186	5
70	SC_CALR_RICCO.2.19	0.001	7102119	10464008	9>4	1.47	Q9FYV2	Calreticulin	99.1	4.9E-324	15
69	SC_CALR_RICCO.2.19	0.009	9973980	13389235	9>4	1.34	Q9FYV2	Calreticulin	99.1	4.9E-324	9
49	SC_CH62_CUCMA.1.9	0.027	3237970	3825086	9>4	1,18	Q05046	Chaperonin CPN60-2, mitochondrial (HSP60-2)	99.0	4.46E-172	6
17	SC_00004509.1.1	0.004	8881755	11145323	9>4	1.25	Q39626	Expansin S2	-	-	-

26	SC_00151055.1.1	0.023	49636921	56671922	9>4	1.14	P42742	proteasome subunit beta type-1	99.1	4.9E-324	16
27	SC_00151055.1.1	0.015	15633321	18505788	9>4	1.18	P42742	proteasome subunit beta type-1	99.1	4.9E-324	13
64	SC_00000119.1.1	0.004	2408421	3939123	9>4	1.64	Q93YR3	FAM10 family protein At4g22670	98.9	2.34E-193	4
4	SC_GLGL1_BETVU.1.2	0.039	14794002	19404652	9>4	1.31	P55233	Glucose-1-phosphate adenylyltransferase large subunit, chloroplastic/amyloplastic (EC 2.7.7.27) (ADP-glucose pyrophosphorylase) (ADP-glucose synthase) (AGPase S) (Alpha-D-glucose-1-phosphate adenyl transferase)	98.5	2.54E-149	3
24	SC_00007308.1.1	0.010	7494340	10118978	9>4	1.35	P41918	GTP-binding nuclear protein Ran-A1	99.0	1.78E-317	8
47	SC_ILV5_SPIOL.1.21	0.020	5378699	6306174	9>4	1.17	O82043	Ketol-acid reductoisomerase, chloroplastic (EC 1.1.1.86) (Acetohydroxy-acid reductoisomerase) (Alpha-keto-beta-hydroxylacil reductoisomerase)	99.1	4.9E-324	14
9	SC_UMC2374_MAIZE.1.3	0.022	9788184	11563722	9>4	1.18	Q69NG5	Probable 6-phosphogluconolactonase 4, chloroplastic (6PGL 4) (EC 3.1.1.31)	99.1	4.9E-324	8
1	SC_PROF1_PHAVU.6.6	0.007	12482534	15756230	9>4	1.26	A9NMR7	Profilin	98.2	1.3E-158	5
71	SC_PDI14_ARATH.1.13	0.037	16904496	19775548	9>4	1.17	A7KY78	Protein disulfide isomerase (Fragment)	99.1	4.9E-324	19
72	SC_PDI14_ARATH.1.13	0.030	7728059	8951249	9>4	1.16	A7KY78	Protein disulfide isomerase (Fragment)	99.1	4.9E-324	17
57	SC_PDI23_ORYSJ.1.5	0.043	1872583	2356422	9>4	1.26	A9CPA7	Protein disulfide isomerase family (Protein disulfide isomerse like protein)	98.5	1.1E-148	4
81	SC_XYL1_ARATH.1.6	0.045	3571950	5079450	9>4	1.42	Q8VWV9	Putative alpha-xylosidase	99.1	4.9E-324	12
16	SC_EXB14_ORYSJ.1.2	0.025	5222297	6461219	9>4	1.24	Q6H677	Putative expansin-B14 (Beta-expansin-14) (OsEXPB14) (OsaEXPb1.18)	-	-	-
23	SC_00146588.1.1	0.001	6949476	9332516	9>4	1.34	Q652Q1	Putative germin –like protein 9-1	99.0	3.76E-284	7
15	SC_SODF_NICPL.4.4	0.050	4539477	5957294	9>4	1.31	Q6QJL0	Superoxide dismutase (EC 1.15.1.1)	98.7	1.8E-184	7

21	SC_SODM_NICPL.1.5	0.010	12622569	15576757	9>4	1.23	A9NNA1	Superoxide dismutase (EC 1.15.1.1)	99.1	4.9E-324	11
66	SC_TBB2_ELEIN.1.4	0.050	9492956	13068855	9>4	1.38	Q9ZPN9	Tubulin beta-2 chain (Beta-2-tubulin)	99.1	4.9E-324	21
8	SC_00009054.1.1	0.040	4799378	6821333	9>4	1.42	O04905	UMP/CMP kinase (UMP/CMPK) (EC 2.7.4.14) (Uridine monophosphate/cytidine monophosphate kinase) (Uridylate kinase) (UK)	98.4	9.91E-194	4



Appendix S5. Assignment of the putative proteins identified in immature somatic embryos of maritime pine matured for 1 week to functional categories, according to Blast2GO (http://www.blast2go.com/b2ghome). For each category, the total number of proteins is indicated in parentheses. Proteins involved in more than one process were assigned to more than one category. Hence the sum of the numbers of proteins in the different classes exceeds the total number of proteins identified.

CHAPITRE III – Les embryons somatiques cotylédonaires de *Pinus pinaster* [Ait.] se développent de façon similaire à des embryons zygotiques cotylédonaires frais : approches physiologique, biochimique et protéomique

III.1 Présentation de la publication

Chez les plantes à graines orthodoxes, la maturation de l'embryon zygotique est considérée comme complète quand celui-ci atteint son maximum de poids sec (formation des organes de réserve que sont les cotylédons et accumulation des réserves), et que la déshydratation est initiée. Un embryon zygotique de qualité se définit par sa capacité à survivre (résistance à la déshydratation, avec parfois entrée en dormance), à germer lorsque les conditions sont favorables et à régénérer une plante vigoureuse et autonome. Les embryons somatiques passent par les mêmes étapes de développement morphologique qu'un embryon zygotique ; ils maturent dans des conditions de culture qui tendent de reproduire (mimer) les conditions de développement *in situ* de leurs homologues zygotiques (Ikeda et Kamada, 2006). Chez le pin maritime, après 12 semaines de maturation l'ES cotylédonaire est prélevé pour obtenir par la suite sa germination et son développement en plant (Lelu-Walter et al., 2006). Cependant comparés à un semis, les plants régénérés par embryogenèse somatique peuvent sur le terrain présenter un manque de vigueur initiale et un retard de croissance (Trontin et al., 2013). *A priori* le problème vient de la mise en germination d'embryons somatiques de « mauvaise qualité », problème à ce jour ne pouvant être révélé qu'à *posteriori*. C'est pourquoi il convient de mieux définir la maturation de nos embryons somatiques.

La sélection des ES matures repose sur une morphologie comparable à celle d'un EZ. Comme pour l'EZ, les profils d'accumulation des composés de réserve (protéines, lipides et sucres) dans l'ES au cours de son développement conditionnent son devenir et renseignent directement sur la qualité de maturation.

Nos objectifs sont **1**) identifier et comprendre les processus physiologiques contrôlant la maturation de l'embryon somatique **2)** déterminer quel stade de maturité l'ES a atteint, comparativement à l'EZ, pour aider à la détermination du meilleur moment de prélever les ES. Une méthode de classification hiérarchique ascendante (HCA), basée sur les neuf variables explicatives mesurées, a été utilisée pour définir quel stade de développement de l'EZ se rapprochait le plus de l'ES de 12 semaines. Cette approche est innovante car jamais

entreprise à ce jour en embryogenèse somatique des conifères et **3**) établir une liste de marqueurs moléculaires potentiels de l'état de maturation de l'embryon cotylédonaires frais.
Nous avons donc choisis d'étudier la physiologie et la biochimique des ES cotylédonaires en fonction de la durée de leur maturation, soit après 10, 12 (témoin) et 14 semaines. Tous ces ES ont été caractérisés à un niveau quantitatif (teneur en eau, sucres solubles, protéines) et à un niveau qualitatif (nature des sucres, profil protéique). Le suivi de la teneur en eau comme l'apparition de certains sucres, tels que le raffinose et le stachyose (Gosslova et al., 2001; Pullman et Buchanan, 2008), nous renseignera sur l'état de déshydratation éventuel de l'ES. Le contenu en sucres et les rapports relatifs qui s'établissent entre eux (en particulier, le ratio hexoses/saccharose et RFO/saccharose) au cours du développement embryonnaire sont des bons marqueurs de la maturité de l'ES (Gosslova et al., 2001; Konradova et al., 2002; Businge et al., 2013). De même, les protéines de réserve sont considérées comme des marqueurs de la phase de maturation (Hughes et Galau, 1991; Tereso et al., 2007) ; leur accumulation est suivie par l'accumulation de nombreuses protéines de l'embryogenèse tardive (LEA) (Dure et al., 1989).

La comparaison des ES et EZ cotylédonaires du pin maritime a déjà été entreprise par des analyses physiologiques et biochimiques (Jordy et Favre, 2003; Breton et al., 2006; Perez Rodriguez et al., 2006), et montraient des différences d'accumulation de réserves. Nous avons choisi de caractériser sept stades de maturation de l'EZ (répartis entre juillet et décembre 2011) quant à leurs teneurs en eau, en sucres solubles, en protéines totales, et en protéines de réserve. Pour aller plus loin dans notre description de la qualité de l'ES cotylédonaire et en s'appuyant sur notre expérience du mélèze hybride (Teyssier et al., 2014), nous avons comparé les protéomes de l'ES cotylédonaire de 12 semaines et celui de son équivalent zygotique. La protéomique étant une technique résolutive qui permet d'obtenir une image des acteurs moléculaires contenus dans les embryons à un instant donné.

Nous avons travaillé sur deux lignées embryogènes d'origine génétique différente pour une meilleure validation des marqueurs de maturité. Le plan expérimental, avec les analyses réalisées, est décrit dans la **Figure III.1**.

L'intégration des données de caractérisation physiologique et biochimique à l'étude protéomique permettra d'avoir une vision plus complète des processus impliqués dans la maturation de l'ES. La comparaison des embryons somatiques et zygotiques à leur stade

mature chez le pin maritime n'a jamais été réalisée à ce jour et constitue l'aspect original du travail.

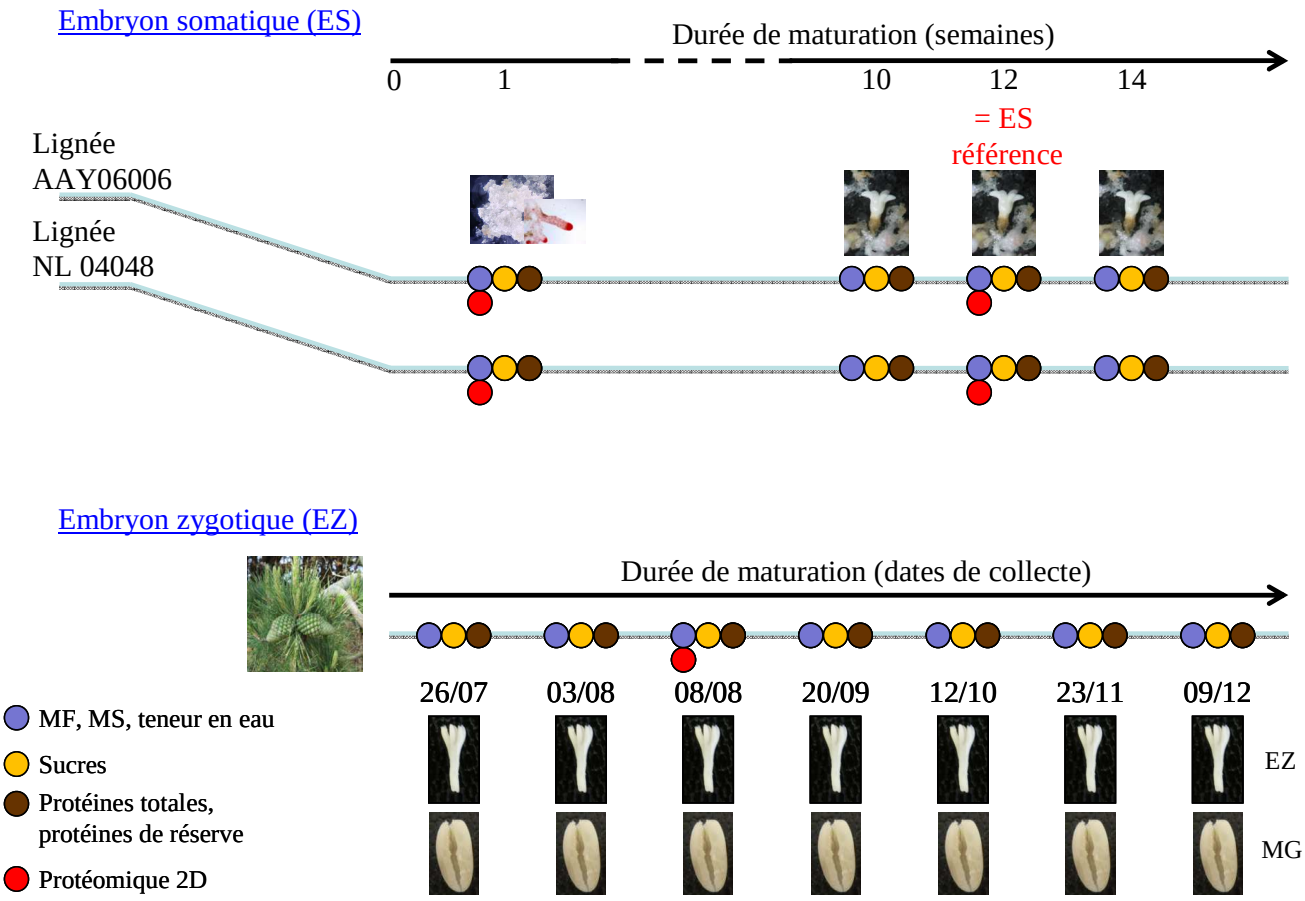


Figure III.1. Plan expérimental des analyses réalisées pour la question de recherche 2 : *Evaluation de la qualité de développement de l'ES cotylédonaire.* **ES :** embryon somatique ; **EZ :** embryon zygotique ; **MG :** mégagamétophyte ; **MF :** masse fraîche ; **MS :** masse sèche

III.2 Publication

Cotyledonary somatic embryos of *Pinus pinaster* Ait. are more similar to fresh maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses.

Article accepté par *Planta*

Alexandre Morel[1], Jean-François Trontin[2], Françoise Corbineau[3], Anne-Marie Lomenech[4], Martine Beaufour[5], Isabelle Reymond[2], Claire Le Metté[1], Kevin Ader [1], Luc Harvengt[2], Martine Cadene[5], Philippe Label[6], Caroline Teyssier[1]†, Marie-Anne Lelu-Walter[1]†**

1: INRA, UR 0588 Unité Amélioration, Génétique et Physiologie Forestières, 2163 Avenue de la Pomme de Pin, CS 4001, Ardon, F-45075 Orléans Cedex 2, France
2: FCBA, Pôle Biotechnologie et Sylviculture Avancée, Equipe Génétique et Biotechnologie, Campus Forêt-Bois de Pierroton, 71 route d'Arcachon, F-33610 Cestas, France
3 : Sorbonne Universités, UPMC Univ Paris 06, UMR 7622, Biologie du Développement, 4 place Jussieu, F-75005, Paris, France
4: Univ. Bordeaux, Centre Génomique Fonctionnelle, Plateforme Protéome, F-33000 Bordeaux, France
5: CNRS, UPR4301, Centre de Biophysique Moléculaire, Rue Charles Sadron, F-45100 Orléans, France.
6: INRA-Université Blaise Pascal, UMR_A 547 PIAF, Les Cézeaux, 24 Avenue des Landais, 63177 Aubière cedex, France.

† These authors contributed equally to this work

** Corresponding author: Marie-Anne Lelu-Walter
Phone : +33 (0) 238417800
Fax : +33 (0) 238417879
e-mail : marie-anne.lelu-walter@orleans.inra.fr

Planta

Cotyledonary somatic embryos of *Pinus pinaster* Ait. most closely resemble fresh, maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses

--Manuscript Draft—Corrected version 17-05-2014

Manuscript Number:	**PLAA-D-14-00183 R1**
Full Title:	Cotyledonary somatic embryos of Pinus pinaster Ait. most closely resemble fresh, maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses
Article Type:	Original Paper
Corresponding Author:	Marie-Anne Lelu-Walter, phD INRA Orleans Cedex 2, FRANCE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	INRA
Corresponding Author's Secondary Institution:	
First Author:	Alexandre MOREL, phD
First Author Secondary Information:	
Order of Authors:	Alexandre MOREL, phD
	Jean-François TRONTIN, phD
	Françoise CORBINEAU, Professor
	Anne-Marie LOMENECH, phD
	Martine BEAUFOUR, phD
	Isabelle REYMOND, Mrs
	Claire LE METTE, Mrs
	Kevin ADER, Mr
	Luc HARVENGT, phD
	Martine CADENE, phD
	Philippe LABEL, phD
	Caroline TEYSSIER, phD
	Marie-Anne Lelu-Walter, phD
Order of Authors Secondary Information:	
Abstract:	Cotyledonary somatic embryos (SEs) of maritime pine are routinely matured for 12 weeks before being germinated and converted to plantlets. Although regeneration success is highly dependent on SEs quality, the date of harvesting is currently determined mainly on the basis of morphological features. This empirical method does not provide any accurate information about embryo quality with respect to storage compounds (proteins, carbohydrates). We first analyzed SEs matured for 10, 12 and 14 weeks by carrying out biological (dry weight, water content) and biochemical measurements (total protein and carbohydrate contents). No difference could be found between collection dates, suggesting that harvesting SEs after 12 weeks is appropriate. Cotyledonary SEs were then compared to various stages, from fresh to fully desiccated, in the development of cotyledonary zygotic embryos (ZEs). We identified profiles that were similar using hierarchical ascendant cluster analysis (HCA). Fresh and dehydrated ZEs could be distinguished, and SEs clustered with fresh ZEs.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	Both types of embryo exhibited similar carbohydrate and protein content and signatures. This high level of similarity (94.5%) was further supported by proteome profiling. Highly expressed proteins included storage, stress-related, late embryogenesis abundant and energy metabolism proteins. By comparing overexpressed proteins in developing and cotyledonary SEs or ZEs, some (23 proteins) could be identified as candidate biomarkers for the late, cotyledonary stage. This is the first report of useful generic protein markers for monitoring embryo development in maritime pine. Our results also suggest that improvements of SEs quality may be achieved if the current maturation conditions are refined.

Cotyledonary somatic embryos of *Pinus pinaster* Ait. most closely resemble fresh, maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses.

Alexandre Morel[1], Jean-François Trontin[2], Françoise Corbineau[3], Anne-Marie Lomenech[4], Martine Beaufour[5], Isabelle Reymond[2], Claire Le Metté[1], Kevin Ader[1], Luc Harvengt[2], Martine Cadene[5], Philippe Label[6], Caroline Teyssier[1]†, Marie-Anne Lelu-Walter[1]†**

1: INRA, UR 0588 Unité Amélioration, Génétique et Physiologie Forestières, 2163 Avenue de la Pomme de Pin, CS 4001, Ardon, F-45075 Orléans Cedex 2, France

2: FCBA, Pôle Biotechnologie et Sylviculture Avancée, Equipe Génétique et Biotechnologie, Campus Forêt-Bois de Pierroton, 71 route d'Arcachon, F-33610 Cestas, France

3 : Sorbonne Universités, Université Pierre et Marie Curie-Paris 6, UMR 7622 CNRS-UPMC, Biologie des semences, Boîte Courrier 24, Bât. C, 2ème étage, 4 place Jussieu, F-75005, Paris, France

4: Université de Bordeaux, Centre Génomique Fonctionnelle, Plateforme Protéome, F-33000 Bordeaux, France

5: CNRS, UPR4301, Centre de Biophysique Moléculaire, Rue Charles Sadron, F-45100 Orléans, France.

6: INRA-Université Blaise Pascal, UMR_A 547 PIAF, Les Cézeaux, 24 Avenue des Landais, 63177 Aubière cedex, France.

† These authors contributed equally to this work

** Corresponding author: Marie-Anne Lelu-Walter

Phone : +33 (0) 238417800

Fax : +33 (0) 238417879

e-mail : marie-anne.lelu-walter@orleans.inra.fr

Abstract
Cotyledonary somatic embryos (SEs) of maritime pine are routinely matured for 12 weeks before being germinated and converted to plantlets. Although regeneration success is highly dependent on SEs quality, the date of harvesting is currently determined mainly on the basis of morphological features. This empirical method does not provide any accurate information about embryo quality with respect to storage compounds (proteins, carbohydrates). We first analyzed SEs matured for 10, 12 and 14 weeks by carrying out biological (dry weight, water content) and biochemical measurements (total protein and carbohydrate contents). No difference could be found between collection dates, suggesting that harvesting SEs after 12 weeks is appropriate. Cotyledonary SEs were then compared to various stages, from fresh to fully desiccated, in the development of cotyledonary zygotic embryos (ZEs). We identified profiles that were similar using hierarchical ascendant cluster analysis (HCA). Fresh and dehydrated ZEs could be distinguished, and SEs clustered with fresh ZEs. Both types of embryo exhibited similar carbohydrate and protein content and signatures. This high level of similarity (94.5%) was further supported by proteome profiling. Highly expressed proteins included storage, stress-related, late embryogenesis abundant and energy metabolism proteins. By comparing overexpressed proteins in developing and cotyledonary SEs or ZEs, some (23 proteins) could be identified as candidate biomarkers for the late, cotyledonary stage. This is the first report of useful generic protein markers for monitoring embryo development in maritime pine. Our results also suggest that improvements of SEs quality may be achieved if the current maturation conditions are refined.

Key words
Carbohydrates; 2-D gel electrophoresis; Embryo quality; Maturation; Protein markers proteome; Somatic embryogenesis; Zygotic embryogenesis.

Abbreviations
ABA, abscisic acid; CI, confidence interval; DW, dry weight; EMs, embryonal masses; Fru, fructose; FW, fresh weight; Glc, Glucose; HSP, heat shock proteins; LEA, late embryogenesis abundant; SE, somatic embryo; ZE, zygotic embryo

Introduction

Softwood species (mainly conifers) accounted for more than 77 % of total timber harvests in Europe in 2011 (375 million m^3). This constitutes the main source of timber supplying an industry which nevertheless continued to be a net importer of both roundwood and pulpwood during that period. In the context of increasing demand for softwood logs worldwide, an important priority for maritime pine (*Pinus pinaster* Ait.), a major species with broad industrial applications (Sanz et al. 2006), is to accelerate the development of more productive varieties for use in sustainable forest plantation.

One way to facilitate the capture of the best genetic stocks from breeding populations is to increase the efficiency of selection for elite genotypes. This can be achieved through early identification of genes (marker-assisted selection), prediction of breeding values (genomic selection) or field comparison of vegetative propagules from different trees (clonal testing). Vegetative propagation is also an effective way of deploying better adapted varieties in long-term plantation forestry, as it facilitates utilization of the whole genetic variance available for a trait of interest, i.e. both additive and non-additive genetic variance. Traditional genetic evaluations approaches, and even the more recently developed genomic prediction methods, are usually based on models that either ignore, or give only an estimate of, non-additive effects (Su et al. 2012). Because of these limitations, conifer breeding would greatly benefit from the implementation of vegetative propagation methods both for selection and for the deployment (Park 2002) in commercial plantation forestry of varieties derived from individuals selected for desirable characteristics. This process is referred to as clonal forestry or, more descriptively, multivarietal forestry. Conifers are usually recalcitrant to clonal propagation (Bonga et al. 2010), in contrast to those tree species (e.g. poplar) that respond spontaneously and cost-effectively to classical methods (e.g. cuttings), even when trees are in their adult vegetative or reproductive phase. Major advances in forest biotechnology have considerably extended the opportunities available for the clonal propagation of woody trees (reviewed in Nehra et al. 2005), especially for somatic embryogenesis in conifers (reviewed in Lelu-Walter et al. 2013).

Somatic embryogenesis provides a powerful tool which offers clear advantages for the rapid renewal of genetically improved varieties and for scaling up production. The process is often restricted to the use of juvenile explants (various stages in the development of zygotic embryos, ZEs) but the amenability of embryogenic tissue

to long-term storage at very low temperature (-196 °C) has opened up the way towards efficient, retroactive clonal propagation of selected, mature trees. In addition, cryopreservation of embryogenic tissue makes it possible to establish long-term collections of improved, native and/or endangered genetic resources in a juvenile state. Somatic embryogenesis is also considered to be a key technique for plant regeneration in research using reverse genetics through genetic transformation of conifers (Trontin et al. 2013). In forest trees, functional genomics is making an increasing contribution to the discovery of markers associated with valuable traits such as wood quality (Mauriat et al. 2014). This is a prerequisite for the development of more efficient methods of selection for breeding programs (e.g. marker-assisted selection). For these reasons, somatic embryogenesis has gradually become the preferred method for the clonal propagation of conifers (Lelu-Walter and Pâques 2009; Hargreaves et al. 2009; Lelu-Walter et al. 2013).

The past decade has seen significant progress in the elucidation of intricacies in the physiology and molecular basis of somatic embryogenesis in both angiosperms (reviewed in Yang and Zhang 2010; Zavattieri et al. 2010) and gymnosperms species (reviewed in Cairney and Pullman 2007; Vestman et al. 2011). Although many genes involved in the initiation and development of somatic embryos (SEs) have been identified and functionally characterized, the basis of the molecular regulation of somatic embryogenesis remains largely unknown (Elhiti et al. 2013).

In maritime pine, somatic embryogenesis was first reported by Jarlet-Hugues (1989), with effective plant regeneration being described 10 years later (Lelu et al. 1999). Extensive research on various steps in the process has since been undertaken in France (Ramarosandratana et al. 2001; Breton et al. 2005; Lelu-Walter et al. 2006), Portugal (Miguel et al. 2004), and Spain (Humánez et al. 2012). Improved protocols are now available for stages ranging from the initiation and proliferation of embryogenic tissue (Park et al. 2006) to somatic plant regeneration (Lelu-Walter et al. 2006). Field trials established in France from somatic plant material have consistently revealed that SEs develop at a lower initial growth rate than that of control seedlings (Trontin et al. 2011). In maritime pine high initial growth of seedlings is vital to cope with rapid weed development. A better understanding of embryo maturation in maritime pine is therefore required in order to produce high-quality, vigorous somatic plants. Current maturation protocols usually result in the development of cotyledonary SEs within 12 weeks. The optimal duration of maturation has hitherto been determined mainly on the basis of morphological features (i.e. visual selection of SEs resembling cotyledonary ZEs), culture productivity (yield of cotyledonary SEs) and the ability of SEs to germinate and convert into plantlets. However, this empirical approach does not provide any accurate information about the quality of the harvested SEs with respect to critical factors such as water content and the accumulation of storage reserves. Certain histological, histochemical or molecular differences between cotyledonary SEs and ZEs have been reported previously (Jordy and Favre 2003; Pérez-Rodriguez et al. 2006; Tereso et al. 2007). These results suggested that harvested SEs could be either substantially different from SEs, or correspond to any of the specific stages in the development of cotyledonary ZEs, from maturation to desiccation. In orthodox seeds, full embryo maturity results from the acquisition of desiccation tolerance during maturation followed by a desiccation phase (Kermode 1990).

There is therefore a need to identify biochemical and/or molecular markers for i) monitoring the development and the quality of cotyledonary SEs during maturation and ii) implementing alternative procedures which will allow SEs to attain the quality of ZEs and thus facilitate optimal conversion to field-grown plants. Qualitative and quantitative variations in carbohydrates and proteins (including the accumulation of storage reserves) are key factors contributing to the whole embryogenesis process, especially cotyledonary embryo development and maturity (reviewed in Lipavská and Konrádová 2004). One approach to discovering markers related to embryo quality is to establish the pattern of accumulation of these compounds in maturing cotyledonary ZEs. Interpretation of such patterns could result in the discovery of biochemical markers (carbohydrates and protein content) and molecular markers (expressed genes and proteins) which could be used to monitor the maturation of SEs.

The pattern of carbohydrate accumulation during the maturation of ZEs has been reported in conifer species such as *Picea abies* (Konradova et al. 2002) and *Pinus taeda* (Pullman and Buchanan 2008). Detailed comparative analyses of developmental stages in ZEs and SEs have been reported only in *Picea abies* (Gösslová et al. 2001; Konradova et al. 2002). Significant variations in carbohydrate profiles were observed during the development of SEs compared to ZEs. These differences may be due to the specific carbohydrate composition of the maturation medium, which has been shown to affect storage reserve accumulation and germination in SEs (Businge et al. 2013).

Storage protein gene expression and accumulation during SE and ZE maturation have been described for conifer species including *Picea abies* (Hakman 1993), *Pinus strobus* (Klimaswezska et al. 2004) and *Pinus taeda* (Brownfield et al. 2007). Vicilin- and legumin-like proteins are usually considered to be the main storage proteins in conifers (Klimaszewska et al. 2004). Protein content is often lower in SEs than in ZEs; however, the content in ZEs is highly dependent upon maturation conditions and developmental stage. Storage protein

concentration has been put forward as a putative biochemical marker of SE quality in pine species such as *Pinus strobus* (Klimaszewska et al. 2004), *Pinus sylvestris* (Lelu-Walter et al. 2008) and *Pinus pinaster* (Tereso et al. 2007).
The development of novel methods for the analysis of genome-wide protein variations (proteomics) has further increased the possibility of systematic screening for proteins putatively involved in embryo maturation and quality. In recent years, proteomics has significantly contributed to an improved understanding of plant somatic embryogenesis (Ehliti et al. 2013; Rocha and Domelas 2013). In conifers, changes in protein expression have been reported to occur during SE maturation in *Cupressus sempervirens* (Sallandrouze et al. 1999), *Picea glauca* (Lippert et al. 2005), and *Larix* x *eurolepis* (Teyssier et al. 2011, 2014). Early changes in protein expression during maturation of SEs were also recently studied in maritime pine (Morel et al. 2014), making it possible to propose predictive protein markers for SE development and for the adaptive response of a culture to maturation conditions. However, none of these proteomic studies included a comparison with zygotic embryogenesis as a reference, nor did they provide information which could lead to methods making it possible to produce SEs of a standard equal to that of cotyledonary, mature ZEs.
In the work presented here, we performed a comparative study of cotyledonary SEs matured for 10-14 weeks and cotyledonary ZEs at various stages of maturity, from fresh embryos collected from green cones to fully developed, desiccated embryos collected from mature cones. Using hierarchical ascendant cluster analysis of quantitative data (content of water, carbohydrates and total protein), we demonstrated that cotyledonary SEs matured for 12 weeks are more similar to fresh cotyledonary ZEs than to ZEs at other stages of development. Proteome profiling confirmed that these two embryo types are roughly equivalent, containing similar cohorts of overexpressed proteins. Some of these proteins are potential biomarkers for the cotyledonary stage of embryos. In addition to providing new insights into molecular events involved in the development of SEs in maritime pine, this is the first report of useful generic protein markers for monitoring embryo development in this species. Our results also suggest that improvements in the quality of maritime pine SEs may be expected if maturation conditions are refined to specifically promote desiccation tolerance in SEs before subjecting harvested embryos to a desiccation phase.

Material and methods

Plant material

Cotyledonary somatic embryos (SEs)

Experiments were conducted with two embryogenic lines of *Pinus pinaster* (FCBA collection) of contrasting genetic origin, a “Landes x Morocco” polycross family (line AAY06006) and a “Landes x Landes” fullsib family (line NL04048). Both AAY06006 and NL04048 are among our best and similarly performing lines (rate ± 95% CI) for germination (91.3 ± 0.9% and 92.8 ± 5.7%, respectively), acclimatization (74.0 ± 5.3% vs. 67.5 ± 6.9%) and final conversion rate of harvested cotyledonary SEs into field-transferable plants (67.6 ± 4.7% vs. 62.9 ± 8.2%).
Embryonal Masses (EMs) initiated from immature ZEs (Park et al. 2006) were proliferated every two weeks in darkness (23 ± 1 °C) using a modified Litvay (mLV) basal formulation (Litvay et al. 1985). The proliferation medium was mLV supplemented with 3 g l^{-1} gellan gum (Gelrite, Sigma, St Louis, MO, USA), 0.087 M sucrose, 2 µM 2,4-dichlorophenoxyacetic acid (2,4-D, Sigma) and 1 µM benzyladenine (BA, Sigma). The EMs were matured in darkness (23 ± 1°C) according to Lelu-Walter et al. (2006). Briefly, vigorous EMs grown for one week on proliferation medium were dissociated in mLV containing no growth regulator, and the resulting cell suspension was poured over a filter paper disc (Whatman N°2, 70 mm in diameter) using a Büchner funnel. The filter with its dissociated EMs, approximately 100 mg fresh weight (FW) per filter, was transferred onto mLV maturation medium supplemented with 0.2 M sucrose, 80 µM cis-trans (±) abscisic acid (ABA, Sigma) and 9 g l^{-1} gellan gum (Gelrite, Sigma). After 12 weeks of maturation, cotyledonary SEs were counted to estimate the embryogenic potential of the EMs (expressed as number of SEs per g FW). Cotyledonary SEs were also sampled after 10, 12 and 14 weeks for subsequent measurements.

Cotyledonary zygotic embryos (ZEs)

Open-pollinated cones from the AAY06006 mother tree were collected in Région Aquitaine, France, on 26 July, 3 and 8 August, 20 September, 12 October, 23 November and 09 December, 2011. The sample collection dates covered the full maturation phase, from recently differentiated cotyledonary ZEs in green cones (fresh ZEs) to mature ZEs in desiccated cones. Cones were stored in plastic bags at 4 °C until required for dissection. Seeds were extracted from the scales with sterile forceps (Lelu-Walter et al. 2006). On each collection date,

cotyledonary ZEs were excised from the surrounding megagametophyte. Samples were immediately frozen in liquid nitrogen prior to storage at -80 °C for subsequent protein and carbohydrate assays.

Dry weight and water content of SEs and ZEs
Samples (SEs, ZEs, approximately 100 mg FW) were immediately weighed to determine their FW. Dry weight (DW) was determined after oven-drying at 70 °C for 24 h and % DW was calculated by multiplying the DW to FW ratio by 100 (Teyssier et al. 2011). Water content was determined as (FW − DW)/ DW and expressed as g H_2O g^{-1} DW (Dronne et al. 1997). Five biological replicates of each embryo type and time-point were used for each assay.

Determination of carbohydrate and total protein contents

Carbohydrates
Soluble sugars were extracted as described by Black et al. (1996). Cotyledonary SEs and ZEs (approximately 40 mg FW) were ground in 1 ml 80 % aqueous ethanol containing 2.5 mg ml^{-1} melezitose as an internal standard, and the extracts were heated for 15 min at 80 °C. After 2 successive extractions with 80 % ethanol without melezitose, the supernatants were combined, and a 0.5 ml aliquot of the total extract was shaken with 0.5 ml water and 0.5 ml chloroform in order to remove lipids. An aliquot of the aqueous phase was freeze-dried and the dried extract was dissolved in 100 µl distilled water, filtered through a 0.45 µm pore filter (Nalgene, Rochester, NY, USA), and then analyzed by HPLC. Samples (10 µl) were injected and eluted with 80:20 (v/v) acetonitrile/H_2O at a flow rate of 1 ml min^{-1} using a Spectra Series P100 pump (Thermo Fisher Scientific, TFS, Illkirch, France) on a 200 x 4.6 mm Spherisorb-NH2 column (TFS). The eluents were analyzed with a differential refractometer (RI Plus Detector, TFS) and the peak areas were electronically integrated using ChromQuest 5.0 (TFS). The various sugars were identified by co-elution with standards (Sigma, St Louis, MO, USA). Results were expressed as µg mg^{-1} DW and corresponded to the mean of 3 measurements ± 95 % confidence interval (CI).

Total protein assay
Total protein extracts were prepared as five replicates for each type of sample (SEs, ZEs). Frozen material (25-50 mg FW) was homogenized with 0.5 ml of lysis buffer (10 % (v/v) glycerol; 2 % (w/v) SDS; 5 % (v/v) β-mercapto-ethanol; 2 % (w/v) poly(vinyl) polypyrrolidone; 50 mM Tris pH 6.8). The samples were incubated for 5 min at 95 °C, and then centrifuged for 10 min at 13,500 rpm. The resulting pellets were re-extracted once and the two supernatants were pooled. Protein concentration was determined using the Bradford assay with bovine serum albumin as a standard. Each of the results, which were expressed as µg mg^{-1} DW, corresponded to the mean of 5 measurements ± 95 % CI.

1D gel electrophoresis
The extracted proteins were separated by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 12 % gel with a stacking gel (4 %) according to standard protocols. The gels were stained for proteins with colloidal Coomassie Brilliant Blue G-250 (CBB-G). The main protein bands observed for SEs (12 weeks maturation) and ZEs (29 September), were cut from a 10 % gel (see excised bands Fig. 3) and subjected to LC-MS/MS identification (see the protocol applied to significant spots during 2-D gel analysis).

2-D gel proteomic analysis
Proteomic study design
Three samples were investigated during the first proteomic analysis: SEs from line AAY06006 collected after 1 (immature stage) and 12 weeks (cotyledonary stage) of maturation, and cotyledonary ZEs collected on 08 August. The second proteomic analysis included 4 samples: SEs collected after 1 (immature stage) and 12 weeks (cotyledonary stage) of maturation for both embryogenic lines (AAY06006 and NL04048).

2-D PAGE
Total protein extracts were prepared as 5 replicates from each sample (approximately 400 mg FW) as previously described in Teyssier et al. (2014). Briefly, after phenol precipitation (Isaacson et al. 2006), protein extracts were assayed as described in Teyssier et al. (2011). First dimension separation was performed with 300 µg protein loaded onto 24-cm IPG strips, pH 4-7 (Protean IEF Cell system, Bio-Rad, France), and the second dimension PAGE was performed with 11 % polyacrylamide gels. Gels were stained with colloidal CBB-G and analysed with the Progenesis software (Nonlinear Dynamics, United Kingdom) as described in Teyssier et al. (2014). The significance was set at a high level of expectation in the second proteomic study when comparing mature *vs*

immature SEs from both lines: only spots showing significant changes, i.e. at least 1.8-fold, and with a normalized spot volume of at least 5.10^6 were considered. This minimum spot volume requirement was not applied in the first proteomic study.

Mass spectrometry and data analysis

Statistically significant spots were excised from the gel and washed in water. Destain was performed with 150 µL wash buffer (50 mM ammonium bicarbonate in water/acetonitrile (50:50, v/v)). Gel pieces were then dried in acetonitrile for 10 min. After acetonitrile removal, digestion was performed by incubating the gel overnight at 37°C with trypsin at 10 ng μL^{-1} in ammonium bicarbonate 40 mM/ acetonitrile 10% (v/v). The resulting peptides were extracted from the gel by incubation once in ammonium bicarbonate 40 mM; acetonitrile 10% (v/v) and twice in H_2O/acetonitrile/formic acid 47.5:47.5:5 (by vol.). All three supernatants were pooled, concentrated down to 25 µL using a vacuum centrifuge and acidified with 1.5 µL of formic acid. Peptide mixtures were analysed by on-line capillary HPLC (LC Packings, Amsterdam, The Netherlands) coupled to an ion trap mass spectrometer (LCQ Deca XP Plus, Thermo-Finnigan, San Jose, CA, USA) in the case of 2-D PAGE spots or to an Orbitrap mass spectrometer (LTQ-Orbitrap XL, Thermo Scientific, Bremen, Germany) in the case of SDS-PAGE protein bands. Ten microliters of peptide digests were loaded onto a 300 µm × 5 mm C18 PepMap™ trap column (Dionex, The Netherlands) at a flow rate of 30 µL min^{-1}. The peptides were eluted from the trap column onto an analytical 75 µm × 15 cm C18 PepMap™ column (Dionex, The Netherlands) with a 5-40 % linear gradient of solvent B in 35 min (solvent A was 0.1 % formic acid in 5 % acetonitrile, and solvent B was 0.1 % formic acid in 80 % acetonitrile). The separation flow rate was set at 200 nl min^{-1}. Ion trap data were acquired in a data-dependent mode alternating a scan survey over the range *m/z* 300–1700 and 3 ion trap MS/MS scans with Collision Induced Dissociation (CID) as activation mode. MS/MS spectra were acquired using a 2-m/z unit ion isolation window and a normalized collision energy of 35. Dynamic exclusion duration was set to 30 s with a repeat count of 2. On the Orbitrap, the data-dependent cycle alternates a 60000-resolution MS scan survey over the range m/z 300–1700 and 6 CID MS/MS. Dynamic exclusion was also enabled for 30 s. Singly-charged ions were rejected from fragmentation.

Peptides were identified by running the SEQUEST™ algorithm through the Proteome Discoverer 1.3 interface (Thermo-Finnigan, Torrence, CA, USA) against a *Pinus pinaster* database consisting of GenoToul (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster3/index.html, following procedure described in Morel et al. 2014) sequences translated in all 6 reading frames (1277568 entries). From the assignments made, the corresponding proteins were identified by matching the m/z values of the peptide ion fragments against the GenoToul or Blast2GO databases.

Spectra from peptides with molecular masses greater than 5000 Da or less than 350 Da were rejected. The search parameters were as follows: the mass accuracy of the monoisotopic peptide precursor and peptide fragments was set to 2 Da and 1 Da, respectively, for ion trap MS data and to 10 ppm and 0.8 Da for Orbitrap MS data. Only b- and y-ions were considered for mass calculation. For spots, methionine oxidation (+16 Da) and cysteine carbamidomethylation (+57 Da) were allowed as variable modifications. In the case of bands, only methionine oxidation was allowed, again as a variable modification. In all cases, 2 missed trypsin cleavages were allowed. Only “high confidence” peptides were retained, corresponding to a 1 % false positive rate at the peptide level. A minimum of two different peptides was required for validation of protein identity.

Statistical analysis

Statistical analysis was carried out with R software (version 2.8.0; R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria). Effects of the treatments on the DW, water content, carbohydrate and total protein contents were evaluated using one-way ANOVA. Variations in these parameters during maturation were analysed using multiple comparisons of means with Tukey contrasts ($P < 0.05$). For 2-D PAGE analysis, the intensity change for each spot was analysed with Student's *t*-test on the basis of the normalized spot volume ($P < 0.001$). In order to compare stages in embryo development, we grouped similar profiles by means of a hierarchical ascendant cluster analysis (HCA) using the Ward’s minimum variance clustering method (Ward 1963), and the Euclidian distance as a measure of similarity. Dataset clustering was carried out using the results of all the assays carried out (DW, water content, fructose, glucose, sucrose, stachyose and raffinose content and total protein content). Each group produced by cluster analysis represented a set of assay results that showed a similar pattern of variation between different stages in embryo development.

Results

Various biological parameters (DW, water content, cotyledonary embryo yield) and quantitative and qualitative biochemical data (contents of carbohydrates and total proteins; proteomic ,profiles) were determined in cotyledonary SEs matured for 10, 12 or 14 weeks and maturing cotyledonary ZE sampled on 7 collection dates (from 26 July to 09 December). SEs and ZEs were produced from closely related genetic backgrounds (embryogenic line AAY06006 and half-sib progenies from the AAY06006 mother tree). To address any genetic bias and confirm the results obtained from AAY06006, we also collected data from one unrelated embryogenic line (NL04048).

Dry weight and water content of cotyledonary SEs and maturing cotyledonary ZEs

Transfer of EMs onto maturation medium resulted in the development of cotyledonary SEs after 10, 12 and 14 weeks for both embryogenic lines (AAY06006 and NL04048). After 12 weeks the embryogenic potential of AAY06006 was significantly higher than that of NL04048 (125.9 ± 47.3 and 49.5± 22.6, respectively, mean ± 95% CI). Throughout the maturation period (i.e. 10, 12 and 14 weeks) and regardless of line, cotyledonary SEs did not show any significant variation in DW (26-32 % FW, Figs 1a, S1a) or in water content (2.24-2.83 g H_2O g^{-1} DW, Fig. 1b), with the exception of line NL04048, which showed significantly lower water content after 10 weeks (2.24 g H_2O g^{-1} DW ± 0.05, Fig. 1b).

Cotyledonary ZEs were collected from open pollinated cones of AAY06006. During the period of harvesting (i.e. from 26 July to 09 December), a significant increase in DW was observed, from around 30 % (late July/early August) to 60-65 % (from 12 October to December), with a peak of about 75 % in September (Fig. 1a). Concomitantly, a significant decrease in water content was observed, from more than 2 g H_2O g^{-1} DW in early August to around 0.5 g H_2O g^{-1} DW from 12 October onwards (Fig. 1b). We concluded that SEs matured for 12 weeks were similar in DW and water content to fresh cotyledonary ZEs harvested before September. There was no significant difference between maturing cotyledonary ZEs harvested on 26 July and 03 August. By the end of the sampling period (December), ZEs showed significantly higher DW values (more than a 2-fold increase; Fig. 1a) and lower water content (4.5 times less) than SEs (Fig. 1b).

Carbohydrate content in cotyledonary SEs and maturing cotyledonary ZEs

Cotyledonary SEs from both lines matured for 10, 12 and 14 weeks contained similar amounts of several carbohydrates (Fig. 2a, Table S1), with sucrose (Su), which was in the range 11.75 to 14.91 µg mg^{-1} DW, being the most abundant. Monosaccharides such as fructose (Fru), glucose (Glc), galactose and the disaccharide maltose were also detected, but at comparatively low levels (Tables 1, S1). Of the oligosaccharides from the raffinose family (RFOs), raffinose was present at a very low level (around 0.4 µg mg^{-1} DW,) whereas stachyose was undetectable. As a result, the hexoses/Su and RFOs/Su ratios were both low and quite consistent (Fig. 2b, Tables 1, S1).

In contrast, significant variations were observed during maturation of cotyledonary ZEs (Fig. 2a, Table 1). Contents of galactose, maltose, fructose and sucrose (the main carbohydrates) decreased gradually from early August to 29 September and were then quite stable until December, showing no significant changes. Mature ZEs had very low and similar contents of galactose (0.19 µg mg^{-1} DW), maltose (0.13 µg mg^{-1} DW) and fructose (0.14 µg mg^{-1} DW), whereas sucrose was again the major carbohydrate (2.83 µg mg^{-1} DW, Table 1). Glucose was undetectable in fresh ZEs (July/August); it then slightly increased to levels similar to those observed for other minor carbohydrates in mature ZEs (0.06 µg mg^{-1} DW, Table 1). Interestingly, RFOs appeared during the course of ZE maturation, with raffinose first being detectable on 03 August and stachyose on 29 September. The RFOs/Su ratio significantly increased in cotyledonary ZEs from a low value in August (0.03-0.05) to 1.42 from October to December (Fig. 2b). The Fru+Glc/Su ratio followed the opposite trend (Table 1). SEs matured for 10-14 weeks were similar to fresh cotyledonary ZEs harvested before September with respect to the contents of sucrose, maltose and RFOs and to the RFOs/Su ratio (Figs 2a, 2b).

Quantitative and qualitative changes in the protein contents of cotyledonary SEs and maturing cotyledonary ZEs

Total protein content increased significantly in cotyledonary SEs of line AAY06006 as a function of the duration of maturation. SEs matured for 12-14 weeks accumulated more protein (276.7-292.4 µg mg^{-1}) than SEs matured for 10 weeks (223.6 µg mg^{-1}; Table 2). Total protein content is quite stable in SEs from line NL04048 (241.3-249.7 µg mg^{-1}, Table S2). In contrast, total protein content gradually increased during the maturation of cotyledonary ZEs to become significantly higher in November than at the earliest sampling date and reaching a maximum in December (357.1 µg mg^{-1} DW, Table 2). As in the case of the carbohydrate and FW/DW data, differences in total protein content between ZEs harvested from 08 August and 12 October, and ZEs harvested at

the end of seed development (09 December) were highly significant. On the basis of these total protein data, SEs matured for 10-14 weeks were again found to be similar to fresh cotyledonary ZEs (Table 2).
An electrophoretic analysis of total protein extracts was performed to produce a qualitative profile of protein content in SEs and ZEs. While no change in protein profile was observed among SEs matured for 10, 12 or 14 weeks (Fig. S2), there were significant differences among ZEs collected from September to December (Fig. 3). Numerous bands were detected in SEs, showing a protein profile very similar to that of fresh ZEs harvested in early August. A few major bands were observed in the profiles of SEs (12 weeks maturation) and fresh ZEs (August). The intensities of these bands increased in ZE profiles from late September to December. They were excised from the gels for further identification by LC-MS/MS. The same storage proteins were identified in both SEs and ZEs. They corresponded to vicilin-like protein, legumin-type protein, pine globulin-1 and/or albumin 3 (Table 3). Two of these protein identifications have been made by a unique peptide, but their MS/MS spectra of its +2 and/or +3 charge state validated them (Fig. S3).

Cotyledonary SEs *vs.* maturing cotyledonary ZEs: a synthesis of quantitative biological and biochemical data

Using the entire set of data available from line AAY06006 (DW, water content, carbohydrate and protein content), a hierarchical cluster analysis (HCA) was performed (Fig. 4) to make a tentative ranking of SE maturity (SE matured for 12 weeks) against the complete range of cotyledonary ZE stages. Two clusters were observed, corresponding to i) ZEs harvested from September (dehydrating ZEs) to December (mature ZEs) and ii) SEs matured for 12 weeks and ZEs harvested in late July and early August (fresh, maturing ZEs). This result further illustrated the fact that SEs matured for 12 weeks are more similar to fresh, maturing cotyledonary ZEs from green cones than to ZEs at other stages of development.
To summarize this part of the work, cotyledonary SEs from 2 genetically contrasted lines matured for 12 weeks (i.e. cotyledonary SEs at the stage at which they are routinely collected for germination) did not differ markedly from those matured for 10 or 14 weeks, according to various biological (DW, water content) and biochemical data (carbohydrate, protein contents). Using hierarchical cluster analysis of quantitative biological and biochemical data, we showed that cotyledonary SEs matured for 12 weeks are most similar to fresh, maturing cotyledonary ZEs collected in late July/early August (Fig. 4). A number of storage proteins expressed at similar levels in cotyledonary SEs and fresh cotyledonary ZEs have also been identified.

2-D gel proteomic analysis of cotyledonary SEs and fresh maturing cotyledonary ZEs

To gain further insights into putative qualitative differences between SEs and ZEs, we performed a proteomic comparison of cotyledonary SEs matured for 12 weeks (line AAY06006) with fresh maturing cotyledonary ZEs (harvested 08 August) from the same group, as revealed by the hierarchical cluster analysis (Fig. 4). A total of 976 reproducible spots were obtained in the corresponding 2-D gels (Fig. S4). Of the 52 spots that showed statistically significant differences ($P<0.001$, ratio ≥ 1.8 between the normalized spot volumes) between SEs and ZEs, 43 were successful at protein identification. A total of 17 of these corresponded to proteins overexpressed in SEs whereas 26 spots represented protein overexpressed in ZEs (Table 4). The proteins identified were classified into functional categories with Blast2GO (Fig. S5). Proteins involved in cellular processes, responses to stimulus and metabolic processes were well represented in both types of embryos. We concluded that there was a similar functional distribution among the proteins identified in cotyledonary SEs matured for 12 weeks old (AAY06006) and in fresh, maturing cotyledonary ZEs.

Identification of candidate protein markers for the fresh, cotyledonary stage of embryo

We were interested in the discovery of proteins that are specifically overexpressed in SEs at the cotyledonary stage and in ZEs that are of similar maturity according to our results (i.e. ZEs harvested on 08 August). Biomarker inference from the results of proteomic analysis at a specific stage is a complex task, as the proteome contains a high proportion of proteins with house-keeping functions. These housekeepers are present at all stages of maturation in embryos of either somatic or zygotic origin. To reduce the background noise arising from these proteins in somatic material, we performed a second proteomic study of immature (1 week maturation) and cotyledonary (12 weeks maturation) SEs to identify spots in 2-D gels corresponding to proteins overexpressed only at the cotyledonary stage. To obtain accurate and reliable candidate protein markers, we integrated data from both AAY06006 and NL04048 lines. A total of 1430 reproducible spots were detected, including 225 spots that were significantly ($P<0.001$) different between SEs matured for 1 week (immature SEs) or 12 weeks (cotyledonary SEs; Fig. S6). Lists of the proteins overexpressed after 1 week and 12 weeks can be found in Table S3 and Table 5, respectively. Among the significant spots over-expressed in SEs matured for 12 weeks, 68 proteins were identified for line AAY06006 and 72 for line NL04048. Most (43) of these spots were common to both lines and of these, 36 proteins could be identified (Table 5, shaded lines). Among these proteins considered

to be specific to the cotyledonary stage of SE development, we identified small numbers of storage proteins (cupin domain-containing protein, legumin-like storage protein, vicilin-like storage protein) and proteins related to embryo tolerance to desiccation (embryonic cell protein 63, late embryogenesis abundant (LEA)-like protein, embryonic protein DC-8, seed maturation protein), two proteins involved in germination (protein MOTHER of FT and TF 1, responsive to abscisic acid 28) and many heat shock proteins (HSP).
We then investigated the whole set of proteomic spots previously observed from fresh, maturing cotyledonary ZEs (Fig. S4). To tentatively eliminate those spots corresponding to house-keeping proteins, we investigated only spots that were significantly (P<0.001) different in ZEs when compared to immature SEs (AAY06006 at 1 week) that were analyzed during the same proteomic experiment (data not shown). We retained only spots corresponding to proteins overexpressed in ZEs. A total of 56 spots corresponding to 45 proteins were found to be overexpressed in cotyledonary ZEs.

By combining the list of the overexpressed spots revealed from our two proteomic studies, those of immature *vs* cotyledonary SEs (Table 5) and immature SEs *vs* fresh, cotyledonary ZEs (data not shown), we were able to identify robust putative generic markers of the fresh, cotyledonary stage of embryo maturation. The results are presented in Table 6. Of the 36 overexpressed protein spots identified in cotyledonary SEs (Table 5, shaded lines), 23 (around 80 %) were also overexpressed in fresh, cotyledonary ZEs. Most (14/23) of them were HSPs (HSP18.2, HSP60, HSP70-4, class II HSP17.6), LEA proteins (LEA-like protein, embryonic protein DC-8, embryonic cell protein 63) and storage proteins (vicilin-like storage protein, legumin-like storage protein, cupin domain-containing protein).

Discussion

Different embryogenic lines produced cotyledonary SEs of similar quality
The two embryogenic lines of maritime pine investigated in this study (AAY06006, NL04048) had different genetic backgrounds. It is therefore no surprise that maturation yield (production of cotyledonary SEs) was dependent on the embryogenic line used, as has been observed in virtually all conifer species (Park 2002). In maritime pine, the influence of genotype on recovery of SEs has been previously reported for Portuguese (Miguel et al. 2004) and French families (Lelu-Walter et al. 2006). Variation in maturation yield has been demonstrated within and between families (Miguel et al. 2004; Lelu-Walter et al. 2006). With regard to SE development, in our maturation conditions both lines matured in a similar way. We found no marked differences in biological (DW, water content) and biochemical parameters (carbohydrates, total protein content). It can be postulated from these data that maturation yield and intrinsic quality of SEs are unconnected in maritime pine as suggested on larger number of representative genotypes (67 lines from 24 families) showing no correlation between maturation yield and germination rate (Trontin unpublished results).

Quality of cotyledonary SEs did not vary as a function of the duration of maturation
Successful SE germination and conversion to plantlets depend on the recovery of high "quality" SEs. In hybrid larch, duration of maturation had a negative impact on subsequent SE desiccation, germination and plant recovery (Lelu et al. 1995). In Scots pine, the protein content of cotyledonary SEs drastically decreased when maturation duration was extended from 8 to 12 weeks (Lelu-Walter et al. 2008). In our optimized culture conditions for maritime pine, SEs matured for 10, 12 or 14 weeks were roughly similar with respect to DW (28.31 % FW on average), water content (2.56 g H_2O g^{-1} DW on average), total protein (254.2 µg mg^{-1} DW on average) and carbohydrate contents (sucrose being the major carbohydrate present). In addition, no changes in storage proteins could be detected in electrophoretic profiles. Using mass spectrometry, we identified some of these proteins for the first time in maritime pine, including 7S vicilin-like protein, 11S legumin-like protein, albumin 3 and pine globulin-1 protein. The latter has only previously been identified in *Pinus strobus* (Klimaszewska et al. 2004); it is quite similar to 11S legumin-like protein, which is a closely related member (though with a slight difference in size) of the globulin family. In angiosperm species, specific globulin-like proteins have been reported to accumulate in SEs; they include citrin in *Citrus sinensis* (Koltunow and Robinson 1996) and convicilin in *Pisum sativum* (Mihsra et al. 2012).
Interestingly, the water content of cotyledonary SEs remained at a similarly high level after 10, 12 or 14 weeks culture (2.2-2.8 g H_2O g^{-1} DW) on maturation medium supplemented with a high concentration of gellan gum (9 g l^{-1}). Comparable water contents were measured in cotyledonary SEs of *Larix* x *eurolepis* matured under high gellan gum conditions (8 g l^{-1}, Teyssier et al. 2011). High gellan gum concentration was shown to result in reduced availability of water from the culture medium in *Pinus strobus* (Klimaszewska et al. 2000) and *Larix* x *eurolepis* (Teyssier et al. 2011). Such conditions induce desiccation, with a concomitant increase in the DW of

maturing SEs and an improved maturation yield in hybrid larch (Lelu-Walter and Pâques 2009). Strikingly, low water availability during maturation was not perceived as a stressful condition for SE development in larch (Teyssier et al. 2011). In maritime pine, a high gellan gum concentration is a pre-requisite for embryo development. The induced desiccation is apparently stabilized at a stage as early as 10 weeks of maturation. In contrast to larch, however, in maritime pine specific stress-related pathways were shown to be activated at an early stage in SE development (1 week maturation; Morel et al. 2014) suggesting that our culture conditions impose some constraints linked with water availability and that there are species-specific requirements.

Cotyledonary SEs resemble fresh, maturing cotyledonary ZEs

Maturation of cotyledonary ZEs of maritime pine from July to December is characterized by a significant increase in DW (from 33.0 % FW in July/August to 67.8 % FW in September/December on average). This can be partly attributed to a concomitant increase in total protein content (from 241 µg mg^{-1} DW in August to 327 µg mg^{-1} DW in September/December on average). A decrease in the water content of ZEs was also clearly observed during the collection period (Fig. 1), and this can be interpreted as reflecting a transition from maturation to the desiccation phase. This transition apparently occurred mainly between 08 August and 29 September and it resulted in a greatly reduced water content (0.49 g H_2O g^{-1} DW in September/December on average), enhanced synthesis of storage proteins and increased content of carbohydrate members of the RFO family (stachyose, raffinose) together with a high RFO/ Su ratio (1.4). RFO accumulation was previously observed in ZEs from *Picea abies*, where it was considered to be a specific event occurring during late embryogenesis (Gösslová et al. 2001). Carbohydrates of the RFO family have also been regarded as being involved in desiccation tolerance (Keller and Ludlow 1993), as they have been observed in the seeds of numerous species (Black et al. 1996, 1999; Corbineau et al. 2000; Bailly et al. 2001). Induction of oligosaccharide accumulation generally occurs during the seed filling phase, and it has been shown to depend on the drying conditions used (Black et al. 1996; Corbineau *et al*. 2000). In maritime pine, results obtained for maturation of ZEs are markedly different from those for SEs, as no clear transition to the desiccation phase could be observed in cotyledonary SEs matured for 10-14 weeks.

A hierarchical cluster analysis of the entire set of quantitative data (DW, water content, carbohydrates and total protein contents) produced during this study resulted in separate groups for fresh and dehydrating ZEs. SEs were shown to be included in the "fresh ZEs" group, which corresponds to green, immature cones collected in late July to August. Interestingly, a transcriptomic analysis by Pullman et al. (2003) similarly demonstrated in *Pinus taeda* that gene expression in the most advanced cotyledonary SEs obtained after appropriate culture time on maturation medium did not conform to that of fully mature ZEs but to earlier, immature cotyledonary ZE stages.

Proteomics revealed similar protein profiles in SEs and in fresh ZEs

Proteomic analysis confirmed that there was a high degree of similarity between fresh ZEs collected on 08 August and cotyledonary SEs matured for 12 weeks. The number of significantly different spots was very low (52) and represented less than 6 % of the spots detected (976). The proteins corresponding to these significant spots (of which 43 proteins could be identified) were overexpressed in either SEs (17 proteins) or ZEs (26 proteins). Interestingly, the functional distribution of overexpressed proteins was quite similar for the two types of embryo. Slight variations in the distribution of functional groups (cellular process, response to stimulus, metabolic process) between the SEs and the ZEs could not be readily interpreted since the total number of significant proteins was low. Nevertheless, we suggest that overexpression of certain proteins in ZEs is balanced by overexpression of a different set of proteins from the same functional groups in SEs. These proteins are mainly related to response to stress, accumulation of storage proteins and energy metabolism (Table 4).

Among the stress-related proteins, three HSPs were found to be overexpressed in SEs (two isoforms of HSP 18.2; HSP 81-1) and one in ZEs (class II HSP 17.6). The presence of protein isoforms with different expression patterns has been previously reported during seed maturation in barley (Finnie et al. 2006). HSP 18.2 and HSP 17.6 belong to a small HSP family (HSP 20) which has been detected in differentiated embryos in *Cyclamen* (Rode et al. 2012), but their function is not yet known. Small HSPs have been also found to be upregulated from early to mature stages of SE development in *P. glauca* (Lippert et al. 2005) and *P. abies* (Businge et al. 2013).

Some storage proteins (legumin-like storage protein; putative germin-like protein subfamily 1 member 9) were apparently depleted in SEs compared to ZEs. The higher level of expression of these proteins observed in ZEs collected on 08 August suggested that storage- and stress-related protein accumulation had been initiated by then and would continue throughout the whole autumn desiccation phase (until December; see protein profiles, Fig. 3). A related observation is that of the accumulation in ZEs of proteins with homology to cupins (cupin domain-containing protein and cupin family protein). These proteins belong to a very large family, which includes globulin-like proteins, whose members are named on the basis of a conserved β-barrel fold. A BLAST search revealed that the cupin domain-containing protein corresponded to vicilin-like protein, while the cupin family

protein matched globulin-like protein. High levels of storage proteins in ZEs harvested on 08 August further confirmed that by this stage the embryos had already entered the desiccation phase (see Fig. 1). Clearly this is not the case for SEs matured for 12 weeks. However a number of the proteins overexpressed in SEs (omega-hydroxypalmitate O-feruloyl transferase, embryonic protein DC-8) suggest that SEs are being prepared for the desiccation phase. Omega-hydroxypalmitate O-feruloyl transferase has been detected during early desiccation in seed coats of *Arabidopsis* (Gou et al. 2009). Embryonic protein DC-8 belongs to the LEA type 1 family, is ABA-inducible and may be involved in seed desiccation, as described for *Zea mays* (Gong et al. 2013). This protein may also play a role in late embryogeny (Franz et al. 1989).
A number of the proteins overexpressed in SEs are involved in the glycolysis pathway (malonate--CoA ligase; glucose-1-phosphate adenylyltransferase small subunit; putative NDP-L-rhamnose synthase and alcohol dehydrogenase 1). This suggests that energy metabolism is more active in SEs compared to ZEs, as previously reported by Sghaier-Hammami et al. (2009) in mature SEs from *Phoenix dactylifera* L. Active energy metabolism may explain why SEs matured for 12 weeks are capable of entering the germination phase without undergoing the desiccation-induced dormancy program observed for ZEs. Previous molecular studies in maritime pine have suggested that the germination program may be activated early during SE maturation (Perez-Rodriguez et al. 2006).

Protein markers of the fresh cotyledonary embryo stage during somatic and zygotic embryogenesis
A comparative analysis of significant proteins that were found to be overexpressed in both cotyledonary SEs (2 lines investigated; AAY06006 and NL04048) and cotyledonary ZEs (collected on 08 August) resulted in a short list of 23 proteins (Fig. S7) including five proteins annotated as HSPs (two isoforms of class II HSP 17.6; HSP 18.2; HSP 60; HSP 70-4) and two other stress-related proteins (aldose reductase, 6-phosphogluconate dehydrogenase family protein), four LEAs (embryonic protein DC-8; embryonic cell protein 63; LEA-like protein; responsive to abscisic acid 28), five energy storage proteins (two isoforms of cupin domain-containing protein; legumin-like storage protein; two isoforms of vicilin-like storage protein), two proteins involved in purine metabolism, i.e. adenosine kinase 2 and S-adenosylmethionine (SAM) synthase, and five proteins of unknown function. The simultaneous occurrence of members of the HSP, LEA and energy storage protein families was observed in *Brassica napus*, where they were described as specific proteins accumulating in mature ZEs (Hajduch et al. 2006). In addition, expression of each of these major protein families mentioned has been reported to be specific to the late embryogenesis stage in gymnosperms as well as in angiosperms. These proteins all contribute to the preparation of the embryo for subsequent desiccation and germination phases.
Stress-related proteins are generally overexpressed during the maturation stage in both angiosperms (Finnie et al. 2006) and gymnosperms (Lippert et al. 2005; Silveira et al. 2008; Shi et al. 2010; Teyssier et al. 2011, 2014; Businge et al. 2013; Morel et al. 2014). Consistent with this observation, recent transcriptomic studies have also revealed that several stress-related genes are up-regulated during early embryogenesis in maritime pine (Morel et al. 2014) and late embryogenesis in Norway spruce (Vestman et al. 2011). HSPs are involved in the acquisition of desiccation tolerance and they have been proposed as markers of ZE maturity in *Phoenix dactylifera* (Sghaier-Hammami et al. 2009).
LEA proteins are known to be involved in protecting cellular components from severe dehydration. LEA protein synthesis has been shown to occur at the late maturation stages of both somatic (*Picea abies*, Businge et al. 2013) and zygotic embryos (*Cunninghamia lanceola*, Shi et al. 2010; *Araucaria angustifolia*, Silveira et al. 2008), and in other species (Kermode and Finch-Savage 2002). A recent study also detected LEA transcripts in both somatic and zygotic embryos from *Pinus taeda* (Lara-Chavez et al. 2012).
It is now well established that seed vigour and potential longevity continue to increase after physiological maturity, and that the final stages in seed maturation are critical for maximizing seed quality. Storage proteins could therefore potentially represent excellent markers of embryo quality. The main type of storage protein present seems to be species dependent. While vicilin-like protein is abundant in *Picea abies* (Businge et al. 2013) and *Picea glauca* (Lippert et al. 2005), it is pine globulin-1 and legumin-like proteins that predominate in *Pinus strobus* (Klimaszewska et al. 2004). Similarly, in our study, pine globulin-1 and legumin-like proteins appeared to be the major storage proteins in fresh, cotyledonary embryos of *Pinus pinaster* (Table 3).
Aldose reductase was also identified as a potential marker of the fresh cotyledonary stage of embryos in maritime pine. To our knowledge this is the first time that this protein has been associated with embryo maturation in conifers. Both the protein and the corresponding transcripts were previously detected in barley during late embryogenesis and in relation to desiccation (Finnie et al. 2006). This enzyme is involved in the synthesis of sorbitol, a component used as an osmolyte for regulating the osmotic potential of embryos during desiccation. Interestingly, the synthesis of aldose reductase protein is ABA inducible.
The term purine salvage refers to the reutilization of purine bases and nucleosides derived from purine containing compounds. This pathway seems to be closely linked to the growth and development of embryos

(Stasolla et al. 2001). Both SAM synthase and adenosine kinase 2 are involved in this salvage pathway, generating SAM and adenosine, respectively. These compounds contribute to the increase in size of the nucleotide pool that is required at the onset of germination. This is consistent with the increase in adenosine activity kinase observed during a partial drying treatment (Stasolla et al. 2001). The presence of SAM synthase in cotyledonary SEs confirms the importance of methionine metabolism in seed development and germination (Shi et al. 2010; Rajjou et al. 2012). Since upregulation of SAM synthase has been previously observed in various conifers at stages from early to late embryogeny (Lippert et al. 2005; Jo et al. 2014; Teyssier et al. 2014), it is likely that it contributes to enhanced amino-acid metabolism (Teyssier et al. 2014) as well as to alterations in both polyamine content and ethylene biosynthesis (Jo et al. 2014).

A final protein highlighted in this study was annotated as 6-phosphogluconate dehydrogenase. This protein is involved in the metabolism of sugars (specifically in pentose metabolism) as well as in the metabolism of glutathione, a non-enzymatic scavenger protecting the cell from oxidative stress, which may occur in relation to drought (see Teyssier et al. 2011 and references therein). 6-phosphogluconate dehydrogenase was found to be overexpressed in developing and mature SEs in both maritime pine (Morel et al. 2014) and hybrid larch (Teyssier et al. 2011, 2014), particularly in response to the reduced water availability in maturation medium that results from high gellan gum concentration.

Few studies have previously identified molecular markers for monitoring embryo development in maritime pine. Two glutamine synthetase gene isoforms (GS1a, GS1b) represent useful molecular markers for checking the quality of cotyledonary SEs in terms of both photosynthetic activity and the extent of vascular formation (Perez-Rodriguez et al. 2006). More recently, predictive markers of SE differentiation (germin 16 like protein, ubiquitin protein ligase) have been identified using a combined transcriptomic/proteomic approach (Morel et al. 2014). The same study also reported on one biomarker (protein phosphatase 2C) of the adaptive response of a culture to maturation conditions. This growing arsenal of molecular markers may be invaluable in facilitating the refinement of somatic embryogenesis protocols for the effective implementation of multivarietal forestry in maritime pine.

Conclusions

We conclude that cotyledonary SEs from different embryogenic lines can reach similar standards of quality on the maturation medium used in this study. This quality did not vary as a function of the duration of maturation within the range 10-14 weeks. Routine harvesting of SEs after 12 weeks of maturation therefore appears appropriate on the basis of our biological and biochemical data.

Using hierarchical ascendant cluster analysis (HCA), we demonstrated that cotyledonary SEs matured for 12 weeks are most similar to fresh cotyledonary ZEs sampled from late July to early August (in terms of DW, water content, sucrose, RFOs content, RFO/Su ratio) or at any time up to October (with respect to total protein). Thus, under our experimental conditions, cotyledonary SEs did not reach a level of maturity equivalent to that of cotyledonary ZEs from dried seeds.

The slight variations observed in the distribution of protein functional groups between SEs and ZEs could not be interpreted unambiguously, since the total number (43) of significant proteins was low. Nevertheless it is noteworthy that the overall classification of these identified proteins into functional groups was quite similar for the two embryo types. This suggests that the overabundance of some proteins in ZEs is counterbalanced by overabundance in SEs of other proteins involved in the same functional groups. This new information will be useful in refining our maturation protocols to optimize the quality of seedlings, which is a prerequisite for integrating somatic embryogenesis into the maritime pine breeding program in order to improve the efficiency of both selection and the deployment of new varieties.

This is the first report on the identification of generic protein markers of the embryo cotyledonary stage during somatic and zygotic embryogenesis in maritime pine. Combined proteomic analyses of developing, immature SEs vs. cotyledonary SEs or ZEs revealed a robust set of 23 candidate proteins that are upregulated at the cotyledonary stage. Of these, 18 belonged to five large families of proteins including five HSPs, four LEA and two other stress-related proteins (aldose reductase, 6-phosphogluconate dehydrogenase), five storage proteins and two proteins involved in purine metabolism (adenosine kinase 2, SAM synthase). Subject to confirmation through additional experimental studies, these biomarkers should be useful in monitoring embryo development in maritime pine for purposes of research (e.g. functional validation of candidate genes), development (e.g. scaling up the process of somatic embryogenesis) or practical applications (e.g. plant production).

Maturation conditions appeared to be of great importance for the subsequent steps. Recently Businge et al. (2013) demonstrated that best maturation conditions produced SEs that did not germinate, which was related to the intolerance to desiccation caused by the lack of storage reserves (LEA proteins, sucrose, raffinose). In

conifers it has been demonstrated that SE desiccation induced *de novo* ABA synthesis in hybrid larch (Dronne et al. 1997) and RFOs accumulation in *Picea abies* and *P. mariana* (Kumstýřová et al. 2000; Bomal et al. 2002 respectively). In consequence to refine maturation of maritime pine SEs, we suggest to apply a desiccation treatment. The rate of drying has been shown to be critical factor in acquisition of desiccation tolerance. In general slow drying is essential as reported in carrot (Tetteroo et al. 1995), as well as in conifers such as spruce and larch species (Lelu et al. 1995).

Acknowledgments

This work was supported by grants from the French "Conseil Régional de la Région Centre" (EMBRYOME project, contract 33639) and "Conseil Régional de la Région Aquitaine" (EMBRYO2011 project, contract 09012579-045). Some equipment used at FCBA to get results presented in this work was funded by ANR-10-EQPX-16 XYLOFOREST. We gratefully acknowledge Gilles Bruneau, Francis Canlet, Sandrine Debille, Karine Durandeau, Séverine Quoniou and Françoise Wan-Fung (FCBA) for their valuable contribution to the collection of somatic and zygotic embryos and Pierre Alazard (FCBA) for cone sampling in the field. We acknowledge Marlène Bailly for her technical assistance in electrophoresis.

References

Bailly C, Audigier C, Ladonne F, Wagner MH, Coste F, Corbineau F, Côme D (2001) Changes in oligosaccharide content and antioxidant enzyme activities in developing bean seeds as related to acquisition of drying tolerance and seed quality. J Exp Bot 52: 701-708

Black M, Corbineau F, Grzesik M, Guy P, Côme D (1996) Carbohydrate metabolism in the developing and maturing wheat embryo in relation to its desiccation tolerance. J Exp Bot 47: 161-169

Black M, Corbineau F, Gee H, Côme D (1999) Water content, raffinose, and dehydrins in the induction of desiccation tolerance in immature wheat embryos. Plant Physiol 120: 463–471

Bomal C, Le VQ, Tremblay FM (2002) Induction of tolerance to fast desiccation in black spruce (*Picea mariana*) somatic embryos: relationship between partial water loss, sugars, and dehydrins. Physiol. Plant. 115:523–530; 2002

Bonga JM, Klimaszewska K, Aderkas P (2010) Recalcitrance in clonal propagation, in particular of conifers. Plant Cell Tissue Organ Cult 100: 241-254

Breton D, Harvengt L, Trontin JF, Bouvet A, Favre JM (2005) High subculture frequency, maltose-based and hormone-free medium sustained early development of somatic embryos in Maritime pine. In Vitro Cell Dev Biol-Plant 41: 494-504

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

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

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

Corbineau F, Picard MA, Fougereux JA, Ladonne F, Côme D (2000) Effects of dehydration conditions on desiccation tolerance of developing pea seeds as related to oligosaccharide content and membrane properties. Seed Sci Research 10: 329-339

Dronne S, Label P, Lelu M-A (1997) Desiccation decreases abscisic acid content in hybrid larch (*Larix* × *leptoeuropaea*) somatic embryos. Physiol Plant 99: 433-438

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

Finnie C, Bak-Jensen KS, Laugesen S, Roepstorff P, Svensson B (2006) Differential appearance of isoforms and cultivar variation in protein temporal profiles revealed in the maturing barley grain proteome. Plant Sci 170: 808-821

Franz G, Hatzopoulos P, Jones T, Krauss M, Sung ZR (1989) Molecular and genetic analysis of an embryonic gene, DC 8, from *Daucus carota* L. Mol Gen Genet 218: 143-151

Gong F, Wu X, Wang W (2013) Comparative proteomic identification of embryo proteins associated with hydropriming induced rapid-germination of maize seeds. Plant Omics, pp 333-339

Gou J-Y, Yu X-H, Liu C-J (2009) A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in *Arabidopsis*. Proc Natl Acad Sci USA 106: 18855-18860

Gösslová M, Svobodová H, Lipavská H, Albrechtová J, Vreugdenhil D (2001) Comparing carbohydrate status during Norway spruce seed development and somatic embryo formation. In Vitro Cell Dev Biol-Plant 37: 20-28

Hajduch M, Casteel JE, Hurrelmeyer KE, Song Z, Agrawal GK, Thelen JJ (2006) Proteomic analysis of seed filling in *Brassica napus*. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. Plant Physiol Biochem 141: 32-46

Hakman I (1993) Embryology in Norway spruce (*Picea abies*). An analysis of the composition of seed storage proteins and deposition of storage reserves during seed development and somatic embryogenesis. Physiol Plant 87: 148-159

Hargreaves CL, Reeves CB, Gough K, Josekutty P, Skudder DB, van der Maas SA, Sigley MR, Menzies MI, Low CB, Mullin TJ (2009) Improving initiation, genotype capture and family representation in somatic embryogenesis of *Pinus radiata* D. Don. by a combination of zygotic embryo maturity, media and plant preparation. Can J Forest 39: 1566-1574

Humánez A, Blasco M, Brisa C, Segura J, Arrillaga I (2012) Somatic embryogenesis from different tissues of Spanish populations of maritime pine. Plant Cell Tissue Organ Cult 111: 373-383

Isaacson T, Damasceno CMB, Saravanan RS, He Y, Catala C, Saladie M, Rose JKC (2006) Sample extraction techniques for enhanced proteomic analysis of plant tissues. Nat. Protocols 1: 769-774

Jarlet-Hugues E (1989) Recherches sur l'aptitude à l'embryogenèse somatique de matériel juvénile et de matériel issu d'arbres adultes de *Pinus pinaster* Sol. Ph. Doctorate in Plant Biology and Physiology, University Paris VI, 135 pp

Jo L, Dos Santos ALW, Bueno CA, Barbosa HR, Floh EIS (2014) Proteomic analysis and polyamines, ethylene and reactive oxygen species levels of *Araucaria angustifolia* (Brazilian pine) embryogenic cultures with different embryogenic potential. Tree Physiol 34: 94-104

Jordy MN, Favre JM (2003) Spatio-temporal variations in starch accumulation during germination and post-germinative growth of zygotic and somatic embryos of *Pinus pinaster*. Biol Plant 46: 507-512

Keller F, Ludlow, MM (1993) Carbohydrate metabolism in drought-stressed leaves of pigeon pea (*Cajanus cajan*). J Exp Bot 44:1351-1359

Kermode A (1990) Regulatory mechanisms involved in the transition from seed development to germination. Crit Rev Plant Sci 9: 155-195

Kermode AR, Finch-Savage BE (2002) Desiccation sensitivity in orthodox and recalcitrant seeds in relation to development. In: Black M, Pritchard HW (eds) Desiccation and survival in plants. Drying without dying. CAB International, Oxon pp149-184

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

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

Koltunow AM, Hidaka T, Robinson SP (1996) Polyembryony in *Citrus*. Accumulation of seed storage proteins in seeds and in embryos cultured *in vitro*. Plant Physiol 110: 599-609.

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

Kumstýřová L, Vágner M, Lipavská H, Gösslová M (2000) Somatic embryogenesis of Norway spruce: anatomical characterization and content of non-structural saccharides. Plant Physiol. Biochem. 38(Suppl.):43

Lara-Chavez A, Egertsdotter U, Flinn B (2012) Comparison of gene expression markers during zygotic and somatic embryogenesis in pine. In Vitro Cell Dev Biol-Plant 48: 341-354

Lelu MA, Klimaszewska K. Pflaum, Bastien C (1995) Effect of maturation duration on desiccation tolerance in hybrid larch (*Larix* x *leptoeuropaea* Dengler) somatic embryos. In Vitro Cell Dev Biol-Plant 31: 15-20

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

Lelu-Walter MA, Pâques L (2009) Simplified and improved somatic embryogenesis of hybrid larches (*Larix* x *eurolepis* and *Larix* x *marschlinsii*). Perspectives for breeding. Ann For Sci 66: 104p1-104p10

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

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

Lelu-Walter M-A, Thompson D, Harvengt L, Sanchez L, Toribio M, Pâques LE (2013) Somatic embryogenesis in forestry with a focus on Europe: state-of-the-art, benefits, challenges and future direction. Tree Genet Genomes 9: 883-899

Lipavská H, Konrádová H (2004) Somatic embryogenesis in conifers: The role of carbohydrate metabolism. In Vitro Cell Dev Biol-Plant 40: 23-30

Lippert D, Jun Z, Ralph S, Ellis DE, Gilbert M, Olafson R, Ritland K, Ellis B, Douglas CJ, Bohlmann J (2005) Proteome analysis of early somatic embryogenesis in *Picea glauca*. PROTEOMICS 5: 461-473

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

Mauriat M, Le Provost G, Rozenberg P, Delzon S, Breda N, Clair B, Coutand C, Domec J-C, Fourcaud T, Grima-Pettenati J, Herrera R, Leplé J-C, Richet N, Trontin J-F, Plomion C (2014) Wood formation in trees. In: Tree biotechnology, Ramawat K.G., Merillon J.-M., Ahuja M.R. (eds), Science Publisher, CRC Press, Boca Raton, USA, pp 56-111. In press (550 pp.)

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

Mishra S, Sanyal I, Amla DV (2012) Changes in protein pattern during different developmental stages of somatic embryos in chickpea. Biol Plant 56: 613-619

Morel A, Teyssier C, Trontin JF, Eliášová K, Pešek B, Beaufour, Morabito D, Boizot N, Le Metté C, Belal-Bessai L, Reymond I, Harvengt L, Cadene M, Corbineau F, Vágner M, Label P, Lelu-Walter MA (2014) Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water availability: transcriptomic and proteomic analyses. Physiol. Plant. (in press) doi: 10.1111/ppl.12158

Nehra NS, Becwar MR, Rottmann WH, Pearson L, Chowdhury K, Chang S, Wilde HD, Kodrzycki RJ, Zhang C, Gause KC, Parks DW, Hinchee MA (2005) Forest biotechnology: innovative methods, emerging opportunities. In Vitro Cell Dev Biol-Plant 41: 701-717

Park YS (2002) Implementation of conifer somatic embryogenesis in clonal forestry: technical requirements and deployment considerations. Ann For Sci 59: 651-656

Park YS, Lelu-Walter M-A, Harvengt L, Trontin JF, McEacheron I, Klimaszewska K, Bonga JM (2006) Initiation of somatic embryogenesis in *Pinus banksiana*, *P. strobus*, *P. pinaster*, and *P. sylvestris* at three laboratories in Canada and France. Plant Cell Tiss Organ Cult 86: 87-101

Perez Rodriguez MJ, Suarez MF, Heredia R, Avila C, Breton D, Trontin JF, Filonova L, Bozhkov P, von Arnold S, Harvengt L, Canovas, FM (2006) Expression patterns of two glutamine synthetase genes in zygotic and somatic pine embryos support specific roles in nitrogen metabolism during embryogenesis. New Phytol 169: 35-44

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

Pullman GS, Buchanan M (2008) Identification and quantitative analysis of stage-specific carbohydrates in loblolly pine (*Pinus taeda*) zygotic embryo and female gametophyte tissues. Tree Physiol 28: 985-996

Rajjou L, Duval M, Gallardo K, Catusse J, Bally J, Job C, Job D (2012) Seed germination and vigor. Annu Rev Plant Biol 63: 507-533

Ramarosandratana A, Harvengt L, Bouvet A, Calvayrac R, Pâques M (2001a) Effects of carbohydrate source, polyethylene glycol and gellan gum concentration on embryonal-suspensor mass (ESM) proliferation and maturation of maritime pine somatic embryos. In Vitro Cell Dev Biol-Plant 37: 29-34

Rocha DI, Dornelas MC (2013) Molecular overview on plant somatic embryogenesis. CAB Reviews 8: 022, 1-17

Rode C, Lindhorst K, Braun H-P, Winkelmann T (2012) From callus to embryo: a proteomic view on the development and maturation of somatic embryos in *Cyclamen persicum*. Planta 235: 995-1011

Sallandrouze A, Faurobert M, El Maataoui M, Espagnac H (1999) Two-dimensional electrophoretic analysis of proteins associated with somatic embryogenesis development in *Cupressus sempervirens* L. Electrophoresis 20:1109-1119

Sanz F, Latour S, Neves M, Bastet E, Pischedda D, Piñeiro G, Gauthier T, Lesbats J, Plantier C, Marques A, Lanvin J-D, Santos JA, Touza M, Pedras F, Parrot J, Reuliong D, Faria C (2006) Industrial applications of *Pinus Pinaster*. 256 pp. Madeira: CIS Madeira, FIBA, AIMMP, CTBA.

Sghaier-Hammami B, Drira N, Jorrín-Novo JV (2009) Comparative 2-DE proteomic analysis of date palm (*Phoenix dactylifera* L.) somatic and zygotic embryos. J Proteomics 73: 161-177

Shi J, Zhen Y, Zheng R-H (2010) Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. J Exp Bot 61: 2367-2381

Silveira V, Santa-Catarina C, Balbuena TS, Moraes FMS, Ricart CAO, Sousa MV, Guerra MP, Handro W, Floh EIS (2008) Endogenous abscisic acid and protein contents during seed development of *Araucaria angustifolia*. Biologia Plantarum 52: 101-104

Stasolla C, Loukanina N, Ashihara H, Yeung EC, Thorpe TA (2001) Purine and pyrimidine metabolism during the partial drying treatment of white spruce (*Picea glauca*) somatic embryos. Physiol Plant 111: 93-101

Su G, Christensen OF, Ostersen T, Henryon M, Lund MS (2012) Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. PLOS One 7 (9): e45293, pp 1-7

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

Tetteroo FAA, Bino RJ, Bergervoet JHW, Hasenack B (1995) Effect of ABA and slow drying on DNA replication in carrot (*Daucus carota*) embryoids. Physiol Plant 95: 154-158

Teyssier C, Grondin C, Bonhomme L, Lomenech A-M, Vallance M, Morabito D, Label P, Lelu-Walter M-A (2011) Increased gelling agent concentration promotes somatic embryo maturation in hybrid larch (*Larix* × *eurolepis*): a 2-DE proteomic analysis. Physiol Plant 141: 152-165

Teyssier C, Maury S, Beaufour M, Grondin C, Delaunay A, Le Metté C, Ader K, Cadene M, Label P, Lelu-Walter M-A (2014) In search of markers for somatic embryo maturation in hybrid larch (*Larix* x *eurolepis*): global DNA methylation and proteomic analyses. Physiol Plant 150: 271-291

Trontin J-F, Reymond I, Quoniou S, Canlet F, Debille S, Bruneau G, Harvengt L, Lelu-Walter M-A, Teyssier C, Le Metté C, Vallance M, Label P (2011) An overview of current achievements and shortcomings in developing Maritime pine somatic embryogenesis and enabling technologies in France. In: YS Park, JM Bonga, SY Park, HK Moon (eds) Advances in somatic embryogenesis of trees and its application for the future forests and plantations. IUFRO Working Party 2.09.02: Somatic embryogenesis and other clonal propagation methods of forest trees, August 19-21 2010 (Suwon, South Korea), pp 100-102

Trontin J-F, Debille S, Canlet F, Harvengt L, Lelu-Walter M-A, Label P, Teyssier C, Lesage-Descauses MC, Le Metté C, Miguel C, De Vega-Bartol J, Tonelli M, Santos R, Rupps A, Hassani SB, Zoglauer K, Carneros E, Diaz-Sala C, Abarca D, Arrillaga I, Mendoza-Poudereux I, Segura J, Avila Saez C, Rueda M, Canales J, Canovas F (2013) Somatic embryogenesis as an effective regeneration support for reverse genetics in maritime pine: the Sustainpine collaborative project as a case study. In: YS Park, JM Bonga (eds) Integrating vegetative propagation, biotechnology and genetic improvement for tree production and sustainable forest management. Second international conference of the IUFRO working party 2.09.02 (somatic embryogenesis and vegetative propagation technologies), June 25-28 2012 (Brno, Czech Republic), pp. 184-187

Vestman D, Larsson E, Uddenberg D, Cairney J, Clapham D, Sundberg E, von Arnold S (2011) Important processes during differentiation and early development of somatic embryos of Norway spruce as revealed by changes in global gene expression. Tree Genet Genomes 7: 347-362

Ward JH (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236-244

Yang X, Zhang X (2010) Regulation of somatic embryogenesis in higher plants. Critical Reviews Plant Sci 29: 36-57

Zavattieri MA, Frederico AM, Lima M, Sabino R, Arnholdt-Schmitt B (2010) Induction of somatic embryogenesis as an example of stress-related plant reactions. Electronic J Biotech 13, 1-9

Figure legends

Fig. 1 Biological characteristics of cotyledonary embryos of somatic (SE, line AAY06006, in white) or zygotic origin (ZE, AAY06006 mother tree, in gray) as a function of time of harvest. Somatic embryos were collected after 10, 12 or 14 weeks of maturation. Zygotic embryos were sampled on 7 collection dates from 26 July to 09 December 2011. (**a**) Dry weight; (**b**) Water content. Bars represent 95 % CI. Letters represent statistical groups defined by the Multiple Comparisons of Means method ($P<0.05$, $n = 10$). DW= dry weight; FW= fresh weight

Fig. 2 Changes in carbohydrate content (**a**) and Raffinose + stachyose/sucrose ratio (**b**) in cotyledonary embryos of somatic (SE ○; line AAY06006) or zygotic origin (ZE ●; AAY06006 mother tree) as a function of time of harvest. Somatic embryos were collected after 10, 12 or 14 weeks of maturation. Zygotic embryos were sampled at 7 collection dates from 26 July to 09 December 2011. Bars represent 95 % CI. Letters represent statistical groups defined by the Multiple Comparisons of Means method ($P<0.05$, $n = 3$). DW= dry weight

Fig. 3 Representative SDS-PAGE gel, showing the total protein profile obtained from somatic embryos matured for 12 weeks (SE 12w, line AAY06006) compared to that of maturing cotyledonary zygotic embryos (ZE, AAY06006 mother tree, 6 collection dates). Molecular masses (kDa) of reference protein markers are indicated (lane MM). Arrows and band numbers show the bands excised and the corresponding proteins identified by mass spectrometry

Fig 4 Hierarchical clustering analysis of maturing cotyledonary zygotic embryos (ZE, AAY06006 mother tree) sampled on 7 collection dates and cotyledonary somatic embryos (SE) matured for 12 weeks (line AAY06006). Clustering was based on the whole quantitative dataset available (water content, DW, fructose, galactose, glucose, maltose, sucrose, stachyose and raffinose content and total protein content)

Supplementary data legends

Table S1 Carbohydrate content of cotyledonary somatic embryos (embryogenic line NL04048) matured for 10, 12 and 14 weeks

Table S2 Quantitative analysis of total proteins in cotyledonary somatic embryos from embryogenic line NL04048 matured for 10, 12 or 14 weeks

Table S3 Characteristics of proteins differentially expressed in somatic embryos from 2 embryogenic lines (AAY06006, NL04048) matured for 1 or 12 weeks. Only those proteins overexpressed in embryogenic tissue matured for one week are shown. Spot number refers to the corresponding excised spot marked in Fig. S6. The protein assignments and accession numbers were retrieved from the GenoToul database (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster2/index.html) using the SEQUEST™ identification engine. The most homologous protein in the UniProt database (http://www.uniprot.org/) and its species affiliation are given. # pep.: number of matching peptides; % cov.: protein sequence coverage percentage

Fig. S1 Biological characteristics of cotyledonary somatic embryos (lines NL04048) matured for 10, 12 or 14 weeks. (**a**), Dry weight; (**b**), Water content. Bars represent 95 % CI. Letters represent statistical groups defined by the Multiple Comparisons of Means method (P <0.05, n = 10). DW= dry weight; FW= fresh weight

Fig.S2 Representative SDS-PAGE gel, showing the total protein profile in cotyledonary somatic embryos from 2 embryogenic lines (AAY06006, AAY; NL04048, NL) matured for 10 (10w), 12 (12w) and 14 weeks (14w). Molecular masses (kDa) of protein markers are indicated (lane MM). Arrows and band numbers show the excised bands and corresponding protein identification determined by mass spectrometry

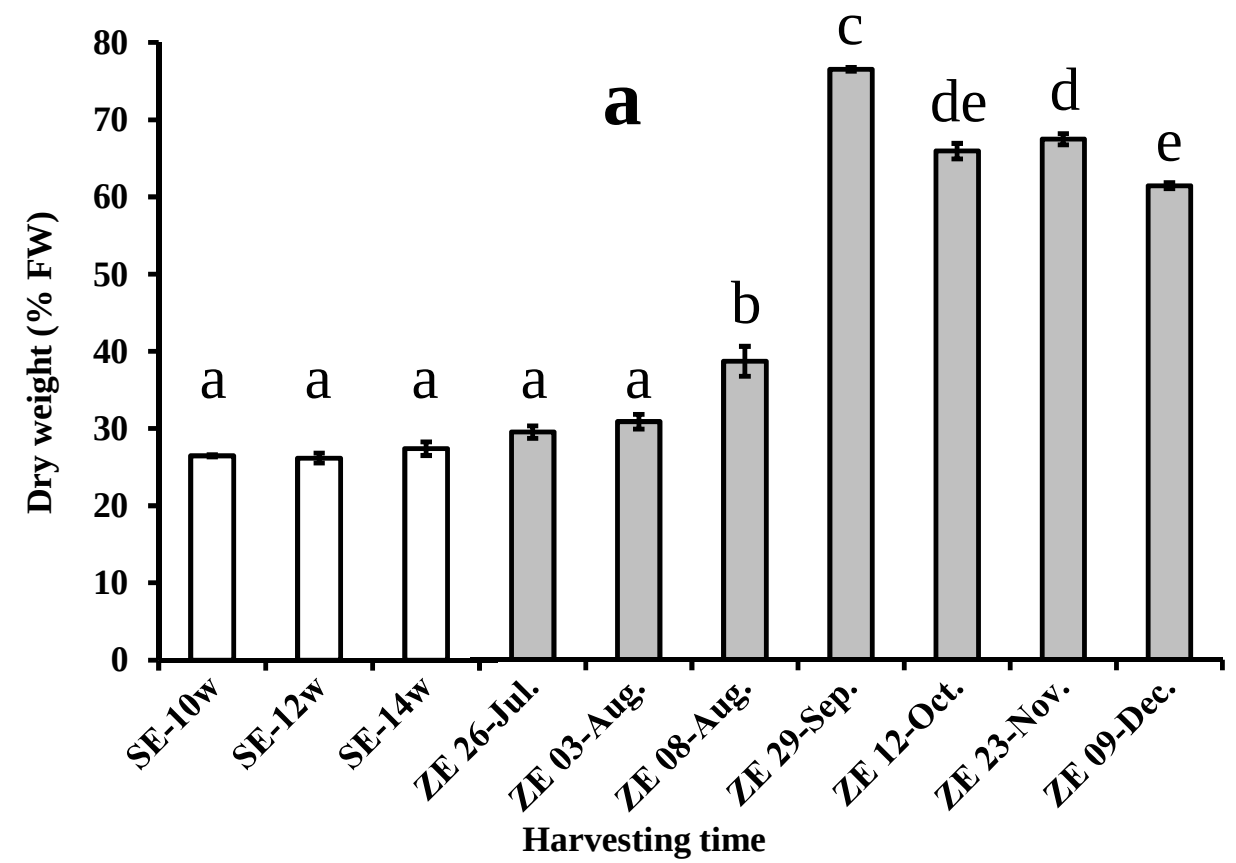
Fig. S3 (**a**) MS/MS spectra of m/z 815,14832 ion precursor corresponding to 2-charge specie of HNADNPEDADIYVR peptide from pine globulin-1 (band 3 of ZEs). (**b**) MS/MS spectra of *m/z* 543,74414 ion precursor corresponding to 3-charge specie of HNADNPEDADIYVR peptide from pine globulin-1 (band 3 of ZEs). (**c**) MS/MS spectra of *m/z* 487,87201 ion precursor corresponding to 3-charge specie of GREEEEEEAVER peptide from albumin 3 (band 4 of SEs)

Fig. S4 2D gel separation of proteins obtained from cotyledonary zygotic embryos (harvested on 08 August). The labeled spots showed significant differences in abundance (P <0.001) between somatic embryos matured for 12 weeks (line AAY06006) and zygotic embryos. The protein identities of these spots are given in Table 4. Molecular masses of protein markers are indicated on the right

Fig S5 Functional classification of identified proteins overexpressed in somatic embryos (**a**) and zygotic embryos (**b**), according to Blast2GO (http://www.blast2go.com/b2ghome). Within each category, the total number of proteins is indicated in parentheses. Proteins involved in more than one process were assigned to more than one category. Hence the sum of the numbers of proteins in the different classes exceeds the total number of proteins identified

Fig. S6 2D gel separation of proteins obtained from cotyledonary somatic embryos matured for 12 weeks (embryogenic line AAY06006). Labelled spots showed significant differences in abundance (P <0.001) between the two somatic embryogenic lines (AAY06006 and NL04048). The protein identities of these spots are given in Table 5 and Table S3. Molecular masses of protein markers are indicated on the right

Fig. S7 Venn diagram showing common proteins specific to certain stages in the development of both cotyledonary somatic embryos (SE) and cotyledonary zygotic embryos (ZE). 12w>1w: proteins overexpressed in SEs matured for 12 weeks *vs* SEs matured for 1 week for both embryogenic lines (AAY006, NL04048); 08-Aug.>AAY06006 1w: proteins overexpressed in cotyledonary ZEs collected on 08 August *vs* immature SEs (1 week maturation)


a
Dry weight (% FW)
80
70
60
50
40
30
20
10
0
a a a a a b c de d e
SE-10w
SE-12w
SE-14w
ZE 26-Jul.
ZE 03-Aug.
ZE 08-Aug.
ZE 29-Sep.
ZE 12-Oct.
ZE 23-Nov.
ZE 09-Dec.
Harvesting time

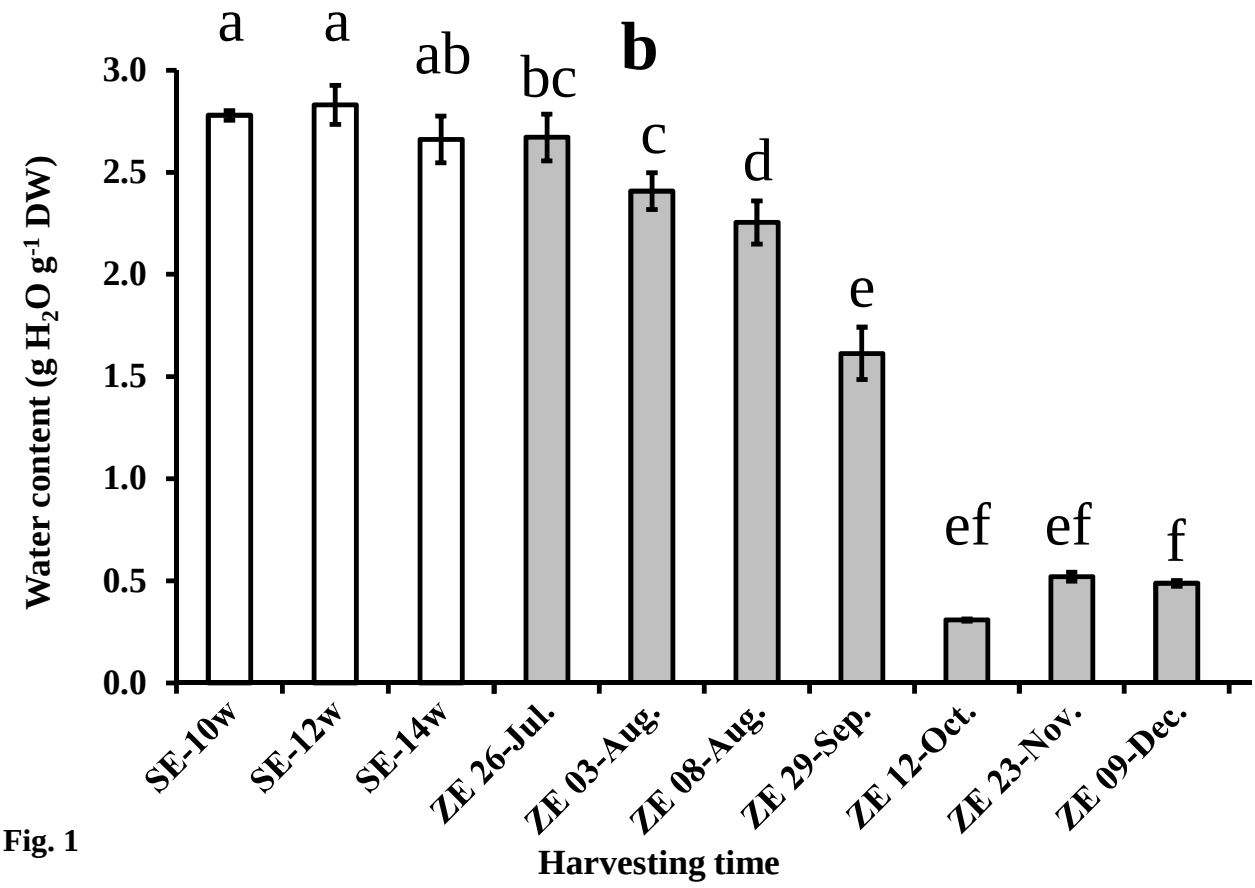
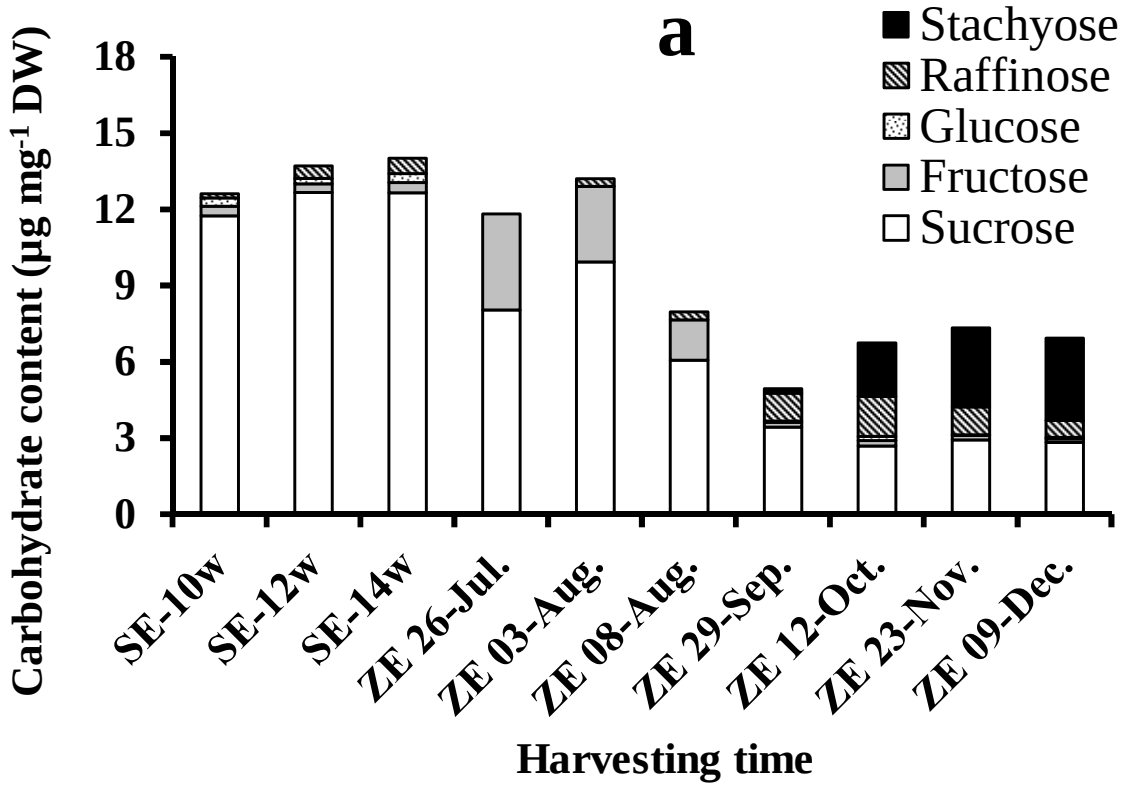

b
Water content (g H2O g-1 DW)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
a a ab bc c d e ef ef f
SE-10w
SE-12w
SE-14w
ZE 26-Jul.
ZE 03-Aug.
ZE 08-Aug.
ZE 29-Sep.
ZE 12-Oct.
ZE 23-Nov.
ZE 09-Dec.
Harvesting time

Fig. 1


a
Stachyose
Raffinose
Glucose
Fructose
Sucrose
Carbohydrate content (µg mg-1 DW)
18
15
12
9
6
3
0
SE-10w
SE-12w
SE-14w
ZE 26-Jul.
ZE 03-Aug.
ZE 08-Aug.
ZE 29-Sep.
ZE 12-Oct.
ZE 23-Nov.
ZE 09-Dec.
Harvesting time

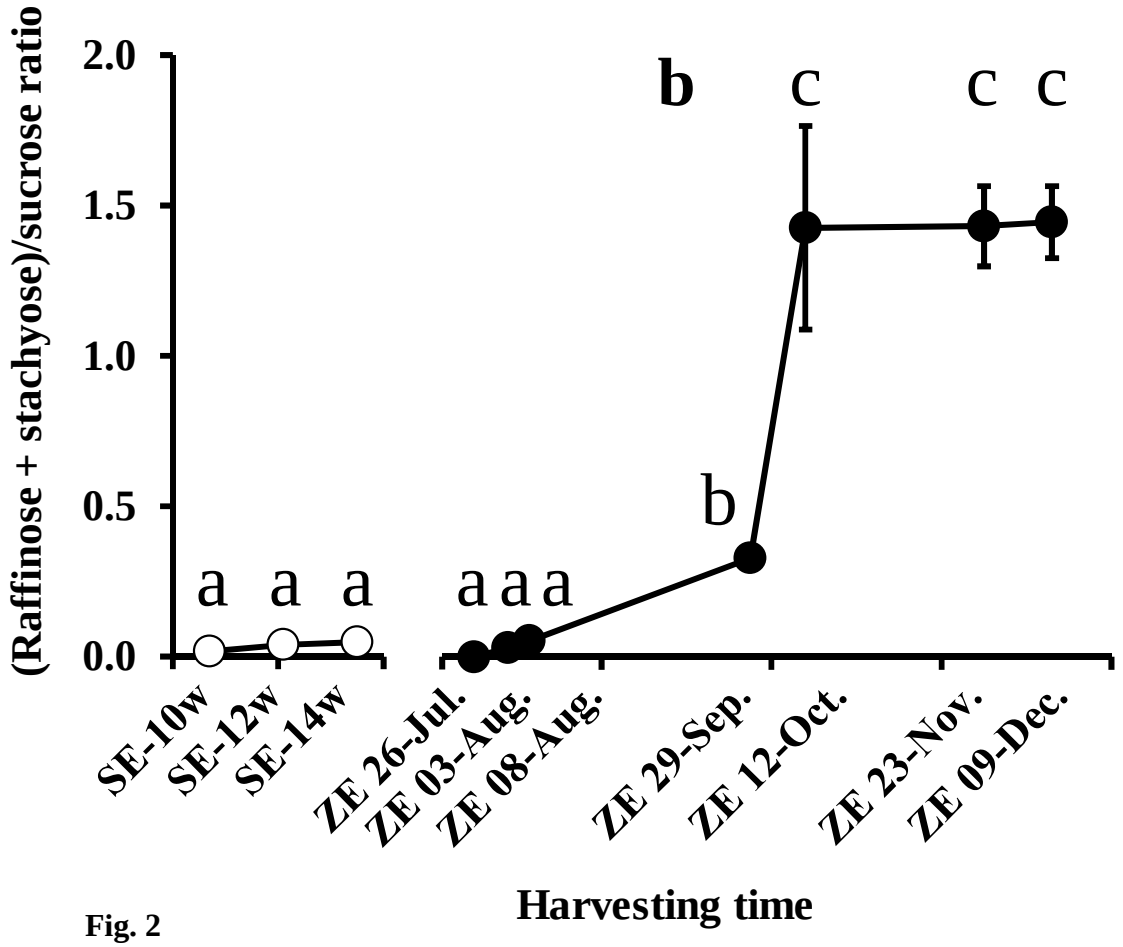

(Raffinose + stachyose)/sucrose ratio
2.0
1.5
1.0
0.5
0.0
b
c
c
c
a
a
a
a
a
a
b
SE-10w
SE-12w
SE-14w
ZE 26-Jul.
ZE 03-Aug.
ZE 08-Aug.
ZE 29-Sep.
ZE 12-Oct.
ZE 23-Nov.
ZE 09-Dec.
Harvesting time

Fig. 2

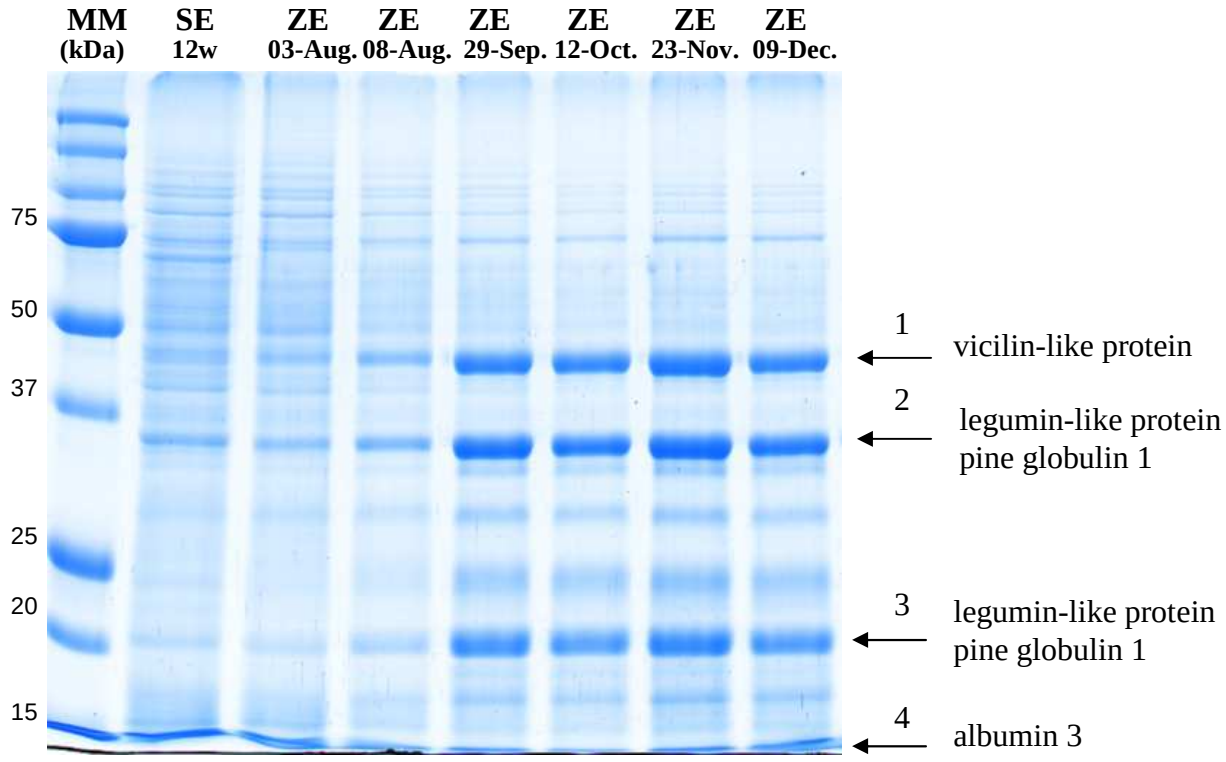

MM
(kDa)
SE
12w
ZE
03-Aug.
ZE
08-Aug.
ZE
29-Sep.
ZE
12-Oct.
ZE
23-Nov.
ZE
09-Dec.
75
50
37
25
20
15
1
vicilin-like protein
2
legumin-like protein
pine globulin 1
3
legumin-like protein
pine globulin 1
4
albumin 3

Fig. 3

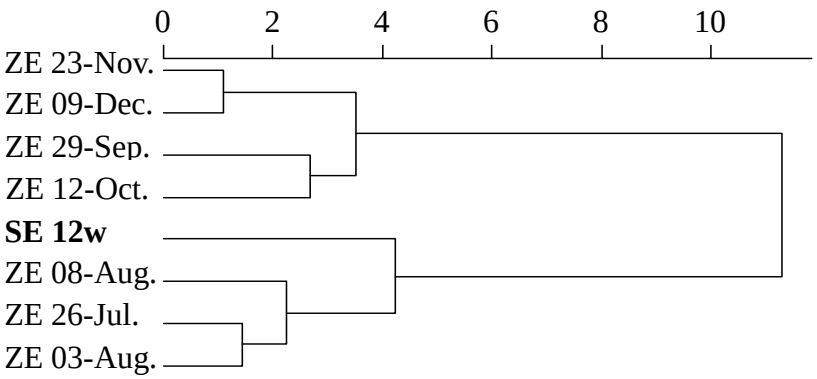

0
2
4
6
8
10
ZE 23-Nov.
ZE 09-Dec.
ZE 29-Sep.
ZE 12-Oct.
SE 12w
ZE 08-Aug.
ZE 26-Jul.
ZE 03-Aug.

Fig 4

Table 1 Carbohydrate content of cotyledonary somatic embryos (SE, line AAY06006) matured for 10, 12 and 14 weeks and of maturing cotyledonary zygotic embryos (ZE, AAY06006 mother tree, 7 collection dates)

Sample	Collection date	Carbohydrate content (µg mg^{-1} DW)							
		Fructose (Fru)	Glucose (Glc)	Sucrose (Su)	Galactose	Maltose	Raffinose (Raf)	Stachyose (Stac)	Ratio (Fru+Glc)/Su
Somatic									
SE	10 weeks	0.37 ± 0.08^{a}	0.33 ± 0.21^{ac}	11.75 ± 1.68^{ac}	0.24 ± 0.3^{a}	1.31 ± 0.16^{bc}	0.16 ± 0.27^{ad}	nd	0.06 ± 0.03^{a}
SE	12 weeks	0.33 ± 0.04^{a}	0.22 ± 0.11^{ab}	12.67 ± 0.96^{a}	0.25 ± 0.09^{a}	1.42 ± 0.1^{ab}	0.49 ± 0.19^{ab}	nd	0.04 ± 0.003^{a}
SE	14 weeks	0.4 ± 0.05^{a}	0.36 ± 0.16^{a}	12.65 ± 0.6^{a}	0.19 ± 0.19^{a}	1.33 ± 0.06^{bc}	0.61 ± 0.07^{abc}	nd	0.06 ± 0.01^{a}
Zygotic									
ZE	26-Jul.	3.78 ± 0.18^{b}	nd	8.05 ± 0.16^{b}	1.85 ± 0.1^{b}	0.73 ± 0.07^{d}	nd	nd	0.47 ± 0.02^{b}
ZE	03-Aug.	2.97 ± 0.28^{c}	nd	9.93 ± 1.07^{c}	2.32 ± 0.23^{b}	1.74 ± 0.21^{a}	0.31 ± 0.31^{bd}	nd	0.3 ± 0.02^{c}
ZE	08-Aug.	1.57 ± 0.2^{d}	nd	6.07 ± 0.37^{d}	0.86 ± 0.24^{c}	1.08 ± 0.04^{c}	0.32 ± 0.08^{bd}	nd	0.26 ± 0.03^{cd}
ZE	29-Sep.	0.18 ± 0.13^{a}	0.04 ± 0.05^{ab}	3.43 ± 0.16^{e}	0.09 ± 0.07^{a}	0.13 ± 0.03^{e}	1.11 ± 0.09^{cef}	0.18 ± 0.3^{a}	0.07 ± 0.04^{a}
ZE	12-Oct.	0.22 ± 0.01^{a}	0.17 ± 0.15^{ab}	2.69 ± 0.19^{e}	0.07 ± 0.09^{a}	0.08 ± 0.01^{e}	1.57 ± 0.22^{e}	2.1 ± 0.56^{b}	0.14 ± 0.05^{ad}
ZE	23-Nov.	0.18 ± 0.07^{a}	0.03 ± 0.02^{b}	2.92 ± 0.26^{e}	0.05 ± 0.02^{a}	0.18 ± 0.05^{e}	1.09 ± 0.14^{f}	3.11 ± 0.69^{b}	0.07 ± 0.02^{a}
ZE	09-Dec.	0.14 ± 0.04^{a}	0.06 ± 0.03^{bc}	2.83 ± 0.11^{e}	0.19 ± 0.04^{a}	0.13 ± 0.05^{e}	0.66 ± 0.02^{b}	3.25 ± 0.24^{b}	0.08 ± 0.03^{a}

*Values are means of 3 measurements ± 95%CI; nd: not detected. Letters represent statistical groups defined by the Multiple Comparisons of Means method ($P<0.05$)

Table 2 Quantitative analysis of total proteins in cotyledonary somatic embryos (SE, embryogenic line AAY06006) matured for 10, 12 and 14 weeks, compared to cotyledonary zygotic embryos at different stages during maturation (ZE, AAY06006 mother tree, 6 collection dates)

Sample	**Collection date**	**Total protein content** ($\mu g\ mg^{-1}$ DW*)
SE	10	223.6 ± 21.1
SE	12	276.7 ± 16.2
SE	14	292.4 ± 16.2
ZE	26-Jul.	nd
ZE	03-Aug.	255.6 ± 10.1
ZE	08-Aug.	227.3 ± 6.4
ZE	29-Sep.	268 ± 22
ZE	12-Oct.	321.6 ± 8.4
ZE	23-Nov.	357.1 ± 24.9
ZE	09-Dec.	363.9 ± 15.2

* Values are means of 5 measurements ± 95 % CI

Table 3 Identification of storage proteins from somatic embryos (SE, line AAY06006) matured for 12 weeks and fresh, maturing cotyledonary zygotic embryos (ZE; AAY06006 mother tree) collected on 08 August. The number of each excised band refers to the corresponding band labeled in Fig. 3. Theoretical molecular masses (Th. M_r) were taken from those of the proteins used to make the assignments. The most homologous protein in the UniProt database (http://www.uniprot.org/) and its species affiliation are given # pep.: number of matching peptides; % cov.: protein coverage percentage

Band N°	Assignment	Accession ID	Species	Th.M_r (kDa)	SE		ZE	
					#pep.	%cov.	#pep.	%cov.
1	vicilin-like storage protein	Q40873	*Picea glauca*	50.2	13	33	11	29
2	pine globulin-1	Q41017	*Pinus strobus*	54.7	3	21	4	28
2	legumin-like storage protein	Q40870	*Picea glauca*	57.4	9	19	9	19
3	pine globulin-1	Q41017	*Pinus strobus*	54.7	2	20	1	10
3	legumin-like storage protein	Q40870	*Picea glauca*	57.4	6	11	8	14
4	albumin 3	Q40997	*Pinus strobus*	19.9	1	8	nd	nd

Table 4 Characteristics of differentially expressed proteins in cotyledonary somatic embryos (line AAY06006) matured for 12 weeks and fresh, maturing cotyledonary zygotic embryos (AAY06006 mother tree, collection date 08 August). Spot number refers to the corresponding excised spot marked in Fig. S4. The protein assignments and accession numbers were retrieved from the GenoToul database (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster2/index.html) by the SEQUEST™ identification engine. The most homologous protein in the UniProt database (http://www.uniprot.org/) and its species affiliation are given. # pep. : Number of matching peptides; % cov.: protein sequence coverage percentage

Putative proteins overexpressed in somatic embryos (17)

Spot n°	Accession	# pep.	% cov.	Homolog	Protein assignment	Species
114	SC_METE_CATRO.1.28	16	84	gi|15238686	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	*Arabidopsis thaliana*
142	SC_ADF6_ORYSJ.19.19	2	48	gi|18411410	actin depolymerizing factor 8	*Arabidopsis thaliana*
139	SC_ARF1_MAIZE.1.3	1	57	Q9LYJ3	ADP-ribosylation factor A1B	*Zea mays*
123	SC_ADH1_PETHY.12.16	10	62	Q43300	alcohol dehydrogenase 1	*Pinus banksiana*
122	SC_LEAD8_DAUCA.1.1	11	52	P20075	embryonic protein DC-8	*Daucus carota*
119	SC_00001504.1.1	9	67	gi|18404934	Euonymus lectin S3	*Arabidopsis thaliana*
121	SC_GLGS_ARATH.1.1	14	74	B8LNV7	glucose-1-phosphate adenylyltransferase small subunit	*Picea sitchensis*
140	SC_00149062.1.1	13	48	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
141	SC_00035783.1.1	10	41	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
103	SC_HSP81_ORYSI.2.25	2	59	Q0J4P2	heat shock protein 81-1	*Oryza sativa subsp. Japonica*
125	SC_00052234.1.1	1	65	gi|18401035	malonate--CoA ligase	*Arabidopsis thaliana*
126	SC_00012145.1.1	4	46	gi|15234993	NmrA-like negative transcriptional regulator family protein	*Arabidopsis thaliana*
108	SC_00001003.1.1	14	54	Q94CD1	omega-hydroxypalmitate O-feruloyl transferase	*Arabidopsis thaliana*
107	SC_PSB1_ARATH.1.8	10	50	A9NMG7	proteasome subunit beta type-1	*Picea sitchensis*
111	SC_00014446.1.1	4	61	Q9SXZ2	protein FLOWERING LOCUS T	*Arabidopsis thaliana*
127	SC_00003578.1.1	9	51	Q9LPG	putative NDP-L-rhamnose synthase	*Arabidopsis thaliana*
128	SC_00003578.1.1	10	51	Q9SYM5	putative NDP-L-rhamnose synthase	*Arabidopsis thaliana*

Putative proteins overexpressed in zygotic embryos (26)

Spot n°	Accession	# pep.	cov.	Homolog	Protein assignment	Species
136	SC_PSB6_ARATH.1.15	2	44	A9NLJ2	20S proteasome subunit PBA1	*Picea sitchensis*
130	SC_00027100.1.1	3	60	gi|15239993	aluminium induced protein with YGL and LRDR motifs	*Arabidopsis thaliana*
117	SC_00000720.1.1	27	42	gi|15226314	chaperonin 60 subunit alpha 1, chloroplastic	*Arabidopsis thaliana*
138	SC_00004359.1.1	5	36	gi|15239846	class II heat shock protein 17.6	*Arabidopsis thaliana*
135	SC_00003455.1.1	4	36	gi|15226403	cupin domain-containing protein	*Arabidopsis thaliana*
120	SC_00114051.1.1	2	38	gi|30690736	cupin family protein	*Arabidopsis thaliana*
100	SC_GALK1_ARATH.1.7	7	48	Q9SEE5	galactokinase	*Arabidopsis thaliana*
133	SC_00013762.1.1	5	56	gi|15222163	glutathione S-transferase DHAR2	*Arabidopsis thaliana*
134	SC_RANA1_TOBAC.6.20	4	52	gi|15240474	GTP-binding nuclear protein Ran-3	*Arabidopsis thaliana*
131	SC_00002599.1.1	3	55	gi|15228389	hydroxyacylglutathione hydrolase	*Arabidopsis thaliana*
101	SC_GLUA1_ORYSJ.1.1	5	75	Q40870	legumin-like storage protein	*Picea glauca*
115	SC_MAOX_VITVI.9.14	3	43	B8LRQ8	NADP-dependent malic enzyme 3	*Picea sitchensis*
110	SC_00007445.1.1	2	65	gi|30691729	peptidase M20/M25/M40 family protein	*Arabidopsis thaliana*
112	SC_00000677.1.1	8	51	gi|15233455	probable 3-hydroxyisobutyrate dehydrogenase-like 1, mitochondrial	*Arabidopsis thaliana*
109	SC_UGPA_SOLTU.12.20	1	79	gi|15228498	probable UTP--glucose-1-phosphate uridylyltransferase 2	*Arabidopsis thaliana*
104	SC_PDI_RICCO.1.8	16	65	gi|15219086	protein disulfide isomerase-like 1-1	*Arabidopsis thaliana*
105	SC_PDI_RICCO.2.8	1	71	gi|15223975	protein disulfide-isomerase 2	*Arabidopsis thaliana*
116	SC_00004333.1.1	2	60	gi|15231398	purple acid phosphatase 15	*Arabidopsis thaliana*
124	SC_00003318.1.1	14	37	gi|30679355	putative aldose reductase	*Arabidopsis thaliana*
102	SC_00146588.1.1	7	65	gi|15241579	putative germin-like protein subfamily 1 member 9	*Arabidopsis thaliana*
132	SC_00016158.1.1	4	45	gi|15240946	putative phosphatidylethanolamine-binding protein	*Arabidopsis thaliana*
106	SC_SYQ_LUPLU.1.3	3	76	gi|22329783	putative tRNA-glutamine ligase	*Arabidopsis thaliana*
118	SC_00004441.1.1	3	61	D5A8K3	serine hydroxymethyltransferase 3	*Picea sitchensis*
129	SC_VATE_CITLI.1.8	9	40	gi|15222641	V-type proton ATPase subunit E3	*Arabidopsis thaliana*
113	SC_00134661.1.1	1		no hit		
137	SC_00047153.1.1	1		no hit		

Table 5 Characteristics of proteins differentially expressed in somatic embryos (SE) from 2 embryogenic lines (AAY06006, NL04048) matured for 1 or 12 weeks. Only proteins overexpressed in SE matured for 12 weeks are presented. Spot number refers to the corresponding excised spot marked in Fig. S6. The protein assignments and accession numbers were retrieved from the GenoToul database (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster2/index.html) using the SEQUEST™ identification engine. The most homologous protein in the UniProt database (http://www.uniprot.org/) and its species affiliation are given. # pep.: number of matching peptides; % cov.: protein coverage percentage. Shaded lines indicate 35 proteins overexpressed in both embryogenic lines

Line AAY06006 (63 significant spots)

Spot n°	Accession	# pep.	% cov.	Homolog	Protein assignment	Species
3	SC_00001921.1.1	11	52	gi|15225307	26S proteasome subunit RPT2B	*Arabidopsis thaliana*
78	SC_00000677.1.1	8	51	gi|15233455	6-phosphogluconate dehydrogenase family protein	*Arabidopsis thaliana*
17	SC_00000826.1.1	5	66	gi|15242717	adenosine kinase 2	*Arabidopsis thaliana*
64	SC_ADH_FRAAN.26.30	9	69	gi|15223838	alcohol dehydrogenase 1	*Arabidopsis thaliana*
2	SC_00003605.1.1	3	34	XM_002530664	bifunctional purine biosynthesis protein, putative	*Ricinus communis*
67	SC_00101372.1.1	10	36	gi|79325071	carboxylate clamp-tetratricopeptide repeat protein	*Arabidopsis thaliana*
1	SC_00024527.1.1	25	71	B9RNX1	chaperone clpb, putative	*Ricinus communis*
36	SC_00004359.1.1	5	36	gi|15239846	class II heat shock protein 17.6	*Arabidopsis thaliana*
37	SC_00146123.1.1	7	33	gi|15239846	class II heat shock protein 17.6	*Arabidopsis thaliana*
44	SC_00100354.1.1	5	60	B9IEK6	coatomer subunit alpha-1	*Populus trichocarpa*
51	SC_00008563.1.1	5	66	gi|15226403	cupin domain-containing protein	*Arabidopsis thaliana*
71	SC_00008563.1.1	2	66	gi|15226403	cupin domain-containing protein	*Arabidopsis thaliana*
29	SC_OS02G0705400_ORYSJ.1.4	5	41	gi|30681471	desiccation responsive protein	*Arabidopsis thaliana*
60	SC_00003529.1.1	6	72	Q9SKP0	embryonic cell protein 63 (LEA)	*Arabidopsis thaliana*
56	SC_LEAD8_DAUCA.1.1	7	52	gi|15228014	embryonic protein DC-8	*Daucus carota*
54	SC_OS01G0633200_ORYSJ.1.1	7	39	gi|15218369	factor of DNA methylation 1	*Arabidopsis thaliana*
70	SC_GLNA_VIGAC.1.7	4	58	A9NSW2	glutamine synthetase	*Picea sitchensis*
14	SC_GSTX4_TOBAC.1.23	10	43	D3YLT8	glutathione S-transferase TAU 19	*Pinus brutia*
9	SC_00035783.1.1	8	41	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
26	SC_00000077.1.1	39	13	gi|18410584	heat shock protein 101	*Arabidopsis thaliana*
35	SC_00035783.1.1	10	41	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
43	SC_HS181_ARATH.23.29	8	55	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
47	SC_00149062.1.1	13	48	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*

7	SC_CH62_CUCMA.2.9	5	69	Q05046	heat shock protein 60	*Cucurbita maxima*
58	SC_CH62_CUCMA.2.9	16	69	Q05046	heat shock protein 60	*Cucurbita maxima*
15	SC_HSP70_SOYBN.1.7	19	79	gi|15230534	heat shock protein 70-4	*Arabidopsis thaliana*
49	SC_HSP12_MEDSA.2.9	13	27	A9NPB3	HSP20-like chaperone	*Picea sitchensis*
52	SC_IDHP_MEDSA.1.16	13	61	A0AR16	isocitrate dehydrogenase [NADP]	*Pinus pinaster*
48	SC_OS03G0723400_ORYSJ.2.2	4	37	Q3I3Y9	late embryogenesis abundant protein	*Picea glauca*
41	SC_LEA14_GOSHI.9.9	2	41	A9NRC1	late embryogenesis abundant-like protein	*Picea sitchensis*
40	SC_GLUA1_ORYSJ.1.1	5	75	Q40870	legumin-like storage protein	*Picea glauca*
5	SC_MAOX_POPTR.1.6	14	62	P34105	NADP-dependent malic enzyme	*Populus trichocarpa*
6	SC_00002037.1.1	5	35	gi|18418333	phosphoglucosamine mutase family protein	*Arabidopsis thaliana*
18	SC_PSB6_ARATH.1.15	2	44	A9NLJ2	proteasome subunit beta type	*Picea sitchensis*
84	SC_PSB1_ARATH.1.8	6	50	A9NMG7	proteasome subunit beta type-1	*Picea sitchensis*
30	SC_MFT_ARATH.1.5	9	39	gi|15220972	protein MOTHER of FT and TF 1	*Arabidopsis thaliana*
12	SC_00003318.1.1	14	37	Q8GXW0	putative aldose reductase	*Arabidopsis thaliana*
4	SC_BADH2_ARATH.7.8	11	53	gi|15228346	putative betaine aldehyde dehydrogenase	*Arabidopsis thaliana*
21	SC_00034040.1.1	6	43	B8LKY4	putative epoxide hydrolase	*Picea sitchensis*
22	SC_00034040.1.1	6	43	B8LKY4	putative epoxide hydrolase	*Picea sitchensis*
82	SC_00012402.1.1	5	45	D2XZY3	putative glutathione S-transferase zeta-class 2	*Pinus brutia*
76	SC_00003578.1.1	10	51	gi|42562732	putative NDP-L-rhamnose synthase	*Arabidopsis thaliana*
77	SC_00003578.1.1	9	51	gi|42562732	putative NDP-L-rhamnose synthase	*Arabidopsis thaliana*
23	SC_00016158.1.1	4	45	gi|15240946	putative phosphatidylethanolamine-binding protein	*Arabidopsis thaliana*
31	SC_PARP3_MEDTR.3.3	11	47	PARP3	putative poly [ADP-ribose] polymerase 3	*Medicago truncatula*
19	SC_00001195.1.1	7	14	D8R837	putative uncharacterized protein	*Selaginella moellendorffii*
74	SC_00026973.1.1	2	76	D7U9K4	putative uncharacterized protein	*Vitis vinifera*
63	SC_00001195.1.1	8	14	D8R837	putative uncharacterized protein	*Selaginella moellendorffii*
45	SC_00015936.1.1	6	41	gi|15218707	responsive to abscisic acid 28	*Arabidopsis thaliana*
34	SC_00074853.1.1	2	45	gi|145359708	RNA recognition motif-containing protein	*Arabidopsis thaliana*
73	SC_METK3_PICSI.17.44	9	50	METK3	S-adenosylmethionine synthase 4	*Picea sitchensis*
75	SC_SFGH_ARATH.1.6	5	51	A9NNG6	S-formylglutathione hydrolase	*Picea sitchensis*
55	SC_00007742.1.1	5	60	B9T0B2	transketolase	*Ricinus communis*
61	SC_IF4A3_ORYSJ.1.23	2	65	gi|18400210	translational initiation factor 4A-1	*Arabidopsis thaliana*

10	SC_00006198.1.1	13	75	B9RHT3	uncharacterized protein	*Ricinus communis*
24	SC_00026973.1.1	8	76	D7U9K4	uncharacterized protein	*Vitis vinifera*
8	SC_VCLA_GOSHI.1.2	11	53	Q40873	vicilin-like storage protein	*Picea glauca*
62	SC_VCLA_GOSHI.1.2	13	33	Q40873	vicilin-like storage protein	*Picea glauca*
80	SC_00028328.1.1	4	40	gi|15222641	V-type proton ATPase subunit E3	*Arabidopsis thaliana*
25	SC_00038758.1.1	7		no hit		
38	SC_00017068.1.1	4		no hit		
39	SC_00087063.1.1	8		no hit		
83	SC_00105062.1.1	8		no hit		

Line NL04048 (58 significant spots)

Spot n°	Accession	# pep.	% cov.	Homolog	Protein assignment	Species
78	SC_00000677.1.1	8	51	gi|15233455	6-phosphogluconate dehydrogenase family protein	*Arabidopsis thaliana*
81	SC_UMC2374_MAIZE.1.3	9	47	gi|18420789	6-phosphogluconolactonase	*Arabidopsis thaliana*
16	SC_ARD4_ARATH.1.1	3	49	A9NPM8	acireductone dioxygenase 4	*Picea sitchensis*
66	SC_STAD_RICCO.1.4	10	52	A9NVD4	acyl-[acyl-carrier-protein] desaturase	*Picea sitchensis*
28	SC_00022934.1.1	7	69	A9NUD5	adenine phosphoribosyl transferase 1	*Picea sitchensis*
17	SC_00000826.1.1	5	66	gi|15242717	adenosine kinase 2	*Arabidopsis thaliana*
64	SC_ADH_FRAAN.26.30	9	69	gi|15223838	alcohol dehydrogenase 1	*Arabidopsis thaliana*
36	SC_00004359.1.1	5	36	gi|15239846	class II heat shock protein 17.6	*Arabidopsis thaliana*
37	SC_00146123.1.1	7	33	gi|15239846	class II heat shock protein 17.6	*Arabidopsis thaliana*
51	SC_00008563.1.1	5	66	gi|15226403	cupin domain-containing protein	*Arabidopsis thaliana*
71	SC_00008563.1.1	2	66	gi|15226403	cupin domain-containing protein	Arabidopsis thaliana
42	SC_00068972.1.1	3	51	gi|42571377	desiccation-induced 1VOC-like protein	*Arabidopsis thaliana*
68	SC_00002262.1.1	4	77	gi|18404397	DJ1-like protein B	*Arabidopsis thaliana*
60	SC_00003529.1.1	6	72	Q9SKP0	embryonic cell protein 63 (LEA)	*Arabidopsis thaliana*
56	SC_LEAD8_DAUCA.1.1	7	52	gi|15228014	embryonic protein DC-8	*Daucus carota*
69	SC_FL3H_MALDO.13.23	7	76	gi|15230433	flavanone 3-hydroxylase	*Arabidopsis thaliana*
70	SC_GLNA_VIGAC.1.7	4	58	A9NSW2	glutamine synthetase	*Picea sitchensis*
14	SC_GSTX4_TOBAC.1.23	10	43	D3YLT8	glutathione S-transferase TAU 19	*Pinus brutia*
13	SC_00114233.1.1	2	33	gi|15239505	glycine-rich RNA-binding protein 3	*Arabidopsis thaliana*

46	SC_GRP1_SINAL.36.42	3	48	Q9XEL4	glycine-rich RNA-binding protein 7	*Picea glauca*
35	SC_00035783.1.1	10	41	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
43	SC_HS181_ARATH.23.29	8	55	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
47	SC_00149062.1.1	13	48	gi|15238509	heat shock protein 18.2	*Arabidopsis thaliana*
7	SC_CH62_CUCMA.2.9	5	69	Q05046	heat shock protein 60	*Cucurbita maxima*
15	SC_HSP70_SOYBN.1.7	19	79	gi|15230534	heat shock protein 70-4	*Arabidopsis thaliana*
52	SC_IDHP_MEDSA.1.16	13	61	A0AR16	isocitrate dehydrogenase [NADP]	*Pinus pinaster*
41	SC_LEA14_GOSHI.9.9	2	41	A9NRC1	late embryogenesis abundant-like protein	*Picea sitchensis*
40	SC_GLUA1_ORYSJ.1.1	5	75	Q40870	legumin-like storage protein	*Picea glauca*
59	SC_00001003.1.1	14	54	D8T9Y4	omega-hydroxypalmitate O-feruloyl transferase	*Selaginella moellendorffii*
50	SC_00017345.1.1	9	17	A9NLD0	peptidyl-prolyl cis-trans isomerase	*Picea sitchensis*
18	SC_PSB6_ARATH.1.15	2	44	A9NLJ2	proteasome subunit beta type	*Picea sitchensis*
30	SC_MFT_ARATH.1.5	9	39	gi|15220972|	protein MOTHER of FT and TF 1	*Arabidopsis thaliana*
12	SC_00003318.1.1	14	37	gi|30679355	putative aldose reductase	*Arabidopsis thaliana*
20	SC_BADH2_ARATH.1.8	4	74	gi|15228346	putative betaine aldehyde dehydrogenase	*Arabidopsis thaliana*
23	SC_00016158.1.1	4	45	gi|15240946	putative phosphatidylethanolamine-binding protein	*Arabidopsis thaliana*
19	SC_00001195.1.1	7	14	D8R837	putative uncharacterized protein	*Selaginella moellendorffii*
74	SC_00026973.1.1	2	76	D7U9K4	putative uncharacterized protein	*Vitis vinifera*
63	SC_00001195.1.1	8	14	D8R837	putative uncharacterized protein	*Selaginella moellendorffii*
45	SC_00015936.1.1	6	41	gi|15218707	responsive to abscisic acid 28	*Arabidopsis thaliana*
11	SC_00000576.1.1	3	41	gi|30682553	RNA recognition motif-containing protein	*Arabidopsis thaliana*
34	SC_00074853.1.1	2	45	gi|145359708	RNA recognition motif-containing protein	*Arabidopsis thaliana*
33	SC_00000156.1.1	5	49	gi|15229321	RNA-binding KH domain-containing protein	*Arabidopsis thaliana*
73	SC_METK3_PICSI.17.44	9	50	METK3	S-adenosylmethionine synthase 4	*Picea sitchensis*
72	SC_00015315.1.1	4	43	Q9LJ97	seed maturation protein (LEA)	*Arabidopsis thaliana*
53	SC_00002400.1.1	2	54	gi|30681414	SUMO-activating enzyme subunit 2	*Arabidopsis thaliana*
27	SC_SODM_NICPL.1.5	11	51	A9NNA1	superoxide dismutase	*Picea sitchensis*
61	SC_IF4A3_ORYSJ.1.23	2	65	gi|18400210	translational initiation factor 4A-1	*Arabidopsis thaliana*
65	SC_00015092.1.1	6	42	A9NKB6	uncharacterized protein	*Picea sitchensis*
8	SC_VCLA_GOSHI.1.2	11	53	Q40873	vicilin-like storage protein	*Picea glauca*
32	SC_VCLA_GOSHI.1.2	3	53	Q40873	vicilin-like storage protein	*Picea glauca*
62	SC_VCLA_GOSHI.1.2	13	33	Q40873	vicilin-like storage protein	*Picea glauca*

57	SC_VATA_MAIZE.6.7	3	70	P49087	V-type proton ATPase catalytic subunit A	*Zea mays*
79	SC_VATE_CITLI.1.8	9	40	gi|15222641	V-type proton ATPase subunit E3	*Arabidopsis thaliana*
80	SC_00028328.1.1	4	40	gi|15222641	V-type proton ATPase subunit E3	*Arabidopsis thaliana*
25	SC_00038758.1.1	7		no hit		
38	SC_00017068.1.1	4		no hit		
83	SC_00105062.1.1	8		no hit		

Table 6 Significant spots (23) corresponding to proteins overexpressed both in cotyledonary somatic embryos(lines AAY06006 and NL04048) matured for 12 weeks and fresh, cotyledonary zygotic embryos from green cones (AAY06006 mother tree, collection date 08 August). More information about each spot is given in Table 5

Spot n°	Assignment
78	6-phosphogluconate dehydrogenase family protein
17	adenosine kinase 2
36	class II heat shock protein 17.6
37	class II heat shock protein 17.6
71	cupin domain-containing protein
51	cupin domain-containing protein
60	embryonic cell protein 63
56	embryonic protein DC-8
43	heat shock protein 18.2
7	heat shock protein 60
15	heat shock protein 70-4
41	late embryogenesis abundant-like protein
40	legumin-like storage protein
12	putative aldose reductase
63	putative uncharacterized protein
74	putative uncharacterized protein
45	responsive to abscisic acid 28
34	RNA recognition motif-containing protein
73	S-adenosylmethionine synthase 4
8	vicilin-like storage protein
62	vicilin-like storage protein
38	no hit
83	no hit

Supplementary Tables

Table S1. Carbohydrate content of cotyledonary somatic embryos (embryogenic line NL04048) matured for 10, 12 and 14 weeks.

Maturation duration (weeks)	Carbohydrate content (µg mg^{-1} DW)								
	Fructose (Fru)	Glucose (Glu)	Sucrose (Su)	Galactose	Maltose	Raffinose (Raf)	Stachyose (Stac)	Ratio (Fru+Glu)/Su	Ratio (Raf+Stac)/Su
10	0.57 ± 0.24[a]	0.62 ± 0.23[a]	14.91 ± 0.93[a]	nd	1.33 ± 0.19[a]	0.5 ± 0.21[a]	nd	0.08 ± 0.02[a]	0.03 ± 0.01[b]
12	0.49 ± 0.16[a]	0.58 ± 0.21[a]	14.83 ± 2.32[a]	0.05 ± 0.08[a]	1.51 ± 0.1[a]	0.17 ± 0.28[a]	nd	0.07 ± 0.03[a]	nd
14	0.42 ± 0.09[a]	0.45 ± 0.09[a]	13.52 ± 2.16[a]	0.04 ± 0.06[a]	1.43 ± 0.09[a]	0.41 ± 0.09[a]	nd	0.07 ± 0.02[a]	0.03 ± 0.01[b]

*Values are means of 3 measurements ± 95 % CI ; nd: not detected. Letters represent statistical groups defined by the Multiple Comparisons of Means method $P<0.05$).

Table S2 Quantitative analysis of total proteins in cotyledonary somatic embryos from embryogenic line NL04048 matured for 10, 12 or 14 weeks.

Maturation duration (weeks)	**Total protein content** ($\mu g\ mg^{-1}$ DW*)
10	249.7 ± 20.4^{a}
12	241.3 ± 23.8^{a}
14	241.4 ± 20.2^{a}

*Values are means of 5 measurements ± 95 % CI. Letters represent statistical groups defined by the Multiple Comparisons of Means method ($P<0.05$)

Table S3 : Characteristics of proteins differentially expressed in somatic embryos from 2 embryogenic lines (AAY06006, NL04048) matured for 1 or 12 weeks. Only those proteins overexpressed in embryogenic tissue matured for one week are shown. Spot number refers to the corresponding excised spot marked in Fig. S6. The protein assignments and accession numbers were retrieved from the GenoToul database (http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster2/index.html) using the SEQUEST™ identification engine. The most homologous protein in the UniProt database (http://www.uniprot.org/) and its species affiliation are given. # pep.: number of matching peptides; % cov.: protein sequence coverage percentage.

Line AAY06006

Spot n°	Accession	# pep.	% cov.	Homolog	Protein assignment	Species
200	SC_METE_CATRO.1.28	8	84	gi|15238686	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	*Arabidopsis thaliana*
201	SC_METE_CATRO.1.28	16	84	gi|15238686	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	*Arabidopsis thaliana*
246	SC_METE_CATRO.1.28	12	84	gi|15238686	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	*Arabidopsis thaliana*
227	SC_ACT1_ACACA.1.2	5	89	ACT1	actin-1	*Acanthamoeba castellanii*
257	SC_ACT7_ARATH.3.10	14	64	ACT7	actin-7	*Arabidopsis thaliana*
205	SC_SAHH_NICSY.1.26	26	49	B8LLL7	adenosylhomocysteinase	*Picea sitchensis*
209	SC_SAHH_NICSY.1.26	10	49	B8LLL7	adenosylhomocysteinase	*Picea sitchensis*
213	SC_SAHH_NICSY.1.26	9	49	B8LLL7	adenosylhomocysteinase	*Picea sitchensis*
214	SC_SAHH_NICSY.1.26	17	49	B8LLL7	adenosylhomocysteinase	*Picea sitchensis*
252	SC_00149928.1.1	16	40	A9NUB8	adenosylhomocysteinase	*Picea sitchensis*
241	SC_APX1_ARATH.1.4	15	41	Q6RY58	ascorbate peroxidase	*Pinus pinaster*
206	SC_CALR_RICCO.1.19	15	36	Q9FYV2	calreticulin	*Pinus taeda*
211	SC_CALR_RICCO.1.19	8	36	Q9FYV2	calreticulin	*Pinus taeda*
212	SC_CALR_RICCO.1.19	9	36	Q9FYV2	calreticulin	*Pinus taeda*
215	SC_CALR_RICCO.2.19	19	37	gi|15217459	calreticulin 1b	*Arabidopsis thaliana*
216	SC_CALR_RICCO.2.19	12	37	gi|15217459	calreticulin 1b	*Arabidopsis thaliana*

251	SC_00000720.1.1	27	42	gi|15226314	chaperonin-60 alpha	*Arabidopsis thaliana*
207	SC_00000168.1.1	4	36	gi|18405251	clathrin light chain protein	*Arabidopsis thaliana*
240	SC_COPE1_ARATH.2.5	10	63	gi|15221463	coatomer subunit epsilon-1	*Arabidopsis thaliana*
233	SC_00004509.1.1	4	48	gi|15223799	expansin-A11	*Arabidopsis thaliana*
219	SC_GLNA_VIGAC.1.7	11	58	A9NSW2	glutamine synthetase	*Picea sitchensis*
242	SC_00000869.1.1	13	35	O23508	growth regulator like protein	*Arabidopsis thaliana*
244	SC_00000093.1.1	18	38	A7YAU9	GRP94	*Pinus taeda*
267	SC_00000019.1.1	10	42	gi|15223288	inorganic pyroph	*Arabidopsis thaliana*
268	SC_00000019.1.1	10	42	gi|15223288	inorganic pyroph	*Arabidopsis thaliana*
236	SC_00005387.1.1	10	59	gi|18422918	lactoylglutathione lyase / glyoxalase I-like protein	*Arabidopsis thaliana*
217	SC_00001300.1.1	5	22	gi|15224810	late embryogenesis abundant protein, group 2	*Arabidopsis thaliana*
263	SC_00001300.1.1	6	22	gi|15224810	late embryogenesis abundant protein, group 2	*Arabidopsis thaliana*
225	SC_NACA_PINTA.1.36	5	41	gi|15230476	nascent polypeptide-associated complex subunit alpha-like protein 1	*Arabidopsis thaliana*
260	SC_00000221.1.1	18	58	gi|22330456	pfkB-like carbohydrate kinase family protein	*Arabidopsis thaliana*
208	SC_00000369.1.1	5	51	B9S6T7	plant ubiquilin, putative	*Ricinus communis*
218	SC_SCRK4_ARATH.1.11	13	64	gi|15231692	probable fructokinase-4	*Arabidopsis thaliana*
234	SC_PROF1_PHAVU.6.6	4	50	A9NMR7	profilin 3	*Picea sitchensis*
204	SC_PDI14_ARATH.1.13	21	47	gi|30697404	protein disulfide isomerase-like 1-4	*Arabidopsis thaliana*
249	SC_PDI14_ARATH.1.13	20	47	gi|30697404	protein disulfide isomerase-like 1-4	*Arabidopsis thaliana*
203	SC_00005357.1.1	4	78	B9SKJ9	protein transport protein sec23, putative	*Ricinus communis*
243	SC_XYL1_ARATH.1.6	7	89	Q8VWV9	putative alpha-xyl	*Pinus pinaster*
228	SC_00002530.1.1	10	49	gi|30687535	quinone reductase family protein	*Arabidopsis thaliana*
220	SC_UPTG2_SOLTU.1.38	8	48	gi|15232865	reversibly glyc	*Arabidopsis thaliana*
221	SC_UPTG_PEA.1.9	23	63	gi|15232865	reversibly glyc	*Arabidopsis thaliana*

259	SC_UPTG2_SOLTU.1.38	20	48	gi|15232865	reversibly glyc	*Arabidopsis thaliana*
255	SC_METK3_PICSI.16.44	11	74	METK3	S-adenosylmethionine synthase 4	*Picea sitchensis*
232	SC_TCTP_PSEMZ.25.123	8	51	TCTP	translationally-controlled tumor protein homolog	*Pseudotsuga menziesii*
254	SC_TBA_PRUDU.1.9	17	63	TBA	tubulin alpha chain	*Prunus dulcis*
210	SC_TBB2_ELEIN.2.4	11	59	TBB2	tubulin beta-2 chain	*Eleusine indica*
253	SC_TBB2_ELEIN.1.4	20	69	TBB2	tubulin beta-2 chain	*Eleusine indica*
269	SC_00009054.1.1	4	44	gi|30690246	uridylate kinase	*Arabidopsis thaliana*

Line NL04048

Spot n°	Accession	# pep.	% cov.	Homolog	Assignment	Species
224	SC433_MESCR.1.7	5		no hit		
226	SC_PSA3_SPIOL.1.2	8	44	A9NPU7	20S proteasome subunit alpha type	*Picea sitchensis*
247	SC_METE_CATRO.1.28	24	84	gi|15238686	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	*Arabidopsis thaliana*
235	SC_ADF6_ORYSJ.19.19	2	48	gi|18411410	actin depolymerizing factor 8	*Arabidopsis thaliana*
248	SC_00000840.1.1	5	63	gi|42563520	acyl-CoA binding protein 4	*Arabidopsis thaliana*
205	SC_SAHH_NICSY.1.26	26	49	B8LLL7	adenosylhomocysteinase	*Picea sitchensis*
209	SC_SAHH_NICSY.1.26	10	49	B8LLL7	adenosylhomocysteinase	*Picea sitchensis*
213	SC_SAHH_NICSY.1.26	9	49	B8LLL7	adenosylhomocysteinase	*Picea sitchensis*
240	SC_COPE1_ARATH.2.5	10	63	gi|15221463	coatomer subunit epsilon-1	*Arabidopsis thaliana*
230	SC_CYSP3_SOLLC.3.8	1	36	gi|79314271	cysteine proteinases superfamily protein	*Arabidopsis thaliana*
237	SC_00148090.1.1	2	60	gi|186507192	embryo-specific protein 3, (ATS3)	*Arabidopsis thaliana*
233	SC_00004509.1.1	4	48	gi|15223799	expansin-A11	*Arabidopsis thaliana*
245	SC_GCSP2_ARATH.1.4	7	51	GCSPA	glycine dehydrogenase [decarboxylating] A, mitochondrial	*Flaveria pringlei*

202	SC_HSP83_IPONI.3.7	1	91	HSP83	Heat shock protein 83	*Ipomoea nil*
268	SC_00000019.1.1	10	42	gi|15223288	inorganic pyroph	*Arabidopsis thaliana*
250	SC_ACEA_PINTA.1.3	5	46	Q41084	isocitrate lyase	*Pinus taeda*
236	SC_00005387.1.1	10	59	gi|18422918	lactoylglutathione lyase / glyoxalase I-like protein	*Arabidopsis thaliana*
238	SC_00150887.1.1	1	63	gi|18422918	lactoylglutathione lyase / glyoxalase I-like protein	*Arabidopsis thaliana*
239	SC_00152425.1.1	3	59	gi|18422918	lactoylglutathione lyase / glyoxalase I-like protein	*Arabidopsis thaliana*
217	SC_00001300.1.1	5	22	gi|15224810	late embryogenesis abundant protein, group 2	*Arabidopsis thaliana*
261	SC_MDHC_MAIZE.1.8	7	70	A9NVT9	malate dehydrogenase	*Picea sitchensis*
258	SC_00001248.1.1	5	53	gi|15220119	methionine-specific aminotransferase	*Arabidopsis thaliana*
264	SC_00000085.1.1	4	53	gi|15234993	NmrA-like negative transcriptional regulator family protein	*Arabidopsis thaliana*
256	SC_00018650.1.1	2	70	gi|15217539	peroxidase 12	*Arabidopsis thaliana*
222	SC_IFRH_TOBAC.9.14	1	36	Q9M524	phenylcoumaran benzylic ether reductase homolog TP5	*Tsuga heterophylla*
204	SC_PDI14_ARATH.1.13	21	47	gi|30697404	protein disulfide isomerase-like 1-4	*Arabidopsis thaliana*
203	SC_00005357.1.1	4	78	B9SKJ9	protein transport protein sec23, putative	*Ricinus communis*
243	SC_XYL1_ARATH.1.6	7	89	Q8VWV9	putative alpha-xyl	*Pinus pinaster*
267	SC_EXB14_ORYSJ.1.2	7	51	EXB14	putative expansin-B14	*Oryza sativa subsp. japonica*
229	SC_00019098.1.1	1	35	A9NYJ8	putative uncharacterized protein	*Picea sitchensis*
266	SC_00106437.1.1	4	41	B8LRN7	putative uncharacterized protein	*Picea sitchensis*
262	SC_00013671.1.1	5	63	gi|15226926	RmlC-like cupins superfamily protein	*Arabidopsis thaliana*
255	SC_METK3_PICSI.16.44	11	74	METK3	S-adenosylmethionine synthase 4	*Picea sitchensis*
231	SC_SKP1B_ARATH.11.13	1	54	Q3ZDI8	SCF ubiquitin ligase	*Picea abies*
223	SC_00001337.1.1	5	47	gi|30693092	translation initiation factor eIF3 subunit	*Arabidopsis thaliana*
232	SC_TCTP_PSEMZ.25.123	8	51	TCTP	translationally-controlled tumor protein homolog	*Pseudotsuga menziesii*
270	SC_00006647.1.1	7	44	gi|186506028	translin family protein	*Arabidopsis thaliana*

254	SC_TBA_PRUDU.1.9	17	63	TBA	tubulin alpha chain	*Prunus dulcis*
265	SC_RHM2_ARATH.1.1	4	40	gi|15218420	UDP-L-rhamn	*Arabidopsis thaliana*
269	SC_00009054.1.1	4	44	gi|30690246	uridylate kinase	*Arabidopsis thaliana*

Supplementary Figures

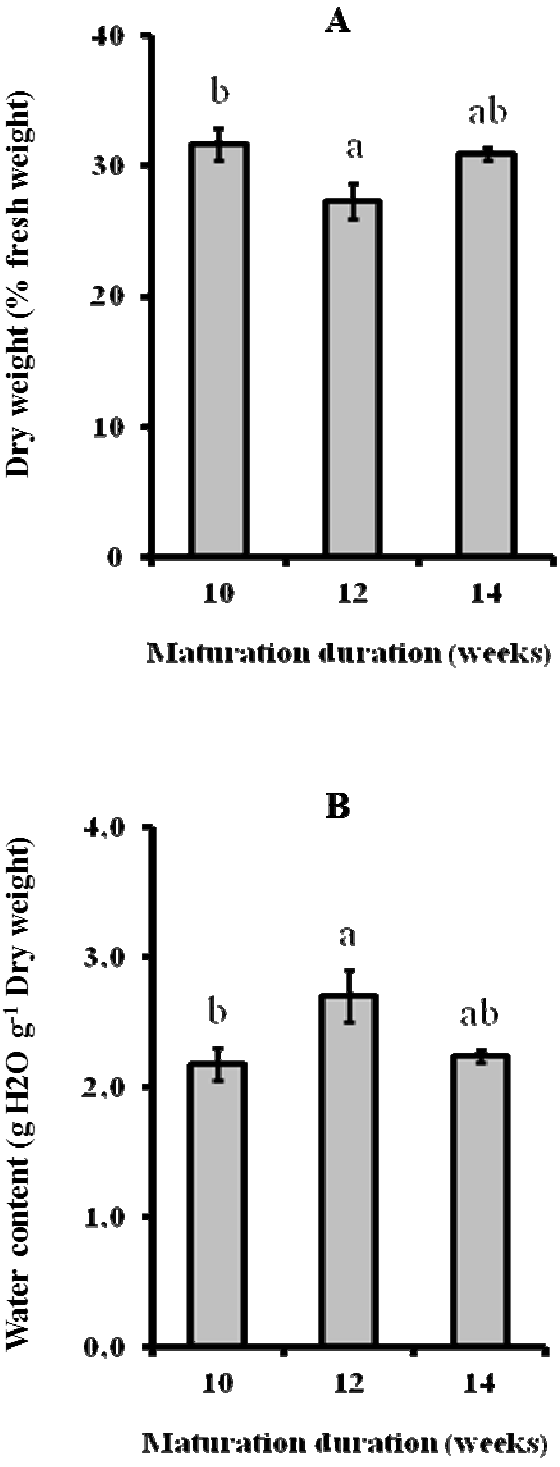


Fig. S1. Biological characteristics of cotyledonary somatic embryos (lines NL04048) matured for 10, 12 or 14 weeks. A, Dry weight; B, Water content. Bars represent 95 % CI. Letters represent statistical groups defined by the Multiple Comparisons of Means method ($P<0.05$, $N = 10$).

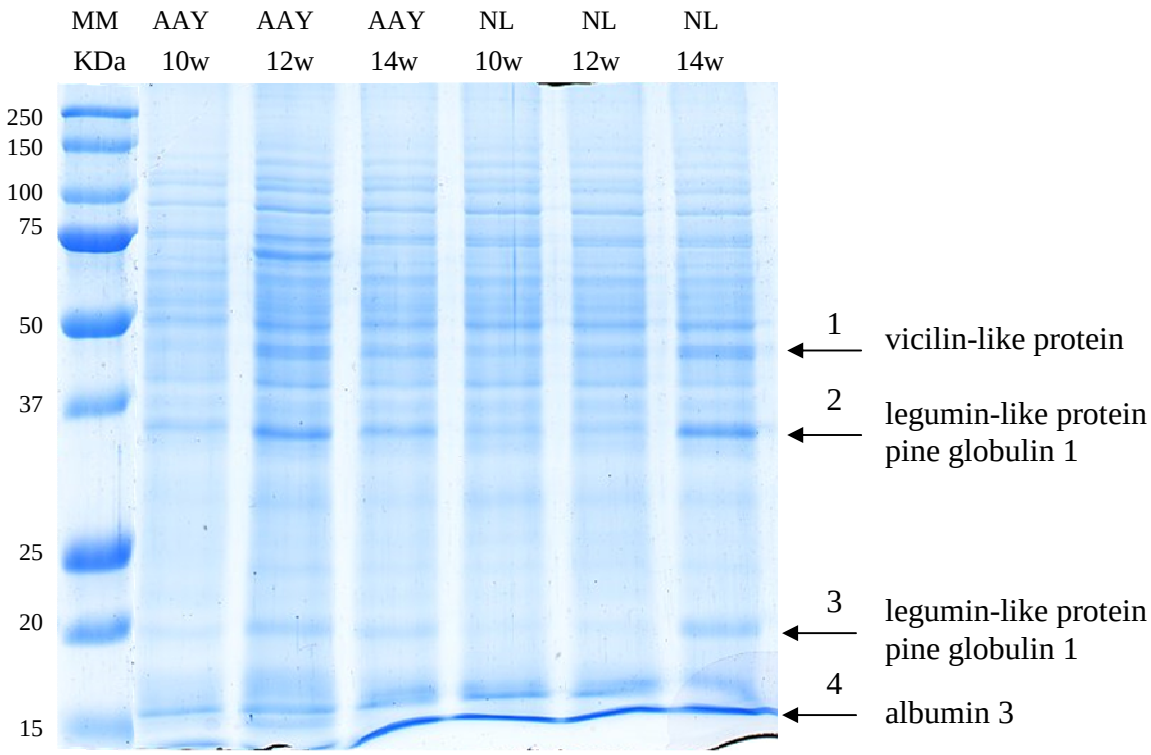


Fig.S2. Representative SDS-PAGE gel, showing the total protein profile in cotyledonary somatic embryos from 2 embryogenic lines (AAY06006, AAY; NL04048, NL) matured for 10 (10w), 12 (12w) and 14 weeks (14w). Molecular masses (KDa) of protein markers are indicated (lane MM). Arrows and band numbers show the excised bands and corresponding protein identification determined by mass spectrometry.

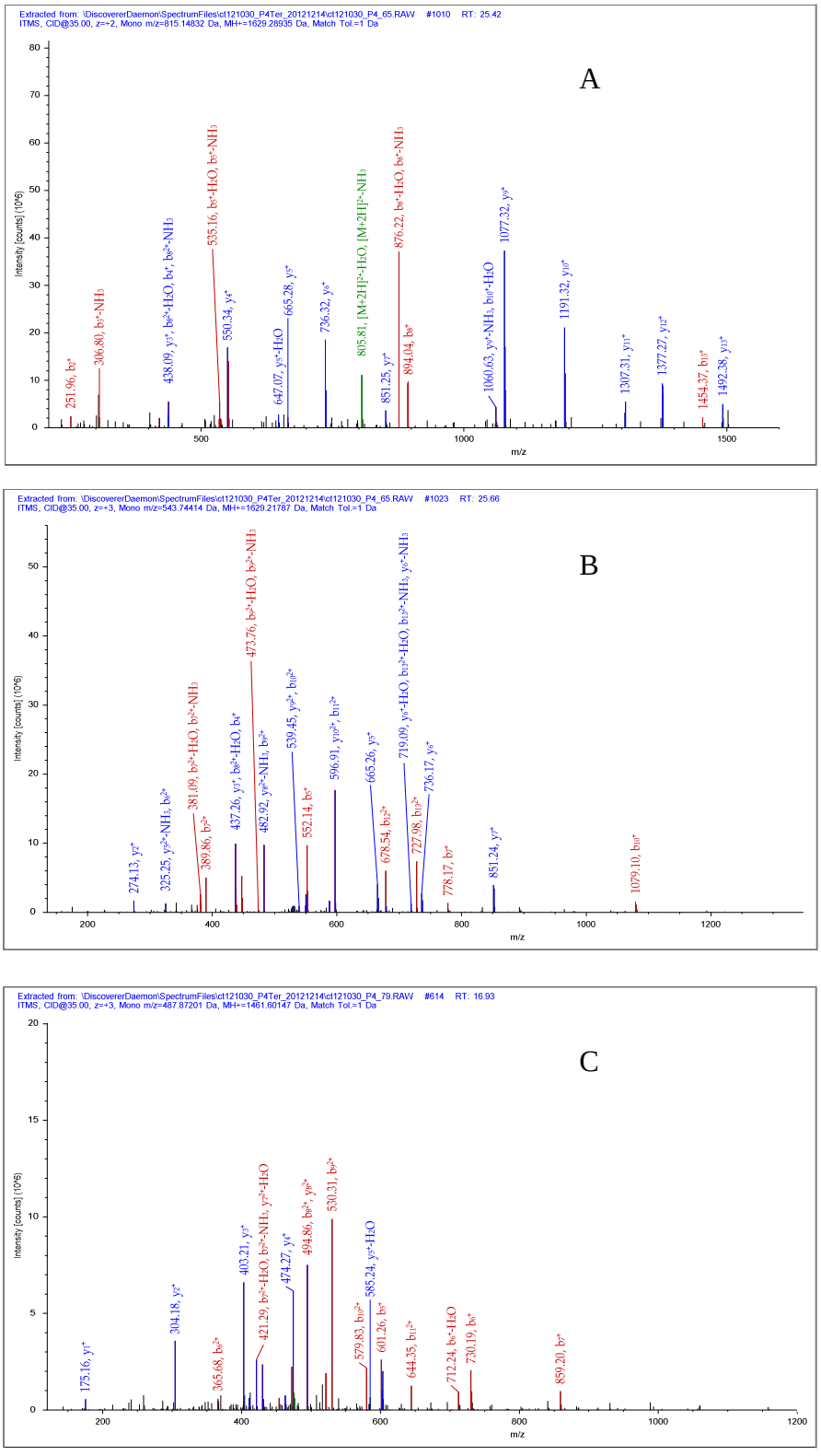


Fig. S3 : (**A**) MS/MS spectra of m/z 815,14832 ion precursor corresponding to 2-charge specie of HNADNPEDADIYVR peptide from pine globulin-1 (band 3 of ZEs). (**B**) MS/MS spectra of m/z 543,74414 ion precursor corresponding to 3-charge specie of HNADNPEDADIYVR peptide from pine globulin-1 (band 3 of ZEs).(**C**) MS/MS spectra of m/z 487,87201 ion precursor corresponding to 3-charge specie of GREEEEEEAVER peptide from albumin 3 (band 4 of SEs).

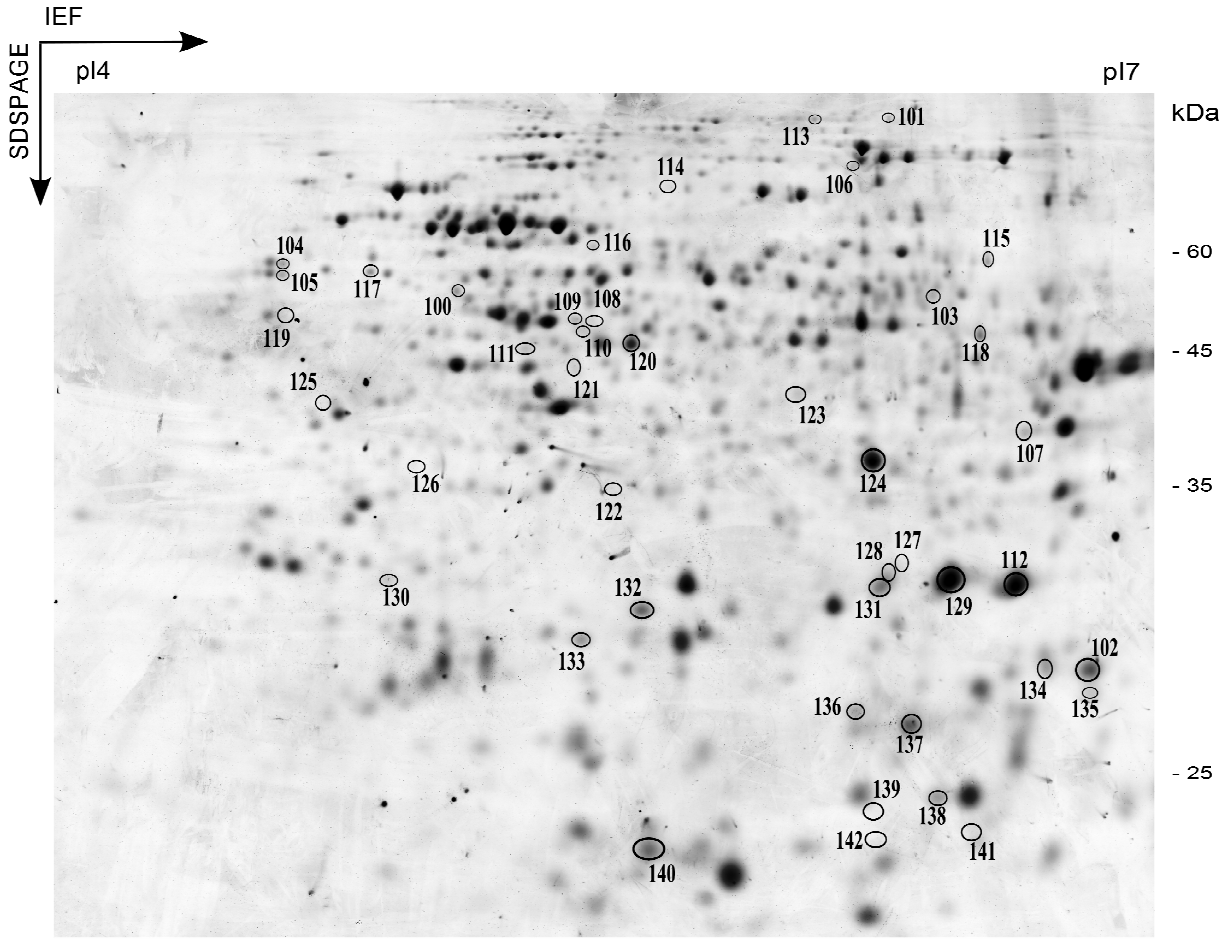


Fig. S4 2D gel separation of proteins obtained from cotyledonary zygotic embryos (harvested on 08 August). The labeled spots showed significant differences in abundance (P <0.001) between somatic embryos matured for 12 weeks (line AAY06006) and zygotic embryos. The protein identities of these spots are given in Table 4. Molecular masses of protein markers are indicated on the right.

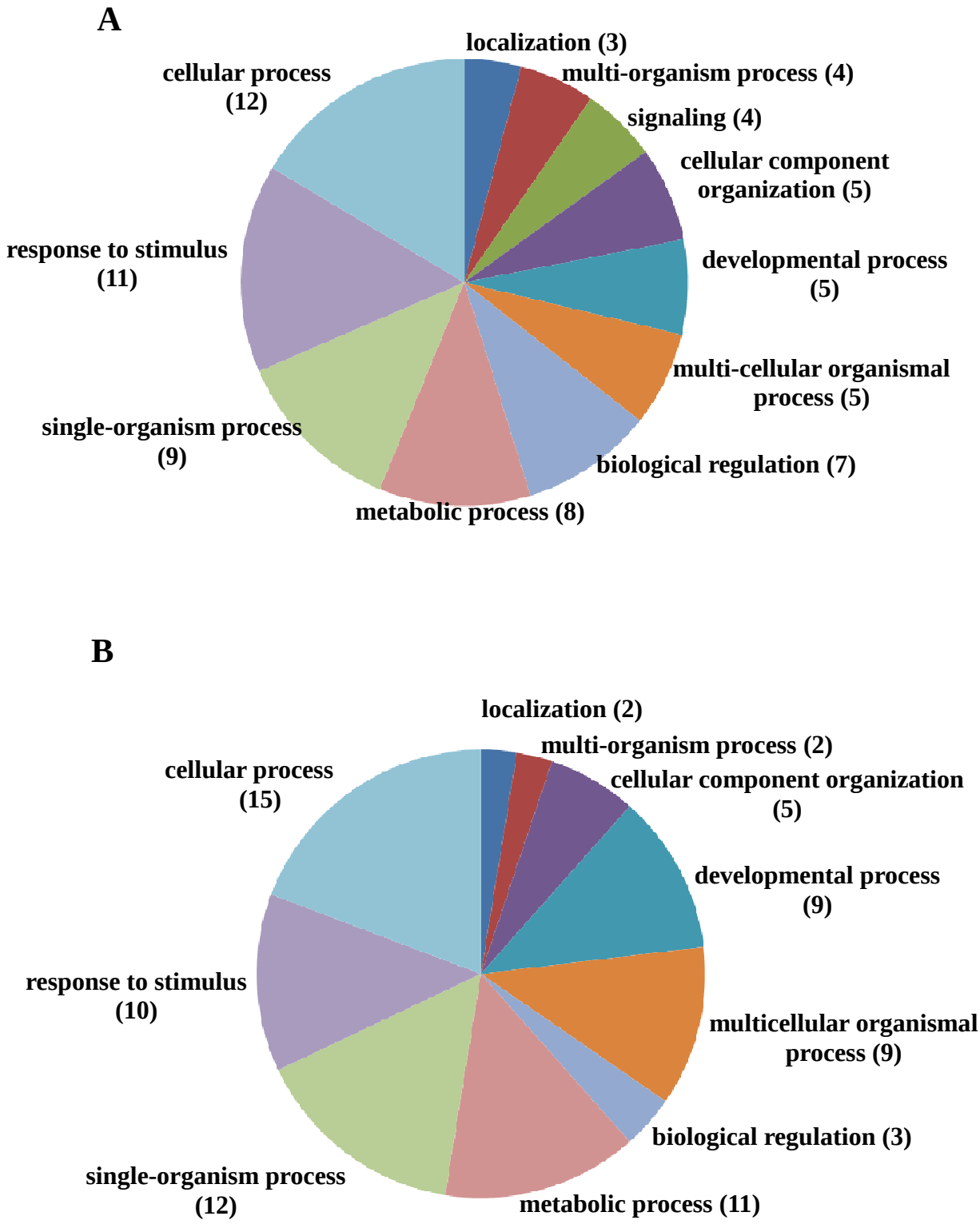


Fig S5: Functional classification of identified proteins overexpressed in somatic embryos (**A**) and zygotic embryos (**B**), according to Blast2GO (http://www.blast2go.com/b2ghome). Within each category, the total number of proteins is indicated in parentheses. Proteins involved in more than one process were assigned to more than one category. Hence the sum of the numbers of proteins in the different classes exceeds the total number of proteins identified.

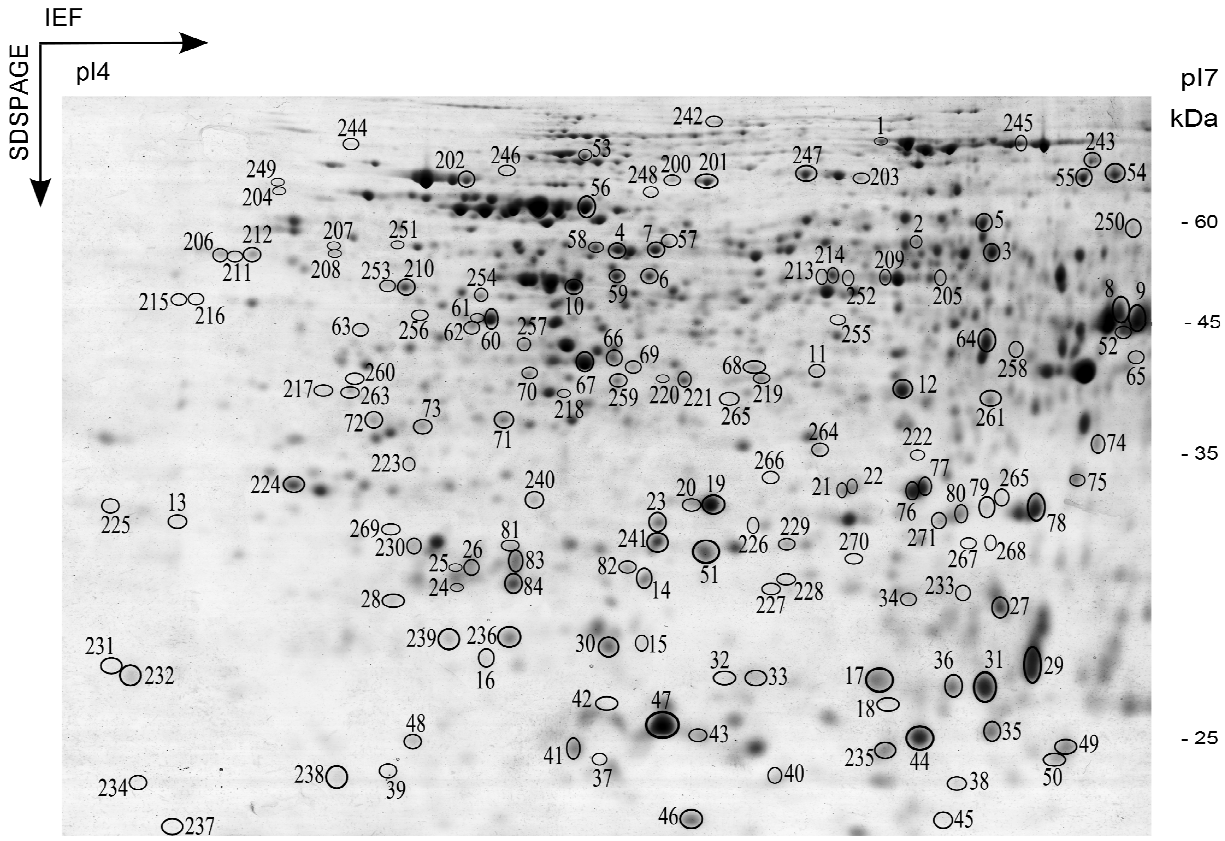


Fig. S6 : 2D gel separation of proteins obtained from cotyledonary somatic embryos matured for 12 weeks (embryogenic line AAY06006). Labelled spots showed significant differences in abundance ($P < 0.001$) between the two somatic embryogenic lines (AAY06006 and NL04048). The protein identities of these spots are given in Table 5 and Table S3. Molecular masses of protein markers are indicated on the right.

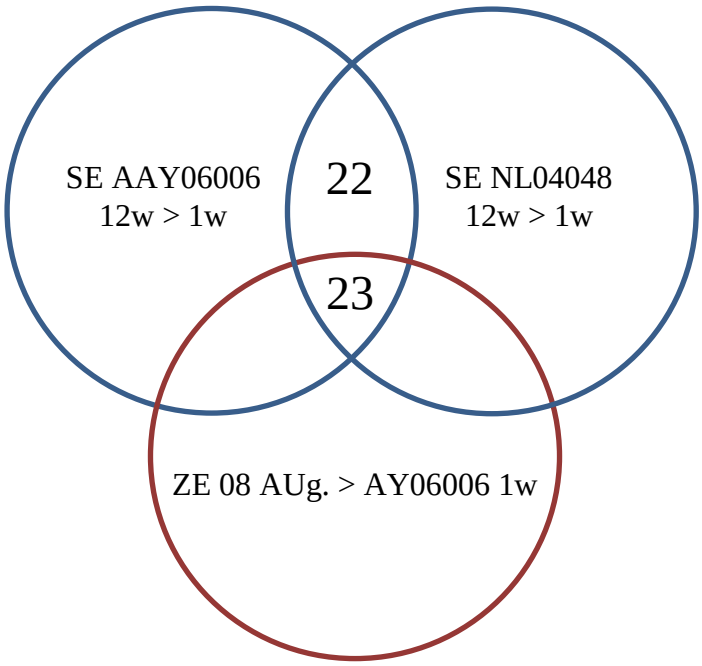


Fig. S7 : Venn diagram showing common proteins specific to certain stages in the development of both cotyledonary somatic embryos (SE) and cotyledonary zygotic embryos (ZE). 12w>1w: proteins overexpressed in SEs matured for 12 weeks *vs* SEs matured for 1 week for both embryogenic lines (AAY006, NL04048); 08 Aug. >AAY06006 1w: proteins overexpressed in cotyledonary ZEs collected on 08 August *vs* immature SEs (1 week maturation).

III.3 Résultats complémentaires

III.3.1 Evolutions physiologique et biochimique de l'EZ de pin maritime au cours de sa maturation

Une analyse en composantes principales (ACP) a été réalisée entre les 7 stades de maturation de l'EZ récoltés entre juillet et décembre 2011 (**Figure III.2**). Les variables explicatives introduites dans l'analyse sont la teneur en eau , la masse sèche, la masse fraîche, le contenu en sucres avec le fructose (fru), le glucose (glu), le saccharose (sac), le raffinose (raf), le stachyose (sta) et contenu en protéines totales (prot) (Morel et al, en préparation, chapitre III.2). L'axe des ordonnées explique 78 % de la variance totale, contre 12 % pour l'axe des abscisses. Deux groupes principaux se détachent de part et d'autre de l'axe des ordonnées, le groupe des EZ frais (EZ_1, EZ_2, EZ_3) et celui des EZ déshydratés (EZ_4, EZ_5, EZ_6, EZ_7). Les variables expliquant la variabilité entre ces 2 groupes, sont surtout : la teneur en eau, le contenu en saccharose et en fructose plus élevés dans le groupe frais, et la masse fraîche, la masse sèche, le contenu en glucose et en raffinose sont plus élevés dans le groupe déshydraté.

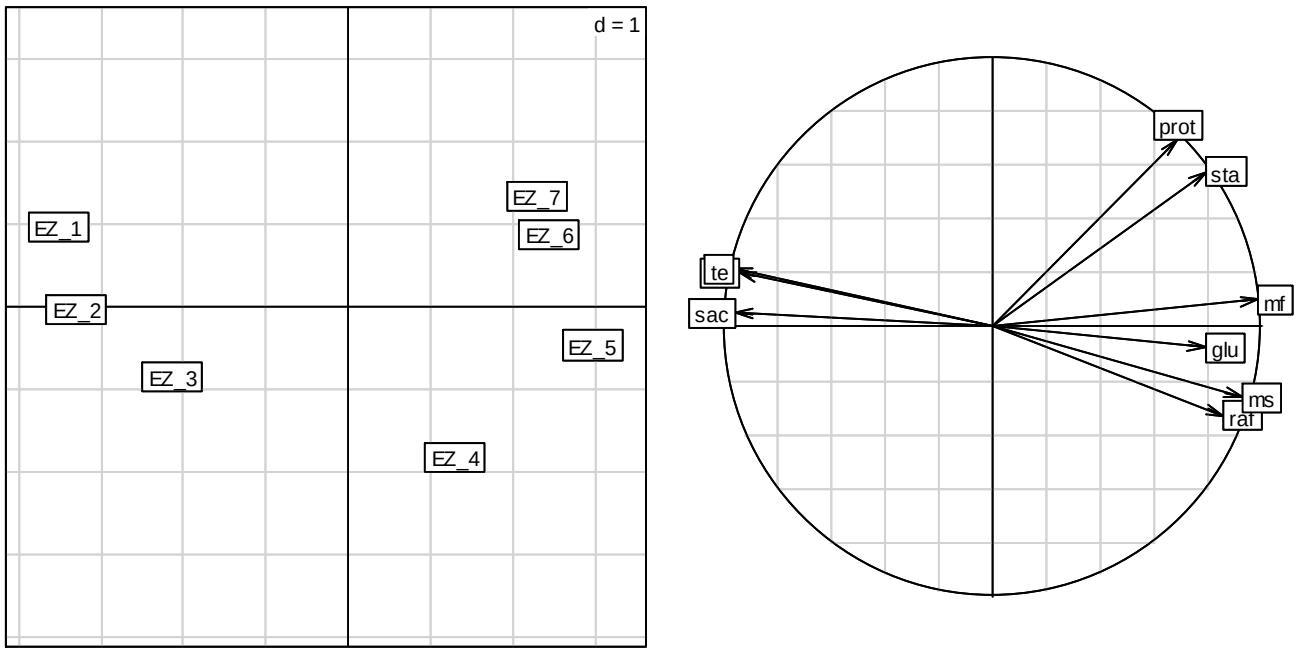


Figure III.2 Représentation graphique de l'analyse en composantes principales de la variabilité de l'EZ de pin maritime. **te** : teneur en eau ; **ms** : masse sèche ; **mf** : masse fraîche ; **fru** : fructose ; **glu** : glucose ; **sac** : saccharose ; **raf** : raffinose ; **sta** : stachyose ; **prot** : protéines totales ; **EZ_1** : 26 juillet ; **EZ_2** : 3 août ; **EZ_3** : 8 août ; **EZ_4** : 29 septembre ; **EZ_5** : 12 octobre ; **EZ_6** : 23 novembre ; **EZ_7** : 9 décembre

III.3.2 Evolutions physiologique et biochimiques du mégagamétophyte de pin maritime au cours de la maturation de la graine

Lors de la campagne de récolte des différents EZ (entre le 26 juillet et le 9 décembre 2011), le tissu nutritif de l'EZ, le mégagamétophyte a été conservé. La disponibilité de ces échantillons, nous a permis de mesurer leur teneur en eau ainsi que leur contenu en sucres solubles, en protéines totales et en protéines de réserve, parallèlement aux mesures effectuées pour l'EZ. Les mêmes méthodologies d'extractions et de mesures ont ainsi été appliquées au mégagamétophyte (Morel et al, en préparation, chapitre III.2, pour la méthodologie). Des échanges et des équilibres de nutriments s'établissent entre l'embryon et son tissu nutritif tout au long de l'embryogenèse. Par conséquent, les variations quantitatives des composés de réserves (sucres et protéines) dans le mégagamétophyte, peuvent nous donner des indices sur des échanges éventuels avec l'EZ.

III.3.2.1 Mesures physiologiques

La proportion de masse sèche moyenne du mégagamétophyte augmente de 15 % environ entre fin juillet - début août et fin septembre, puis se stabilise jusqu'à décembre (à environ 76 % de la masse fraîche). Dans le même temps, la masse sèche de l'EZ augmente de 100 % environ entre fin juillet - début août et fin septembre, puis se stabilise jusqu'à décembre (à environ 68 % de la masse fraîche). Les courbes de masses sèches de l'EZ et du mégagamétophyte montrent une évolution parallèle à partir de fin septembre, même si la proportion de masse sèche du mégagamétophyte reste supérieure d'environ 10 % à celle de l'EZ (**Figure III.3.A**) Une courbe de même allure se dessine concernant la teneur en eau moyenne du mégagamétophyte et de l'EZ au cours de leur développement (**Figure III.3.B**).

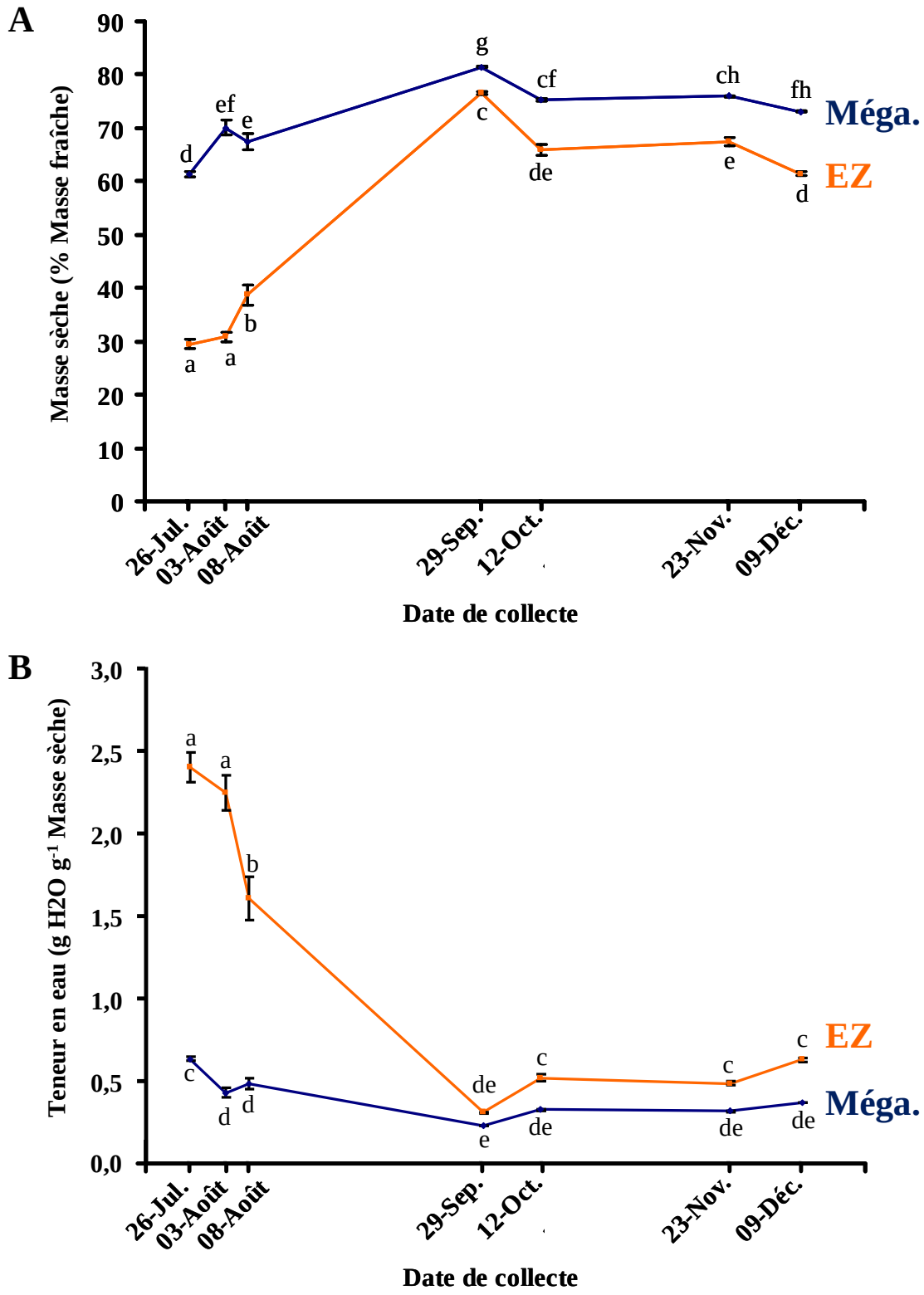


Figure III.3 Evolution de la masse sèche (**A**) et de la teneur en eau (**B**) de l'embryon zygotique cotylédonaire **(— EZ)** (tirée de Morel et al, en préparation, chapitre III.2, **Figure 1**) et du mégagamétophyte **(— Méga.)** au cours du développement de la graine de pin maritime. Les barres représentent l'intervalle de confiance à 95 %. Les lettres représentent les groupes statistiques définis par méthode de comparaison multiple de moyenne (P<0.05, n = 5).

III.3.2.2 Contenu en sucres solubles

La proportion de sucres solubles du mégagamétophyte est bien plus faible que pour l'EZ. Les sucres majeurs constituant le mégagamétophyte sont le saccharose, le raffinose et le stachyose, à des teneurs plus faibles que l'EZ. Contrairement au profil de l'EZ, le raffinose semble présent depuis le début de la maturation (juillet-août) chez le mégagamétophyte, alors que le stachyose ne commence à s'accumuler qu'à partir de novembre (**Figure III.4.A** et **III.4.B**).

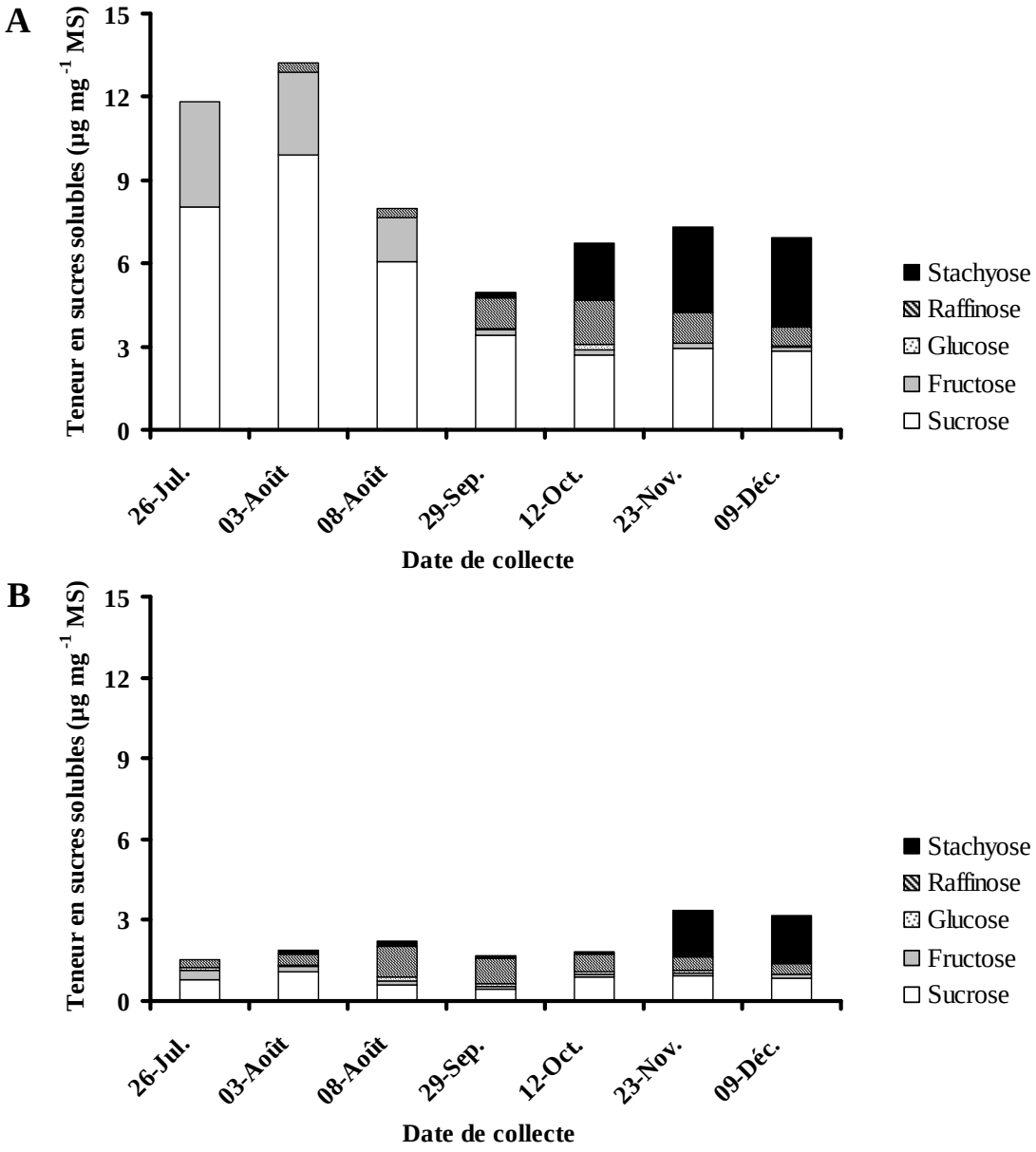


Figure III.4 Evolution du contenu en sucres solubles de l'embryon zygotique cotylédonaire (**A**) (tirée de Morel et al, en préparation, chapitre III.2, **Figure 2**) et du mégagamétophyte (**B**) au cours du développement de la graine de pin maritime. **MS :** masse sèche

III.3.2.3 Contenu en protéines totales et en protéines de réserve

La teneur en protéines totales mesurée dans le mégagamétophyte en début de maturation augmente brusquement (d'environ 70 %) entre le 3 août et le 8 août. Elle va ensuite graduellement augmenter jusqu'à un maximum atteint en milieu de maturation de la graine (fin septembre) pour diminuer lentement jusqu'au stade final (d'environ 25 % entre septembre et décembre) (**Figure III.5**).

Les courbes de teneur en protéines totales de l'EZ et du mégagamétophyte au cours de la maturation de la graine évoluent de façon opposée. A partir de fin septembre, la teneur en protéines totales de l'EZ augmente à mesure que celle du mégagamétophyte diminue, ces teneurs finissent par se recouvrir aux deux dernières dates analysées (novembre et décembre). La teneur en protéines totales de l'EZ semble évoluer conjointement avec celle du mégagamétophyte (**Figure III.5**).

Une évolution similaire se retrouve pour les protéines de réserves qui sont les protéines majoritaires constituant la graine aussi bien dans l'EZ que dans le mégagamétophyte. L'analyse des profils de migration éléctrophorétique des extraits de protéines totales du mégagamétophyte révèle quatre bandes majoritaires qui ne varient que faiblement au cours de la maturation. Ces 4 bandes ont été excisées et identifiées par LC-MS/MS. Les protéines correspondantes appartiennent toutes à la famille des protéines de réserve :

-l'homologue d'une protéine de type viciline de masse moléculaire 50,2 kDa de *Picea glauca* (accession Q40873) est présente dans la bande n°1

- l'homologue d'une légumine de pin de masse moléculaire 54,7 kDa de *Pinus strobus* (accession Q41017), dissociée dans nos conditions de migration en deux sous-unités de masses moléculaires 35 et 20 kDa est présente dans les bandes n°2 et n°3

- l'homologue d'une albumine de type 3 de masse moléculaire 19,9 kDa de *Pinus strobus* (accession Q40997) est présente dans la bande n°4.

Le maximum d'intensité pour ces 4 bandes majoritaires est obtenu aux dates du 08 août, 29 septembre et 12 octobre, qui correspondent aux dates où les teneurs en protéines totales les plus fortes (**Figure III.6**).

Les profils de migration éléctrophorétique des extraits de protéines totales de l'EZ sont très similaires à ceux du mégagamétophyte, présentant les 4 mêmes bandes majoritaires. La différence notable est l'accumulation plus tardive dans l'EZ. Ces bandes ont aussi été excisées et présentent les mêmes identifications en spectrométrie de masse que celles du mégagamétophyte (tirée de Morel et al, en préparation, chapitre III.2, **Table 3**).

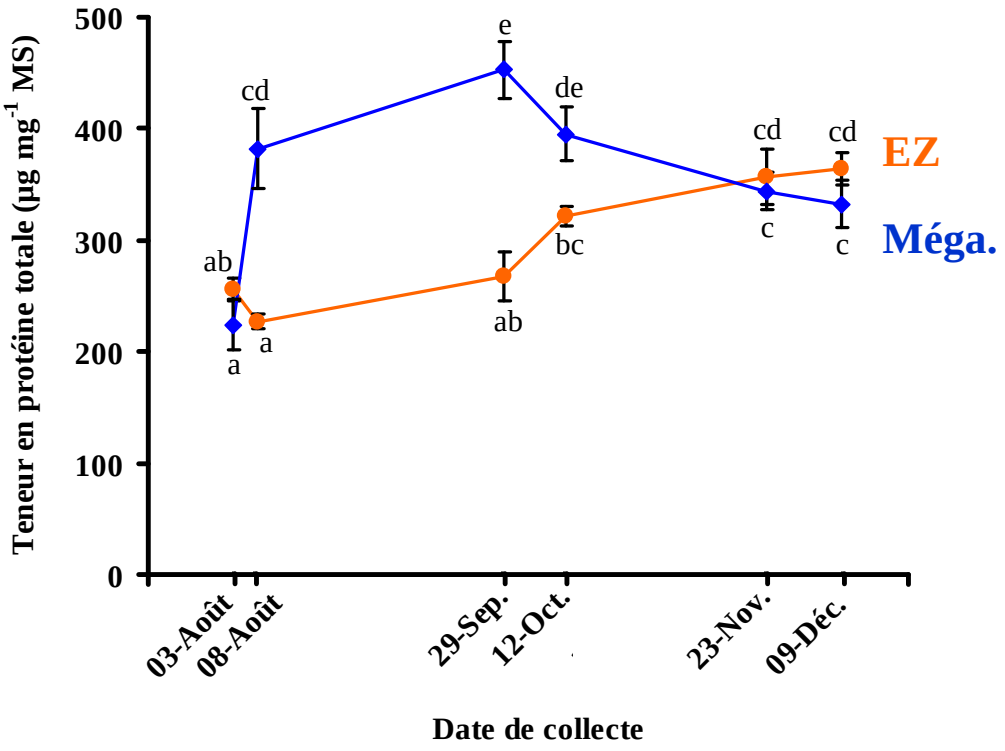


Figure III.5 Evolution des teneurs en protéines totales de l'embryon zygotique cotylédonaire **(— EZ)** (tirée de Morel et al, en préparation, chapitre III.2, **Table 2**) et du mégagamétophyte au cours du développement de la graine et du mégagamétophyte **(— Méga.)** au cours du développement de la graine de pin maritime. Les barres représentent l'intervalle de confiance à 95 %. Les lettres représentent les groupes statistiques définis par la méthode de comparaison multiple de moyenne (P<0.05, n=5).

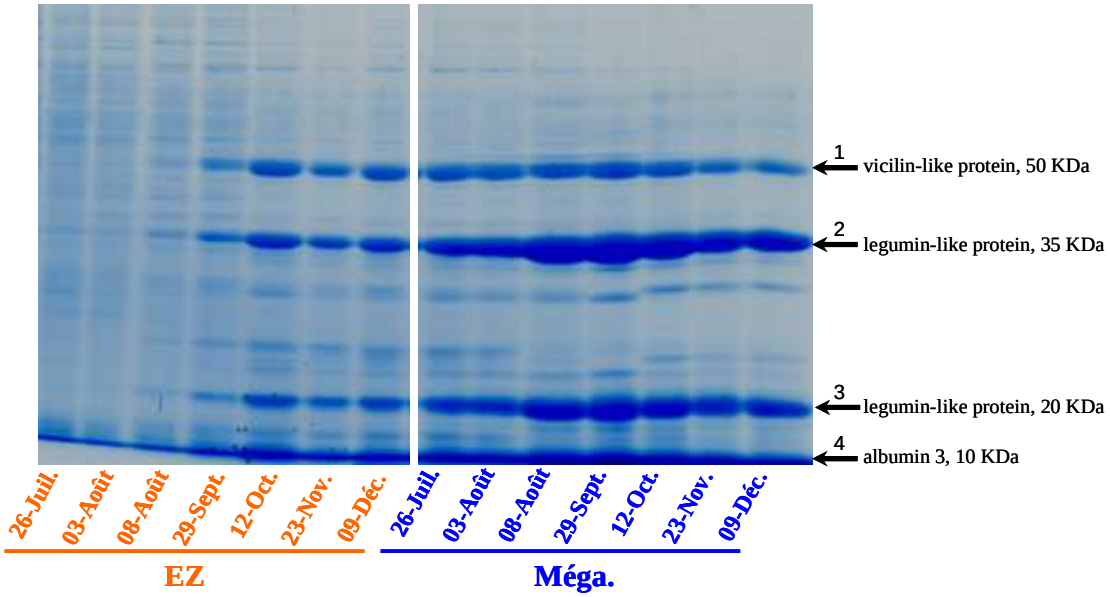


Figure III.6 Profils de migration d'extraits de protéines totales en SDS-PAGE de l'embryon zygotique cotylédonaire **(— EZ)** (tirée de Morel et al, en préparation, chapitre III.2, **Figure 3**) et du mégagamétophyte **(— Méga.)** au cours de la maturation de la graine de pin maritime. Les flèches indiquent les bandes excisées et les protéines correspondantes identifiées par spectrométrie de masse (LC-MS/MS).

CHAPITRE IV – Discussion générale

IV.1 La diminution de la disponibilité en eau n'est pas perçue comme une contrainte par les ES indifférenciés

En se basant sur les observations macroscopiques et histologiques ainsi que sur les analyses de masse sèche et de teneur en eau, les ES apparaissent mieux développés sur milieu à forte teneur en gélifiant (Morel et al, 2014, sous presse, chapitre II.2, **Figure 1**). De plus, le ratio hexose : saccharose est significativement plus faible sur ce milieu ; il a été montré que ce ratio diminuait au cours du développement de l'ES chez *Picea mariana* et *Picea glauca* (Iraqi et Tremblay, 2001b) et chez *Picea abies* (Lipavskà et al., 2000), alors qu'il augmentait au cours de la prolifération des ES indifférenciés (Lipavskà et Konradova, 2004). La maturation des ES est accompagnée par des changements dans la teneur en glucides à la fois dans le milieu de culture et dans les ES (Lipavskà et Konradova, 2004). Ainsi, la teneur faible en hexoses des ES immatures de notre étude marquerait le début de leur différenciation observée en microscopie.

Nos résultats confirment ainsi pour le pin maritime, ce qui a été précédemment décrit au cours de l'embryogenèse somatique de *Pinus strobus* (Klimaszewska et al., 2000) et *Pinus sylvestris* (Lelu-Walter et al., 2008) où les fortes concentrations en gélifiant entrainent une disponibilité en eau réduite et permettent l'amélioration du développement des ES. La teneur en gélifiant impacte ainsi significativement le développement embryonnaire chez les conifères.

L'augmentation de la teneur en gélifiant réduit l'eau du milieu disponible pour les ES immatures en différenciation, entrainant la diminution significative de leur teneur en eau par rapport au milieu à faible teneur en gélifiant. Cette limitation de l'eau pour les ES indifférenciés est-elle perçue comme une contrainte hydrique ? L'ABA est connu pour être un facteur de réponse aux stress abiotiques et en particulier à la déshydratation de la graine (Finkelstein et al., 2002) et dans les ES (Rai et al., 2011), les réponses à ces stress étant généralement rapides. Ainsi chez les ES de mélèze hybride, la synthèse endogène d'ABA est observable quelques heures après leur déshydratation, et est associée à une très faible teneur en eau des ES (Dronne et al., 1997). Dans notre étude, le pic d'ABA endogène n'est détecté qu'après 3 semaines de maturation et pour des ES ne présentant pas de réduction forte de leur

teneur en eau. Par conséquent, nous faisons l'hypothèse que ce signal ABA serait plutôt lié à la mise en place de l'embryogenèse *sensus stricto* plutôt qu'à la seule réponse à un déficit hydrique. La différenciation des têtes polaires embryonnaires et la structuration des suspenseurs, visibles en histologie après 3 semaines de maturation, confortent cette hypothèse. L'environnement hydrique ne serait donc pas perçu comme une condition stressante par les ES indifférenciés au cours de leur maturation. Ce fait est étayé par les récents résultats obtenus chez le mélèze hybride pour des ES indifférenciés (1 semaine maturation). En effet, l'analyse protéomique de ces ES a montré que comparées aux conditions de culture en présence d'une faible teneur en gélifiant (4 g L^{-1}), les conditions de culture en présence d'une forte teneur en gélifiant (8 g L^{-1}) n'étaient pas stressantes (Teyssier et al., 2011).

IV.2 La diminution de la teneur en eau impacte-elle le métabolisme des sucres ?

Chez la carotte, une faible disponibilité en eau et en source de carbone améliore fortement la maturation et la germination des ES (Lee et al., 2001). Les résultats de cette étude suggèrent qu'un déficit nutritionnel temporaire pourrait promouvoir le développement des ES *in vitro* pour un large éventail d'espèces végétales. De la même façon, l'augmentation du gélifiant dans notre système d'étude, qui contribue à limiter l'apport en eau aux ES, pourrait limiter en même temps l'apport en nutriments qui sont sous forme soluble dans le milieu. Cette limitation en eau et en nutriments deviendrait alors favorable au développement des ES. De plus chez l'hévéa il a été montré qu'un apport limité en maltose en début de maturation contribuait à la mise en place de l'embryogenèse somatique (Blanc et al., 2002). Dans notre étude, les ES indifférenciés cultivés sur **9G** montrent une composition en sucres solubles (fructose et glucose) significativement plus faible que celle des ES indifférenciés cultivés sur milieu **4G**, après 1 semaine de maturation (Morel et al, 2014, sous presse, chapitre II.2).

L'importation des sucres contenus dans le milieu de culture joue un rôle primordial pour les embryons en culture. En effet, le fructose et le glucose sont libérés dans le milieu à partir de l'invertase acide de la paroi cellulaire des ES indifférenciés en culture qui hydrolyse partiellement le saccharose du milieu provoquant un potentiel osmotique plus négatif (Iraqi et Tremblay, 2001b; Pullman et al., 2003). Cette hydrolyse des sucres est importante car elle leur permet (**1**) d'entrer directement dans des voies métaboliques spécifiques, (**2**) de modifier

le potentiel osmotique à la fois du milieu et des cellules des ES et (**3**) et de jouer un rôle dans la signalisation cellulaire par des systèmes de « sugar-sensing ». Le contrôle du potentiel osmotique, est une des conditions à la mise en place de l'embryogenèse somatique chez les conifères (Gupta et Pullman, 1990; Gupta et Pullman, 1991; Pullman et Buchanan, 2008). Dans notre système de culture, le saccharose n'est pas en concentration limitante (0.2 M). Cette forte concentration qui augmente la pression osmotique du milieu, associée à la réduction de la disponibilité en eau et l'ajout d'un signal hormonal (ABA, 80 µM) induisent la différenciation des ES immatures. Cependant sur **4G**, ces fortes conditions osmotiques ne sont pas suffisantes pour induire la différenciation. Plusieurs travaux montrent que la disponibilité en nutriments, et particulièrement en hexoses, agit comme le premier signal pour l'induction de l'expression des gènes chez les plantes (Graham, 1996; Gibson, 2004) La disponibilité limitée en saccharose sur **9G** pourrait ainsi agir en tant que signal de l'embryogenèse. La baisse de la teneur en hexoses endogènes pourrait être un élément clé dans la réorientation du métabolisme cellulaire. Le saccharose non limitant dans le milieu de culture (0.2 M) sur milieu **4G** pourrait être rapidement assimilé dans les ES leur permettant de proliférer. Prolifération qui, très consommatrice en énergie, serait impossible sur milieu **9G**, compte tenu de la disponibilité réduite en sucres. Les ES indifférenciés s'adapteraient alors à cette limitation en modifiant leur programme d'expression génique vers le développement embryonnaire.

La teneur en amidon est significativement plus forte dans les ES indifférenciés cultivés sur milieu **9G** comme nous l'avons observé en microscopie (voir chapitre II.2). Chez *Picea mariana* et *Picea. glauca*, une accumulation d'amidon a lieu principalement pendant les 2 premières semaines de maturation, la teneur restant à peu près constante par la suite (Iraqi et Tremblay, 2001b). Une caractérisation des premiers stades du développement embryonnaire de l'ES du dattier a montré que la synthèse de l'amidon prédomine et puis diminue progressivement (Sghaier-Hammami et al., 2009). Une observation semblable a été faite dans différentes graines en développement (soja, sapin chinois, luzerne) pour lesquelles l'expression transitoire de protéines de synthèse de l'amidon a été observée à un stade précoce (Adams et al., 1980; Gallardo et al., 2007; Shi et al., 2010). De la même façon, chez le pin maritime, l'accumulation d'amidon est un processus mis en place de façon précoce dans le développement de l'ES. De plus, des auteurs travaillant sur l'embryogenèse chez *Arabidopsis* ont suggéré que l'accumulation précoce d'amidon aurait une fonction directe pour la division cellulaire et à la différenciation cellulaire plutôt qu'une fonction de réserve spécifiques pour

les embryons en développement comme classiquement admis (Andriotis et al., 2010). Ensuite, une autre étude chez le pin maritime a montré que l'accumulation de l'amidon est en corrélation avec l'activité organogénique du méristème apical et le développement et la croissance de nouveaux organes de l'embryon depuis l'apex (Jordy, 2004). Par conséquent, l'accumulation précoce d'amidon pourrait dans notre étude marquer la différenciation cellulaire des embryons.

IV.3 Les processus cellulaires et moléculaires clés impliqués dans le développement précoce des embryons somatiques

Les différents résultats que nous avons obtenus lors des 4 premières semaines de maturation des ES de pin maritime ont été résumés dans la **Figure IV.1** Si l'approche multi-échelle que nous avons développé est originale, l'intégration des différents niveaux d'études entre eux demeure difficile et limite les interprétations. Ainsi, relier entre eux les différents niveaux d'observation (macroscopique, cellulaire, moléculaire) n'est pas toujours évident, la temporalité de réponse n'étant pas la même selon l'échelle observée. Par exemple, de nombreux transcrits impliqués dans la différenciation embryonnaire ont été identifiés à 1 semaine de maturation, alors qu'aucun changement n'est visible aux niveaux histologiques ou physiologiques (MS et TE). Par opposition, il s'est avéré plus évident de relier certains transcrits et protéines identifiés impliqués dans le métabolisme des sucres et de l'amidon aux différences de teneur en sucres et en amidon observées entre les 2 conditions testées à 1 semaine de maturation. Ainsi les données moléculaires, métaboliques et cellulaires ne peuvent pas toujours être mises en relation directe avec les données physiologiques et biochimiques. Par conséquent nous n'avons pu reconstituer qu'une partie des interactions qui se créent entre les différents niveaux de complexité biologiques au cours de l'embryogenèse.

Parmi les 212 298 séquences assemblées lors du travail de transcriptomique, 632 (0.3 %) ont été considérées comme différentiellement exprimées dans les ES cultivés sur milieu **9G**, parmi lesquels 239 (38 %) sont ressorti de l'analyse fonctionnelle par enrichissement GO (Gene Ontology). La majorité de ces transcrits étaient impliqués dans la croissance cellulaire et le développement relatifs à l'embryogenèse, la régulation de la réponse aux stimuli et le métabolisme secondaire et le rôle fonctionnel de 34 d'entre eux (5 %) a été discuté dans Morel et al, 2014 (sous presse, chapitre II.2). Tous n'ont pas pu être discutés soulignant la richesse d'information apportée par les données transcriptomiques de ce travail. C'est

pourquoi, j'ai réutilisé ces données afin d'aller plus loin dans leur exploitation. La **Table IV.1** regroupe 194 de ces transcrits classés selon leur fonction potentielle dans le développement embryonnaire. J'ai regroupé ces fonctions en 5 catégories : la reprogrammation génétique, les voies de signalisation et de transduction du signal, la structuration cellulaire la synthèse, l'adressage, le repliement et le stockage des protéines et la protection cellulaire. Les valeurs moyennes de FPKM (moyFPKM9), calculées en regroupant les transcrits codant des protéines appartenant à la même famille ou de même nom, donnent une meilleure estimation des différences d'expression entre les 2 conditions de maturation testées. Le ratio 9/4 est calculé pour chaque transcrit isolé par le rapport du FPKM moyen attribué au transcrit exprimé dans les ES cultivées sur milieu **9G** sur le FPKM moyen attribué au transcrit exprimé dans les ES cultivées sur milieu **4G** (FPKM9/ FPKM4). De la même façon, j'ai réutilisé les données de protéomique et 19 protéines ont été classées dans la **Table IV.2** selon les mêmes catégories. Le ratio 9/4 est calculé pour chaque protéine identifiée par le rapport du volume de spot sur **9G** sur le volume de spot sur **4G**. Je vais maintenant discuter de l'implication de certains de ces transcrits et protéines (en **bleu** dans les tables) dans le développement embryonnaire. Cette analyse sous un nouvel angle des données moléculaires (transcriptomique et protéomique) vient compléter et renforcer les conclusions définies dans Morel et al, 2014 (sous presse, chapitre II.2).

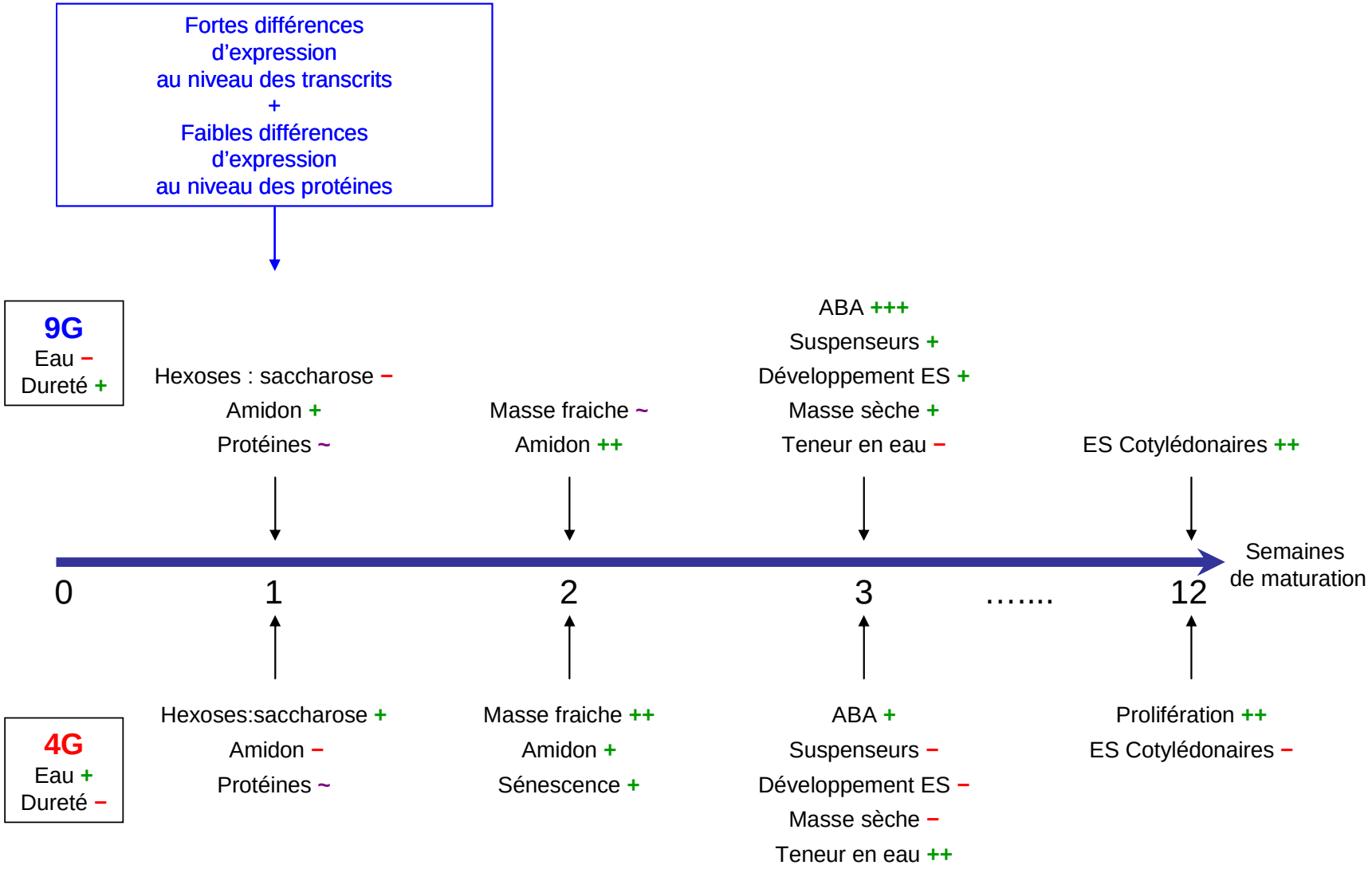


Figure IV.1. Mise en place chronologique des événements intervenants dans l'embryogenèse somatique précoce du pin maritime au cours de la maturation : les symboles représentent les variations entre conditions (**4G** et **9G**) et d'une semaine de culture à l'autre : **(–)** diminution ; **(+)** augmentation ; **(~)** pas de variation ; **(9G)** milieu de maturation solidifié avec 9 g L^{-1} de gélifiant ; **(4G)** milieu de maturation solidifié avec 4 g L^{-1} de gélifiant.

Table IV.1. Transcrits identifiés surexprimés dans les ES après 1 semaine de maturation sur milieu à forte teneur en gélifiant **(9G)**

Description	Accession	Contig	FPKM9	FPKM4	P-Value	Ratio 9/4	moy FPKM9	Fonctions	Catégories
methyl esterase 6	Q8S8S9	SC_00001884.1.1	47	19	0,0362	2,5	*52*	défense de la plante	protection cellulaire
methyl esterase 6	Q9LUV2	SC_00042371.1.1	56	27	0,0244	2,1			
tir-nbs-lrr class disease resistance protein	F4J910	SC_00009446.1.1	10	3	0,0374	3,7	*35*	défense de la plante	protection cellulaire
tir-nbs-lrr class disease resistance protein	F4J361	SC_00058122.1.1	3	1	0,0407	3,1			
tir-nbs-lrr class disease resistance protein	F4J910	SC_00098360.1.1	123	46	0,0429	2,7			
tir-nbs-lrr class disease resistance protein	F4IFF6	SC_LRC40_DANRE.1.2	3	1	0,0221	2,2			
wound-responsive-related protein	Q9SFV3	SC_PR4_PHAVU.9.12	39	15	0,0463	2,6		défense de la plante	protection cellulaire
heavy metal transport detoxification domain-containing	Q94BT9	SC_00003722.1.1	7	4	0,0303	2,0	*12*	détoxication	protection cellulaire
heavy metal transport detoxification domain-containing	Q94BT9	SC_00003952.1.1	11	5	0,0467	2,4			
heavy metal transport detoxification domain-containing	Q94BT9	SC_00160008.1.1	18	5	0,0292	3,4			
em-like protein gea6	Q02973	SC_LE193_HORVU.1.1	50	12	0,0205	4,3		protection contre la dessiccation	protection cellulaire
hsp20-like chaperone	P19037	SC_00151185.1.1	10	3	0,0303	3,6	*17*	protection contre la dessiccation	protection cellulaire
hsp20-like chaperone	P19037	SC_HS181_ORYSJ.3.10	34	8	0,0437	4,5			
hsp20-like chaperone	P19037	SC_HS181_ORYSJ.5.10	8	2	0,0267	3,6			
late embryogenesis abundant domain-containing protein	Q9SUB2	SC_CYP306A1_BOMMO.407.3286	22	7	0,0283	3,2	*23*	protection contre la dessiccation	protection cellulaire
late embryogenesis abundant group 1 protein	Q39138	SC_LEA11_HELAN.2.5	32	9	0,0360	3,4			
late embryogenesis abundant group 6	Q8S8R1	SC_00025014.1.1	7	2	0,0063	3,5			
late embryogenesis abundant protein	F4JB74	SC_00013261.1.1	7	3	0,0027	2,4			
late embryogenesis abundant protein	F4JB74	SC_00033538.1.1	44	16	0,0382	2,8			

adenine nucleotide alpha hydrolases-like protein	Q93W91	SC_00003596.1.1	16	6	0,0130	2,6		réponse au stress	protection cellulaire
myb proto-oncogene protein	Q9SDS8	SC_00041622.1.1	62	27	0,0008	2,3		réponse au stress	protection cellulaire
zinc finger (c3hc4-type ring finger) family protein	Q8GW10	SC_00001272.1.1	22	9	0,0486	2,3	*11*	réponse au stress	protection cellulaire
zinc finger a20 and an1 domain-containing stress-associated	Q9SJM6	SC_SAP6_ARATH.1.3	3	2	0,0307	2,0			
zinc finger a20 and an1 domain-containing stress-associated	Q9SJM6	SC_SAP6_ARATH.3.3	8	3	0,0025	2,5			
bax inhibitor-1 motif-containing protein	Q9LD45	SC_BI1_ORYSJ.11.15	4	1	0,0064	2,5		réponse au stress hydrique	protection cellulaire
dehydration-induced protein erd15	Q39096	SC_00000544.1.1	***7***	3	0,0343	2,1		réponse au stress hydrique	protection cellulaire
dehydrin 2	E1A569	SC_00081475.1.1	71	24	0,0084	2,9	***56***	réponse au stress hydrique	protection cellulaire
dehydrin cor47	P31168	SC_00001231.1.1	40	17	0,0325	2,3			
water deficit inducible protein LP3	O24314	SC_00144857.1.1	13	3	0,0000	3,9	***35***	réponse au stress hydrique	protection cellulaire
water deficit inducible protein LP3-3	Q41095	SC_ASR2_SOLLC.2.31	19	3	0,0099	6,9			
water-stress inducible protein 3	E0A741	SC_ASR2_SOLLC.14.31	73	26	0,0173	2,8			
glutathione s-transferase (class phi) 14	Q9C6C8	SC_00026575.1.1	14	7	0,0235	2,2	***17***	réponse au stress oxydatif	protection cellulaire
glutathione s-transferase dhar1	Q93VJ2	SC_UMC1467_MAIZE.1.1	20	9	0,0295	2,1			
peroxidase 12	Q96520	SC_PER12_ARATH.2.14	10	3	0,0231	3,4	***14***	réponse au stress oxydatif	protection cellulaire
peroxidase 51	O23609	SC_00055627.1.1	15	5	0,0004	2,8			
peroxidase 57	Q43729	SC_00077646.1.1	19	3	0,0272	5,5			
peroxiredoxin -like 2 domain-containing protein	Q8L5Y8	SC_00003923.1.1	3	1	0,0139	2,2	*3*	réponse au stress oxydatif	protection cellulaire
peroxiredoxin-2C	Q9SRZ4	SC_PRX2C_ORYSJ.2.15	3	1	0,0462	2,6			
2og-fe oxygenase family protein	Q9S818	SC_FL3H_ARATH.5.7	7	2	0,0147	3,8	*7*	métabolisme secondaire	protection cellulaire
2og-fe oxygenase family protein	Q9S818	SC_FL3H_MALDO.12.23	5	1	0,0112	3,2			
2og-fe oxygenase family protein	Q9S818	SC_FL3H_MALDO.2.23	8	2	0,0319	4,5			

2og-fe oxygenase family protein	B1GV57	SC_FLS_CITUN.2.7	8	3	0,0385	2,7			
2-oxoglutarate family protein	Q9S818	SC_FL3H_MALDO.19.23	13	3	0,0004	4,7	*11*	métabolisme secondaire	protection cellulaire
2-oxoglutarate and fe -dependent oxygenase-like protein	Q9S818	SC_FL3H_PETHY.5.8	8	2	0,0033	3,4			
disease resistance- dirigent domain-containing protein	Q84TH6	SC_00006247.1.1	4	1	0,0066	2,9	*3*	métabolisme secondaire	protection cellulaire
disease resistance- dirigent domain-containing protein	Q84TH6	SC_00105899.1.1	3	1	0,0056	2,2			
aspartyl protease family protein	Q9ZVS5	SC_00025099.1.1	2	1	0,0439	2,2	*9*	recyclage protéique	recyclage/synthèse/modification des protéines
aspartyl protease family protein	Q9ZVS4	SC_00028565.1.1	8	4	0,0445	2,2			
aspartyl protease-like protein	Q9LTW4	SC_00027439.1.1	16	8	0,0361	2,0			
carboxypeptidase d	Q0WRX3	SC_CYP306A1_BOMMO.843.3286	13	3	0,0022	4,0	*8*	recyclage protéique	recyclage/synthèse/modification des protéines
carboxypeptidase d	Q93Y09	SC_SCP45_ARATH.2.9	4	1	0,0168	3,0			
polyubiquitin 14	Q9SHE7	SC_RUB1_ARATH.2.6	9	4	0,0325	2,0	*9*	recyclage protéique	recyclage/synthèse/modification des protéines
polyubiquitin 14	Q9SHE7	SC_RUB1_ARATH.3.6	9	4	0,0466	2,3			
pppde thiol peptidase family protein	Q9SFT9	SC_PCO141631_MAIZE.1.2	11	3	0,0115	3,4		recyclage protéique	recyclage/synthèse/modification des protéines
rhomboid-related intramembrane serine protease-like protein	Q9FX61	SC_00004453.1.1	2	1	0,0146	2,9		recyclage protéique	recyclage/synthèse/modification des protéines
thiol protease aleurain	Q058I7	SC_RD19A_ARATH.8.18	7	4	0,0107	2,0		recyclage protéique	recyclage/synthèse/modification des protéines
ubiquitin family protein	Q9M1U1	SC_UBL5_ARATH.26.27	12	6	0,0096	2,0	*27*	recyclage protéique	recyclage/synthèse/modification des protéines
ubiquitin family protein	Q9M1U1	SC_UBL5_ARATH.27.27	22	11	0,0028	2,1			
ubiquitin family protein	Q9M1U1	SC_UBL5_ARATH.8.27	103	51	0,0251	2,0			
ubiquitin-conjugating enzyme e2 35	Q94A97	SC_UBC35_ARATH.9.14	4	2	0,0176	2,1			
ubiquitin-conjugating enzyme e2 8	P35132	SC_00000189.1.1	19	8	0,0314	2,3			
ubiquitin-conjugating enzyme e2 8	P42745	SC_00013883.1.1	3	1	0,0409	2,0			
ubiquitin-conjugating enzyme e2 8	Q94F47	SC_00016734.1.1	4	2	0,0146	2,5			

ubiquitin-conjugating enzyme e28	Q9FKT3	SC_00034141.1.1	49	15	0,0274	3,2			
ubiquitin-conjugating enzyme e28	Q9FMM0	SC_00040622.1.1	4	2	0,0113	2,4			
ubiquitin-conjugating enzyme e28	Q9FI61	SC_00155240.1.1	4	1	0,0037	4,0			
ubiquitin-conjugating enzyme e28	P35131	SC_UBC10_ARATH.2.4	41	17	0,0418	2,4			
ubiquitin-conjugating enzyme e28	Q9FKT3	SC_UBC30_ARATH.10.10	9	4	0,0243	2,4			
ubiquitin-conjugating enzyme e28	Q94F47	SC_UBC30_ARATH.3.10	87	36	0,0172	2,4			
ubiquitin-conjugating enzyme e28	P35132	SC_UBC8_ARATH.1.24	24	12	0,0114	2,0			
ubiquitin-conjugating enzyme e28	P35131	SC_UBC8_ARATH.10.24	14	7	0,0355	2,1			
ubiquitin-conjugating enzyme e28	P35131	SC_UBC8_ARATH.12.24	15	6	0,0063	2,5			
ubiquitin-conjugating enzyme e28	F4JKF3	SC_UBC8_ARATH.17.24	48	19	0,0038	2,5			
ubiquitin-conjugating enzyme e28	P35131	SC_UBC8_ARATH.3.24	31	11	0,0011	2,9			
ubiquitin-conjugating enzyme e28	P35132	SC_UBC8_ARATH.4.24	33	16	0,0160	2,1			
ubiquitin-conjugating enzyme e28	P35131	SC_UBC8_ARATH.6.24	26	13	0,0309	2,0			
ubiquitin-conjugating enzyme e28	P35131	SC_UBC9_ARATH.14.24	8	4	0,0010	2,2			
26S ribosomal RNA gene		SC_LOC100275739_MAIZE.44.160	6	2	0,0229	2,3	***25***	synthèse protéique	recyclage/synthèse/modification des protéines
26S ribosomal RNA gene		SC_LOC100275739_MAIZE.49.160	100	32	0,0293	3,1			
26S ribosomal RNA gene		SC_LOC100275739_MAIZE.55.160	3	1	0,0197	2,2			
26S ribosomal RNA gene		SC_LOC100275739_MAIZE.63.160	7	3	0,0450	2,5			
26S ribosomal RNA gene		SC_LOC100275739_MAIZE.95.160	11	4	0,0403	2,4			
elongation factor 1-alpha	F4HUA0	SC_00000762.1.1	9	4	0,0357	2,0	***6***	synthèse protéique	recyclage/synthèse/modification des protéines
elongation factor 1-alpha	Q0WL56	SC_00149373.1.1	4	2	0,0498	2,5			
elongation factor 1-delta 1	Q9SI20	SC_00027980.1.1	4	2	0,0079	2,5			
elongation factor 1-delta 1	Q9SI20	SC_EF1D_PIMBR.17.70	5	2	0,0332	2,6			

translation initiation factor eif-2 beta subunit	F4K6W6	SC_IF2B_WHEAT.6.7	4	2	0,0011	2,2	*11*	synthèse protéique	recyclage/synthèse/modification des protéines
translation initiation factor eif-5	Q9S825	SC_IF5Z_ARATH.3.6	18	8	0,0154	2,3			
syntaxin 1b 2 3	Q8VZU2	SC_SY132_ARATH.1.3	6	3	0,0061	2,3	*7*	transport protéique	recyclage/synthèse/modification des protéines
syntaxin of plants syp5	Q94KK7	SC_00161130.1.1	8	3	0,0414	2,5			
methionine synthase 2	Q9SRV5	SC_METE_MESCR.27.30	5	2	0,0180	2,2	*13*	méthylation de l'ADN	reprogrammation génétique
methionine synthase 2	O50008	SC_METE_SOLSC.2.72	10	3	0,0445	2,9			
methionine synthase 2	O50008	SC_METE_SOLSC.4.72	14	5	0,0019	2,9			
methionine synthase 2	O50008	SC_METE_SOLSC.54.72	9	4	0,0459	2,1			
methionine synthase 2	Q9SRV5	SC_OS12G0623900_ORYSJ.8.15	29	9	0,0133	3,3			
s-adenosyl-l-methionine-dependent methyltransferase-like	Q9ZVC3	SC_00007904.1.1	88	44	0,0025	2,0		méthylation de l'ADN	reprogrammation génétique
histone h2a	Q9LHQ5	SC_H2A6_ORYSJ.1.9	3	2	0,0155	2,1	***8***	remodelage de la chromatine	reprogrammation génétique
histone h3	B9DGR9	SC_H33_ARATH.29.70	15	6	0,0036	2,4			
histone h4	P59259	SC_00079928.1.1	6	2	0,0379	3,0			
histone h4	A8MRV1	SC_H4_OLILU.1.5	8	4	0,0234	2,1			
cbl-interacting serine threonine-protein kinase 3	B3H6F9	SC_SAPK1_ORYSJ.1.1	8	2	0,0494	4,5	*24*	voie de signalisation dépendante de l'ABA	signal/transduction/régulation
cbl-interacting serine threonine-protein kinase 3	Q9FFP9	SC_SRK2E_ARATH.1.3	55	12	0,0475	4,5			
cbl-interacting serine threonine-protein kinase 3	Q940H6	SC_SRK2E_ARATH.3.3	29	5	0,0221	5,4			
cbl-interacting serine threonine-protein kinase 9	Q9SN43	SC_CIPKC_ARATH.4.4	4	2	0,0112	2,4			
protein phosphatase	Q9CAJ0	SC_P2C08_ORYSJ.7.9	14	5	0,0130	3,0	*28*	voie de signalisation dépendante de l'ABA	signal/transduction/régulation
protein phosphatase 2c 77	Q9CAJ0	SC_P2C08_ORYSJ.2.9	43	19	0,0149	2,2			
12-oxophytodienoate reductase 3	Q9FUP0	SC_00061885.1.1	8	3	0,0211	2,4	*8*	biosynthèse de l'acide jasmonique	signal/transduction/régulation
12-oxophytodienoate reductase 3	Q9FUP0	SC_00147094.1.1	10	3	0,0427	2,8			

12-oxophytodienoate reductase 3	Q9FUP0	SC_OPR3_ARATH.2.6	7	3	0,0058	2,4			
12-oxophytodienoate reductase 3	Q9FUP0	SC_OPR3_SOLLC.2.3	8	2	0,0253	3,5			
Homeodomain transcription factor	Q14FJ6	SC_CYP306A1_BOMMO.1085.3286	21	8	0,0021	2,8		contrôle la division cellulaire	signal/transduction/régulation
wuschel related homeobox 11	Q6X7K1	SC_00005814.1.1	32	16	0,0053	2,0	***20***	contrôle la division cellulaire	signal/transduction/régulation
wuschel related homeobox 11	Q6X7J3	SC_00011070.1.1	8	4	0,0142	2,2			
lob domain-containing protein 6	Q8L5T5	SC_LBD15_ARATH.2.6	2	1	0,0252	2,9		contrôle la séparation des organes	signal/transduction/régulation
nac domain containing protein 47	Q9SQQ5	SC_00004156.1.1	4	2	0,0149	2,4	*3*	contrôle le développement des téguments	signal/transduction/régulation
nac domain-containing protein 72	Q9CAR0	SC_00027363.1.1	3	1	0,0073	2,4			
ethylene-responsive transcription factor 2	O80338	SC_EF101_ARATH.1.1	6	2	0,0218	2,5	***5***	voie de signalisation dépendante de l'éthylène	signal/transduction/régulation
ethylene-responsive transcription factor 8	Q9MAI5	SC_00157909.1.1	4	2	0,0079	2,3			
gibberellin-regulated protein	Q8GWK5	SC_00030584.1.1	23	11	0,0164	2,1	*44*	voie de signalisation dépendante des gibbérellines	signal/transduction/régulation
gibberellin-regulated protein	F4IQJ4	SC_00033457.1.1	6	3	0,0291	2,3			
gibberellin-regulated protein	Q6NMQ7	SC_00053117.1.1	112	43	0,0032	2,6			
gibberellin-regulated protein	Q6NMQ7	SC_00151200.1.1	21	8	0,0430	2,7			
gibberellin-regulated protein	Q6NMQ7	SC_AT3G10185_ARATH.3.3	57	19	0,0069	3,0			
wrky transcription factor 18	Q1PEB9	SC_00015833.1.1	2	1	0,0236	2,6		privation en phosphate et en bore	signal/transduction/régulation
dormancy auxin associated protein	F4I3L7	SC_00018728.1.1	14	6	0,0010	2,3	*22*	réponse à l'auxine	signal/transduction/régulation
dormancy auxin associated protein	F4I3L7	SC_UMC2505_MAIZE.3.5	29	14	0,0034	2,1			
alpha- -glucan-protein synthase family protein	Q9LFW1	SC_00112388.1.1	14	6	0,0357	2,5	*6*	signalisation cellulaire	signal/transduction/régulation
alpha- -glucan-protein synthase family protein	Q9SRT9	SC_UPTG_MAIZE.10.12	3	1	0,0186	2,1			
alpha- -glucan-protein synthase family protein	Q9LFW1	SC_UPTG_MAIZE.2.12	3	1	0,0128	2,6			

alpha- -glucan-protein synthase family protein	Q9SRT9	SC_UPTG2_SOLTU.36.38	4	2	0,0459	2,2			
chitinase-like protein	O24658	SC_00008805.1.1	25	12	0,0208	2,1	***20***	signalisation cellulaire	signal/transduction/régulation
chitinase-like protein	Q9M2U5	SC_00025909.1.1	6	2	0,0216	3,7			
chitinase-like protein	Q9ZSI6	SC_CHI1_ORYSJ.4.9	18	6	0,0181	2,9			
chitinase-like protein	O22842	SC_CHI4_ORYSJ.1.7	25	4	0,0224	6,6			
chitinase-like protein	Q9ZSI6	SC_CHI8_ORYSJ.2.2	25	7	0,0165	3,5			
nodulin 21 -like transporter family protein	Q9FL41	SC_OS05G0493800_ORYSJ.1.2	5	1	0,0162	3,7	*11*	signalisation cellulaire	signal/transduction/régulation
nodulin 3-like protein	Q9FM10	SC_00016623.1.1	16	5	0,0081	3,2			
14-3-3-like protein gf14 epsilon	Q96300	SC_1433_LILLO.6.9	6	2	0,0279	2,9		transduction du signal	signal/transduction/régulation
calcium-binding protein cml27	Q9ZR02	SC_00024580.1.1	4	2	0,0459	2,1	***16***	transduction du signal	signal/transduction/régulation
calcium-binding protein cml27	Q9FYK2	SC_POLC2_JUNOX.2.31	45	20	0,0267	2,3			
calmodulin 5	Q1H5F3	SC_00043177.1.1	8	4	0,0413	2,1			
calmodulin 5	Q9FYK2	SC_CML10_ORYSJ.2.26	12	5	0,0135	2,4			
calmodulin 5	Q9FYK2	SC_POLC2_JUNOX.18.31	14	5	0,0013	2,9			
phospholipase a2 family protein	Q8S8N6	SC_LOC100284361_MAIZE.1.1	18	9	0,0308	2,0	***12***	catabolisme des lipides	structuration cellulaire
phospholipase c	Q8H965	SC_00034320.1.1	11	3	0,0158	4,5			
phospholipase c	O81020	SC_00122309.1.1	12	4	0,0152	3,3			
phospholipase c	O81020	SC_OS03G0852800_ORYSJ.1.1	8	2	0,0031	3,4			
aquaporin pip2-1	P25818	SC_00001064.1.1	9	4	0,0001	2,1	***6***	élongation cellulaire	structuration cellulaire
aquaporin pip2-1	P25818	SC_00002754.1.1	4	2	0,0075	2,4			
aquaporin pip2-1	P25818	SC_00015040.1.1	4	1	0,0135	3,1			
aquaporin pip2-1	P25818	SC_TIP11_ARATH.2.14	6	2	0,0109	2,9			

extensin-1	Q38913	SC_EXTN1_ARATH.1.4	10	1	0,0039	6,8		modification de la paroi cellulaire	structuration cellulaire
germin-like protein subfamily 1 member 1	P92998	SC_00010373.1.1	69	30	0,0035	2,3	*44*	modification de la paroi cellulaire	structuration cellulaire
germin-like protein subfamily 1 member 18	Q9FMB0	SC_GL91_ORYSJ.4.4	10	2	0,0035	4,2			
germin-like protein subfamily 2 member 2	Q9FZ27	SC_00006936.1.1	5	1	0,0143	3,1			
germin-like protein subfamily 2 member 4	A0JQ79	SC_GL11_ARATH.1.2	94	39	0,0084	2,4			
inositol oxygenase	B4F7Q2	SC_MIOX_ORYSJ.3.8	155	48	0,0332	3,3	***49***	modification de la paroi cellulaire	structuration cellulaire
inositol oxygenase 1	Q8L799	SC_00014095.1.1	24	8	0,0044	2,9			
inositol oxygenase 1	Q8L799	SC_00026140.1.1	11	5	0,0291	2,3			
inositol oxygenase 1	Q8L799	SC_00040126.1.1	34	16	0,0314	2,2			
inositol oxygenase 1	Q8L799	SC_00045575.1.1	14	6	0,0498	2,5			
inositol oxygenase 1	Q8L799	SC_MIOX_ORYSJ.1.8	59	25	0,0159	2,4			
lipid transfer protein	O22110	SC_00011326.1.1	5	3	0,0249	2,1	***5***	modification de la paroi cellulaire	structuration cellulaire
lipid transfer protein	O22111	SC_00154588.1.1	5	3	0,0365	2,0			
non-specific lipid-transfer protein 7	Q9LLR7	SC_NLTP_PINTA.219.283	12	5	0,0287	2,3	***11***	modification de la paroi cellulaire	structuration cellulaire
non-specific lipid-transfer protein 7	Q9XFS7	SC_NLTP_PINTA.5.283	9	4	0,0242	2,4			
non-specific lipid-transfer protein 7	Q9SCZ0	SC_NLTP2_LENCU.9.30	13	4	0,0204	3,4			
omega-hydroxypalmitate O-feruloyl transferase	F4JWW6	SC_00014965.1.1	6	2	0,0321	3,3	*9*	modification de la paroi cellulaire	structuration cellulaire
omega-hydroxypalmitate O-feruloyl transferase	Q9XEF2	SC_00061026.1.1	13	6	0,0249	2,3			
omega-hydroxypalmitate O-feruloyl transferase	Q9LFB5	SC_LOC100191955_MAIZE.1.1	9	4	0,0244	2,1			
pectinesterase 68	Q8LPF3	SC_00037025.1.1	4	2	0,0152	2,3			
xyloglucan:xyloglucosyl transferase	Q67Z90	SC_XTHA_PHAAN.3.10	13	4	0,0104	3,3			
mitotic-spindle organizing protein 1b	Q9M0N8	SC_MZT1_PICSI.1.3	6	3	0,0005	2,1			

protein brick 1	Q94JY4	SC_00063786.1.1	6	2	0,0332	3,7	*46*	modifications du cytosquelette	structuration cellulaire
protein brick 1	Q94JY4	SC_BRK1_ARATH.1.2	86	43	0,0195	2,0			
protein exordium like 2	Q9FE06	SC_00037132.1.1	6	1	0,0473	4,2			
protein scar4	Q5XPJ6	SC_CHIB_VITVI.2.3	122	41	0,0399	3,0			
protein spiral1-like2	Q9LE54	SC_00000089.1.1	85	39	0,0381	2,2	*42*	modifications du cytosquelette	structuration cellulaire
protein spiral1-like2	Q9LE54	SC_00003107.1.1	25	10	0,0178	2,5			
protein spiral1-like2	Q9LE54	SC_00089066.1.1	46	20	0,0380	2,3			
protein spiral1-like2	Q9LE54	SC_00134731.1.1	33	11	0,0206	3,1			
protein spiral1-like2	Q9LE54	SC_00150816.1.1	61	25	0,0330	2,4			
protein spiral1-like2	Q9LE54	SC_OS03G0417800_ORYSJ.1.6	38	19	0,0217	2,0			
protein spiral1-like2	Q9LE54	SC_OS03G0417800_ORYSJ.2.6	9	4	0,0482	2,1			
pyridoxal phosphate -dependent transferase-like protein	Q9FE98	SC_00009805.1.1	2	1	0,0405	2,0	*4*	privation du phosphate	autres
pyridoxal phosphate phosphatase-related protein	Q67YC0	SC_00008655.1.1	5	1	0,0403	4,9			
spx domain-containing protein 4	Q94A21	SC_00016812.1.1	4	1	0,0132	3,0		privation du phosphate	autres
strictosidine synthase-like 4 protein	Q9SD07	SC_00029068.1.1	20	6	0,0007	3,3		privation du phosphate	autres
12s seed storage protein cru4	Q9SIA7	SC_00008432.1.1	16	6	0,0069	2,5	*9*	réserve en acides aminés et en azote	autres
12s seed storage protein cru4	Q9SIA7	SC_00011399.1.1	3	1	0,0022	2,3			
galactinol--sucrose galactosyltransferase 2	Q84VX0	SC_RFS1_ARATH.5.8	5	2	0,0056	2,4	*4*	synthèse du raffinose	autres
galactinol--sucrose galactosyltransferase 6	Q94A08	SC_RFS2_ARATH.16.16	3	1	0,0111	2,6			
plant glycogenin-like starch initiation protein 1	O22693	SC_00009357.1.1	6	2	0,0359	3,2		synthèse du raffinose	autres

Table IV.2. Protéines identifiées surabondantes dans les ES après 1 semaine de maturation sur milieu à forte teneur en gélifiant (**9G**).

Spot	Contigs	p-Value	Ratio 9/4	Accession	Noms	Fonctions	Catégories
9	SC_UMC2374_MAIZE.1.3	0,022	1,18	Q69NG5	probable 6-phosphogluconolactonase 4, chloroplastic	réponse cellulaire à l'état rédox	protection cellulaire
15	SC_SODF_NICPL.4.4	0,05	1,31	Q6QJL0	superoxide dismutase (EC 1.15.1.1)	détoxication	
21	SC_SODM_NICPL.1.5	0,01	1,23	A9NNA1	superoxide dismutase (EC 1.15.1.1)	détoxication	
42	SC_SAHH_NICSY.1.26	0,022	1,26	B8LLL7	adenosylhomocysteinase (EC 3.3.1.1)	modification des protéines	recyclage/synthèse/modification des protéines
26	SC_00151055.1.1	0,023	1,14	P42742	proteasome subunit beta type-1	protéolyse	
27	SC_00151055.1.1	0,015	1,18	P42742	proteasome subunit beta type-1	protéolyse	
71	SC_PDI14_ARATH.1.13	0,037	1,17	A7KY78	protein disulfide isomerase	réparation des protéines	
72	SC_PDI14_ARATH.1.13	0,03	1,16	A7KY78	protein disulfide isomerase	réparation des protéines	
57	SC_PDI23_ORYSJ.1.5	0,043	1,26	A9CPA7	protein disulfide isomerase family	réparation des protéines	
6	SC_00003988.1.1	0,038	1,27	Q9LUG8	endo-1,3-1,4-beta-D-glucanase	signalisation cellulaire	signal/transduction/régulation
70	SC_CALR_RICCO.2.19	0,001	1,47	Q9FYV2	calreticulin	transduction du signal	
69	SC_CALR_RICCO.2.19	0,009	1,34	Q9FYV2	calreticulin	transduction du signal	
17	SC_00004509.1.1	0,004	1,25	Q39626	expansin S2	modification de la paroi cellulaire	structuration cellulaire
1	SC_PROF1_PHAVU.6.6	0,007	1,26	A9NMR7	profilin	Modifications du cytosquelette	
81	SC_XYL1_ARATH.1.6	0,045	1,42	Q8VWV9	putative alpha-xylosidase	modification de la paroi cellulaire	
16	SC_EXB14_ORYSJ.1.2	0,025	1,24	Q6H677	putative expansin-B14	modification de la paroi cellulaire	
66	SC_TBB2_ELEIN.1.4	0,05	1,38	Q9ZPN9	tubulin beta-2 chain	modification du cytosquelette	
76	SC_DPE2_ARATH.1.1	0,025	1,13	Q69Q02	4-alpha-glucanotransferase DPE2 (EC 2.4.1.25)		synthèse amidon
4	SC_GLGL1_BETVU.1.2	0,039	1,31	P55233	Glucose-1-phosphate adenylyltransferase chloroplastic/amyloplastic (EC 2.7.7.27)		

Des transcrits intervenant dans le remodelage de la chromatine sont surexprimés sur **9G** : deux histones H4, une histone H2a et une histone H3 présentent un FPKM moyen de 8 (SC_H2A6_ORYSJ.1.9, SC_H33_ARATH.29.70, SC_00079928.1.1, SC_H4_OLILU.1.5, **Table IV.1**). Ces transcrits ont été retrouvés à des stades précoces de l'ES du soja (Thibaud-Nissen et al., 2003) et de la luzerne (Kapros et al., 1992). Des transcrits, codant des protéines histones responsables de la mise en place des nucléosomes, et des protéines de type méthionine synthase, permettant la méthylation de l'ADN, peuvent entrainer le remaniement de la chromatine qui a lieu au cours de la mise en place d'un nouveau programme génétique.

Dans notre étude ce nouveau programme, déclenché par les nouvelles conditions de culture, aboutit au développement embryonnaire, par l'intermédiaire de protéines impliquées dans les voies de transduction du signal et par des facteurs de transcription spécifiques (**Tables IV.1 et IV.2**). Parmi les transcrits, 2 putatives wuschel homeobox protein présentent un FPKM moyen de 20 (SC_00005814.1.1 et SC_00011070.1.1) ; ce sont des facteurs de transcription impliqués dans le développement précoce des embryons (Cairney, et al., 2006 ; Palovaara et Hakman, 2008, chapitre I, partie III.1.4.2). Chez *Arabidopsis*, le gène WUSCHEL codant une protéine impliquée dans la mise en place des méristèmes apicaux caulinaires, qui pourrait également jouer un rôle clé au cours de l'ES en permettant la différenciation cellulaire (Zuo et al., 2002). Ensuite 2 ethylene-responsive element binding factor-like protein présentant un FPKM moyen de 5, sont des facteurs de transcription impliqués dans les voies de réponse aux stress (SC_EF101_ARATH.1.1 et SC_00157909.1.1) Chez le coton, un gène codant un élément de réponse à l'éthylène est exprimé au cours des phases précoces de l'ES (Zeng et al., 2006). Chez le pin sylvestre, l'activité d'une enzyme limitant la production d'éthylène (1-aminocyclopropane-1-carboxylate synthase, ACS) marque l'embryogenèse précoce de l'ES (Lu et al., 2011). Des transcrits impliqués dans la signalisation cellulaire, comme ceux codant des calmodulines, protéines fixatrices du calcium présentent un FPKM moyen de 16 (SC_00024580.1.1, SC_POLC2_JUNOX.2.31, SC_00043177.1.1, SC_CML10_ORYSJ.2.26, SC_POLC2_JUNOX.18.31). Deux calmodulines ont aussi été identifiées dans l'étude du protéome des ES précoces de *Picea glauca* (Lippert et al., 2005) ; leur rôle est de moduler la teneur en Ca^{2+} intracellulaire qui est un messager secondaire important pour la mise ne place de l'embryogenèse somatique (Overvoorde et Grimes, 1994; Anil et Rao, 2000). Des transcrits codant des protéines extracellulaires, dont 5 chitinases (FPKM moyen de 20, SC_00008805.1.1, SC_00025909.1.1, SC_CHI1_ORYSJ.4.9, SC_CHI4_ORYSJ.1.7, SC_CHI8_ORYSJ.2.2 FPKM), sont

surexprimés sur **9G**. Chez la chicorée, l'accumulation de protéines extracellulaires joue un rôle dans l'induction de l'ES (Hilbert et al., 1992; Boyer et al., 1993). Chez la carotte, les chitinases EP3-1 et EP3-2 sont capables de restaurer le développement des ES d'une lignée cellulaire produisant pas d'embryon (van Hengel et al., 1998, voir chapitre I, partie III.1.5.1).

Le nouveau programme génétique se traduit par d'intenses modifications cellulaires et métaboliques au sein des embryons somatiques. Les modifications du cycle cellulaire, tout comme les divisions et les expansions cellulaires, sont accompagnées de modifications du cytosquelette et de la paroi cellulaire ainsi que des membranes lipidiques. Une des principales modifications structurales observée au cours du développement embryonnaire, concerne l'altération des propriétés physico-chimiques des parois cellulaires. Ces changements peuvent être induits par des modifications au niveau des enzymes impliquées dans la biosynthèse des monolignols ou dans leur assemblage dans la paroi (Dixon et Paiva, 1995). Des transcrits codant pour 6 inositol oxygenase (SC_MIOX_ORYSJ.3.8, SC_00014095.1.1, SC_00026140.1.1, SC_00040126.1.1, SC_00045575.1.1, SC_MIOX_ORYSJ.1.8) présentant un FPKM moyen de 49, sont impliqués dans la biosynthèse des précurseurs des nucléotide-sucre de la matrice de polysaccharides de la paroi cellulaire (Aquea et Arce-Johnson, 2008). Trois non-specific lipid-transfer protein, (FPKM moyen de 11, SC_NLTP_PINTA.219.283, SC_NLTP_PINTA.5.283, SC_NLTP2_LENCU.9.30), et 2 lipid transfer protein, (FPKM moyen de 5, SC_00011326.1.1, SC_00154588.1.1) jouent un rôle dans les modifications membranaires par le transfert de phospholipides ainsi que de galactolipides à travers les membranes. Les protéines de transfert des lipides sont aussi fortement exprimées dans les stades précoces de l'embryogenèse (Sterk et al., 1991; Zeng et al., 2006, chapitre I, partie III.1.5.2). Cinq phospholipase C (SC_LOC100284361_MAIZE.1.1, SC_00034320.1.1, SC_00122309.1.1, SC_OS03G0852800_ORYSJ.1.1) présentent un FPKM moyen de 11. Ces enzymes interviennent dans le catabolisme des lipides et sont capables d'hydrolyser les arabinogalactane protéine (AGPs) pour les libérer dans la matrice extracellulaire, permettant leur sécrétion (Johnson et al., 2003). Les AGPs sont impliqués dans la signalisation cellulaire de l'embryogenèse (Kreuger et Holst, 1993; von Arnold et al., 2005, voir chapitre III.1.5.1).

L'embryogenèse se traduit par une augmentation rapide de la teneur en protéines solubles endogènes, ce qui indique un métabolisme cellulaire très actif et «nouvellement orienté» (Blanc et al., 2002). L'accumulation de protéines a lieu graduellement au cours du développement embryonnaire, pour devenir constant en fin de maturation chez *Picea mariana*

et *Picea glauca* (Iraqi et Tremblay, 2001a). Des transcrits participant à la synthèse protéique et en particulier cinq 26S ribosomal RNA (FPKM moyen de 25, SC_LOC100275739_MAIZE.44.160, SC_LOC100275739_MAIZE.49.160, SC_LOC100275739_MAIZE.55.160, SC_LOC100275739_MAIZE.63.160, SC_LOC100275739_MAIZE.95.160), deux elongation factor 1-alpha et 2 elongation factor 1-delta (FPKM moyen de 6, SC_00000762.1.1, SC_00149373.1.1, SC_00027980.1.1, SC_EF1D_PIMBR.17.70) ont été mis en évidence (**Table IV.1)**. La synthèse des protéines est indispensable à la transition de l'état indifférencié à l'état différencié des ES (action sur l'organogenèse). Le gène CEM1 codant le facteur d'élongation 1-alpha est surexprimé dans les cellules se divisant activement au stade globulaire de l'ES de carotte (Kawahara et al., 1992).

Concernant la dégradation des protéines, deux proteasome subunit beta type ont été identifiés en protéomique (spot 26 et 27, **Table IV.2**). De nombreuses protéines présentant des similarités avec des protéases ont été isolées au cours du développement embryonnaire. Par exemple chez le colza (*Brassica napus*) la surexpression d'une protéase à cystéine a été mise en évidence (Stroeher et al., 1997) ; son rôle est d'éliminer les protéines redondantes, les acides aminés issus de cette dégradation peuvent alors être réutilisés pour la synthèse de nouvelles protéines (Alamillo et Bartels, 1996; Bauer et al., 1997).

La réponse des ES indifférencié aux conditions de culture est visible par un nombre conséquent de transcrits identifiés, impliqués dans la protection cellulaire et dans la réponse au stress, en particulier au stress hydrique. De plus cette réponse adaptative aux conditions culturales se traduit par la mise en place d'un métabolisme secondaire intense (**Tables IV.1 et IV.2**). Parmi ces transcrits beaucoup sont impliqués dans la réponse à un déficit en eau, qui peut être directement relié à la plus faible disponibilité en eau du milieu **9G**. Concernant la protection cellulaire, des transcrits codant des protéines de réponse au stress, et en particulier au stress hydrique, comme une dehydration-induced protein erd15 (SC_00000544.1.1, FPKM moyen de 7), deux dehydrin 2 (SC_00081475.1.1 et SC_00001231.1.1, FPKM moyen de 56) et trois water deficit inducible protein LP3 (SC_00144857.1.1, SC_ASR2_SOLLC.2.31, SC_ASR2_SOLLC.14.31, FPKM moyen de 35) ont été isolés. Les ES en différenciation semblent répondre très vite (dès 1 semaine) au changement d'environnement hydrique occasionné par le changement de milieu. Quatre aquaporin pip2-1 impliquées dans les transporteurs d'ions (SC_00001064.1.1, SC_00002754.1.1, SC_00015040.1.1, SC_TIP11_ARATH.2.14, FPKM moyen de 6) sont surexprimées dans les ES cultivés sur **9G**.

Les canaux hydriques sont importants dans la régulation du potentiel osmotique de la cellule, *via* la régulation du flux hydrique. Chez *Atriplex canescens* un transcrit surexprimé au cours d'un stress hydrique et codant une aquaporine a été isolé (Cairney et al., 1995). Cependant les aquaporines peuvent aussi marquer l'élongation du suspenseur au cours du développement précoce. En effet, l'expression d'une aquaglycéroporine localisée dans le suspenseur (Ciavatta et al., 2002), augmente en début de maturation de l'ES chez *Pinus taeda*, indiquant une très rapide différenciation des cellules du suspenseur. Cette expression est augmentée à la fois lors de la différenciation précoce des embryons et lors du développement tardif des embryons, en concordance avec l'augmentation de la taille du suspenseur.

Enfin, nous avons identifié des transcrits codant pour des enzymes impliquées dans la réponse au stress oxydatif : 3 peroxydases (SC_PER12_ARATH.2.14, SC_00055627.1.1 et SC_00077646.1.1, FPKM moyen à 14) et deux glutathione S-transferase (SC_00026575.1.1, SC_UMC1467_MAIZE.1.1, FPKM moyen à 17) sont surexprimées en réponse aux variations des conditions de culture. Les activités peroxydase et glutathione sont fortement impliquées dans les réactions de défense des cellules contre les espèces réactives à l'oxygène (ROS). L'implication de ces enzymes dans le processus morphogénétique et dans la prévention de la sénescence a également été rapportée (van Engelen et de Vries, 1992). La présence de peroxydases dans le milieu de culture est nécessaire pour maintenir la taille et la forme des cellules du protoderme chez les jeunes embryons somatiques de carotte (Cordewener et al., 1991). Ces enzymes semblent garder la paroi cellulaire rigide (Cella et Carbonera, 1997). L'équilibre rédox des tissus embryonnaires serait un des facteurs importants régissant l'initiation et la différenciation de l'embryogenèse somatique (Thibaud-Nissen et al., 2003; Belmonte et al., 2005). Pour certains auteurs, un certain seuil de stress oxydant est nécessaire pour favoriser la formation des cellules embryogènes (Kairong et al., 2002; Imin et al., 2005; Pasternak et al., 2005). Dans ce contexte, l'évolution des systèmes enzymatiques antioxydants pourraient servir de signal requis pour l'acquisition de la compétence embryonnaire (Blazquez et al., 2009). En outre, une augmentation du nombre de publications ont lié les espèces réactives de l'oxygène (ROS) à l'embryogenèse somatique (Zavattieri et al., 2010; Rodriguez-Serrano et al., 2012; Gomez-Garay et al., 2013).

En conclusion, l'intégration de nos données obtenues à plusieurs nivaux d'étude nous a permis d'établir des relations entre ces différents niveaux et de proposer un modèle du contrôle physiologique et moléculaire de l'embryogenèse somatique chez le pin maritime

(**Figure IV.2**). La teneur en gélifiant du milieu contrôle les conditions biochimiques et physiologiques de l'environnement de maturation des ES qui répondent différemment selon que l'environnement est limitant ou non à l'accès à l'eau. La réponse sur milieu **4G**, potentiellement peu limitatif en eau, passe d'abord par une réponse moléculaire d'activation des voies de la glycolyse, de la néoglycogenèse, de l'anoxie et du stress oxydatif, et une inhibition des processus de mort cellulaire programmée. Cela entraine une réponse cellulaire caractérisée par l'augmentation de la réplication de l'ADN, des divisions cellulaires, de la respiration, de la mise en place de phénomènes actifs de polyembryogenèse de clivage, et par une réponse cellulaire aux stress abiotiques subits. Cette réponse cellulaire se traduit au niveau biologique par le blocage du développement embryonnaire et la prolifération entrainant des évènements de sénescence marqués. A l'opposé, la réponse des ES sur milieu **9G**, potentiellement très limitatif en eau, passe d'abord par une réponse moléculaire d'activation des voies de recyclage et de resynthèse des protéines, de la mort cellulaire programmée, et du métabolisme secondaire. Les cellules embryonnaires subissent une restructuration marquée au niveau de leur paroi cellulaire et de leur membrane, ou meurent par PCD pour permettre la différenciation embryonnaire. La différenciation leur permet aussi d'exprimer de nouveaux types de gènes, comme de métaboliser des composés secondaires pour se protéger contre les stress abiotiques imposés par l'environnement de maturation. Cette réponse cellulaire permet au niveau biologique la maturation des embryons indifférenciés.

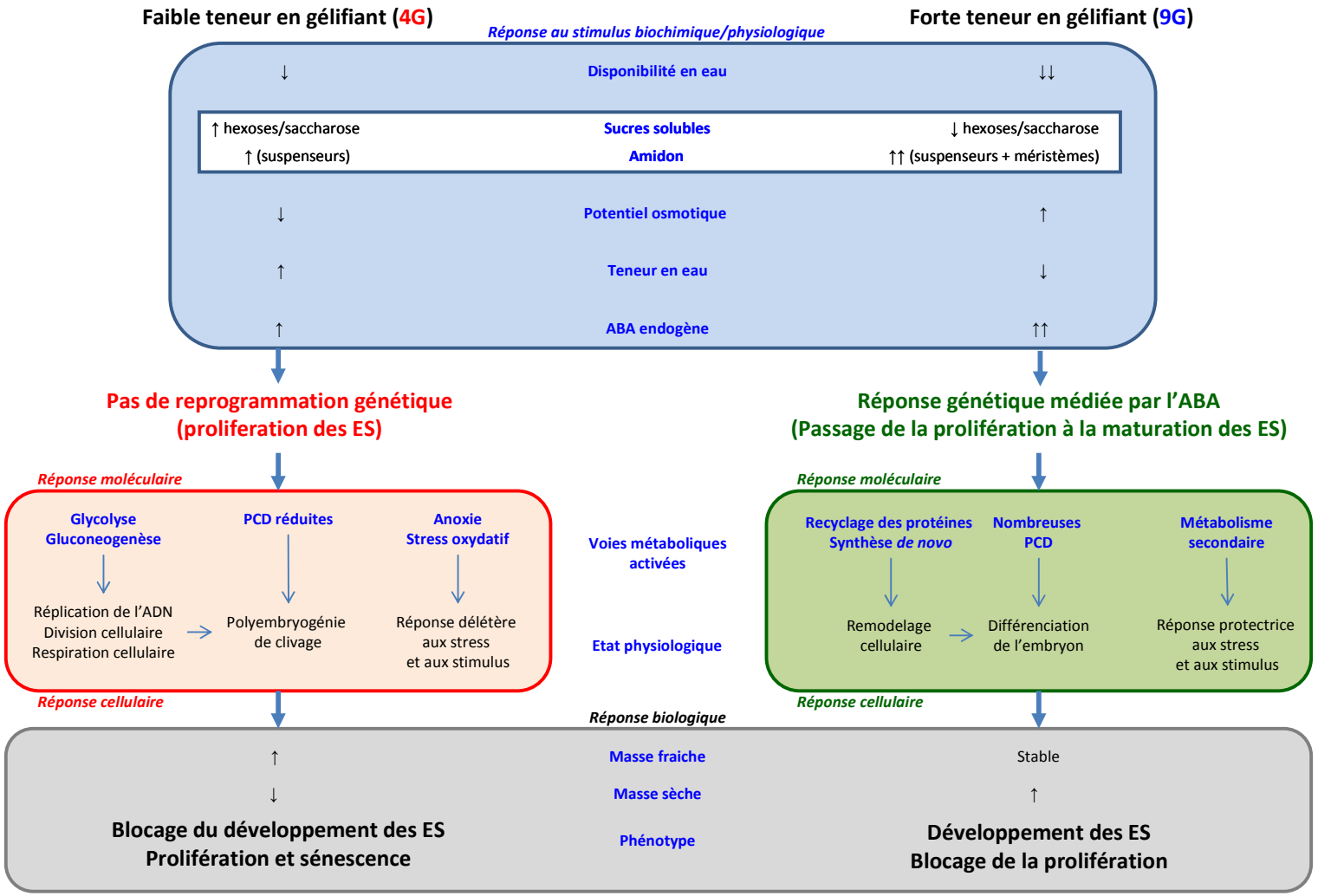


Figure IV.2 Proposition de modélisation des interactions physiologiques, cellulaires et moléculaires entre les processus impliqués dans l'embryogenèse somatique précoce du pin maritime en fonction de la teneur en gélifiant du milieu de culture. **PCD :** mort cellulaire programmée ; **↑ :** augmentation ; **↓ :** diminution.

IV.4 Comparaison du développement de l'EZ et de son mégagamétophyte au cours de la maturation de la graine de pin maritime

La représentation de l'analyse en composantes principales (ACP) des différentes mesures réalisées sur l'EZ tout au long de sa maturation (chapitre III.3.1), nous permet de regrouper les EZ analysés en deux groupes : le stade cotylédonaire frais, et le stade cotylédonaire déshydraté. Une transition très nette est visible entre août et septembre, où la teneur en eau commence à chuter et où la composition en sucres se modifie drastiquement. Les ratios hexose : saccharose et RFO : saccharose s'inversent à partir de ce point clé de septembre (Morel et al, en préparation, chapitre III.2, **Table 3**). Fait intéressant, l'EZ de septembre (EZ_4) se situe à la transition entre les 2 groupes ; même s'il reste plus proche du groupe déshydraté, il n'accumule pas encore de stachyose alors que le raffinose est déjà présent. La TE de l'EZ ne semble plus évoluer à partir de fin septembre, par contre la composition en sucres, en protéines totales et en protéines de réserve continue de varier jusqu'à décembre. Il aurait été intéressant de collecter des EZ entre début août et fin septembre, où des modifications métaboliques semblent se mettre en place très rapidement.

Globalement au niveau MS, TE et teneur en sucres solubles, le mégagamétophyte semble subir moins de modifications que l'EZ. Le tissu nutritif semble agir comme une zone tampon protègeant l'embryon. La diminution de TE du mégagamétophyte aux derniers stades de développement (novembre et décembre) est corrélée avec l'apparition du stachyose, connu pour son rôle dans l'acquisition de la tolérance à la déshydratation (Obendorf, 1997; Muzquiz et al., 1999) (**Figures III.3** et **III.4**). L'accumulation de raffinose et de stachyose par l'EZ est aussi corrélée avec la perte d'eau aux mêmes stades tardifs de maturation (Morel et al, en préparation, chapitre III.2). Le mégagamétophyte est un tissu nutritif plus riche en MS que l'EZ (le rapport $MS_{EZ}/MS_{Méga}$ est de 19 fin juillet et diminue jusqu'à 5 en décembre), laissant supposer une réserve importante notamment en sucres pouvant alimenter l'EZ. Par exemple, la teneur en raffinose pour le mégagamétophyte maximale au 8 août diminue fin septembre, en même temps que ce sucre apparait dans l'EZ. Ce même type d'observation est valable pour la quantité de protéines totales qui commence à diminuer dans le mégagamétophyte alors qu'elle augmente dans l'EZ (octobre), ou encore la diminution de la quantité de protéines de réserve (bandes 2 et 3 pour le mégagamétophyte, **Figure III.6**) en même temps qu'elle augmente pour l'EZ, à cette même date (octobre). Ces données illustrent le rôle de support

nutritif du mégagamétophyte pour l'embryon en maturation, en termes d'apports en sucres solubles et en protéines de réserve. Les réserves du mégagamétophyte disponibles pour l'EZ ne sont *de facto* indisponibles pour l'ES *in vitro*. Cela peut expliquer les différences de vigueur observée entre les ES et les EZ au niveau de leur germination et de leur développement en plantes ES (Trontin et al 2013). Enfin la composition lipidique de l'EZ et du mégagamétophyte pourrait être importante à analyser au cours de la maturation, ces autres composés de réserve du mégagamétophyte pouvant supporter le développement de l'EZ (Silveira et al., 2004; Terskikh et al., 2005; Brownfield et al., 2007; Pescador et al., 2012).

CHAPITRE V – Conclusions et Perspectives

V.1 Conclusions

Ce travail de thèse portant sur l'embryogenèse somatique du pin maritime avait deux objectifs majeurs: **1-** le premier s'intéressait aux mécanismes physiologiques, cellulaires et moléculaires précoces contrôlant la différenciation des ES en réponse à une disponibilité en eau réduite **2-** le second consistait à évaluer l'état de maturité physiologique et la composition moléculaire (protéome) de l'ES pour ensuite le comparer à l'EZ. Afin de répondre à ces objectifs, nous avons utilisé des méthodologies capables d'analyser l'expression géniques de notre matériel végétal. L'extraction, le séquençage et l'annotation des pools de transcrits et de protéines par des techniques novatrices de transcriptomique et de protéomique nous a permis de caractériser notre matériel végétal au niveau moléculaire.

1- Pour atteindre notre premier objectif, nous avons fait le rapprochement des données transcriptomique et protéomique permettant l'analyse de deux niveaux de régulation de l'embryogenèse somatique du pin maritime. Chez les conifères cette **approche est originale car jamais entreprise à ce jour**; elle est d'autant plus innovante qu'elle a été complétée par des observations histologiques, biologiques et biochimiques, nous permettant d'obtenir une **image globale des processus impliqués** lors des premières phases de l'embryogenèse somatique. Nous avons montré au niveau moléculaire que l''augmentation de la teneur en gélifiant du milieu de maturation diminue l'eau disponible pour les embryons en culture et oriente le métabolisme des ES indifférenciés vers l'embryogenèse précoce. Des marqueurs précoces du développement (transcrits et protéines) isolés au cours de ces travaux associés aux paramètres physiologiques et biochimiques mesurés nous permettent de dresser un certain nombre de conclusions. En présence d'une **forte concentration en gélifiant** les faits suivants sont observés :

- Une forte augmentation de l'ABA endogène à la quatrième semaine de maturation suggère un signal pour le développement ontogénétique des ES. La forte augmentation de l'ABA endogène à la quatrième semaine de maturation n'est pas *à priori* qu'une réponse à un stress mais suggère un signal pour le développement ontogénétique des ES
- Deux marqueurs clés du développement embryonnaire (SRK2 et PPC2) surexprimés dans les ES en développement, associés à l'augmentation de la teneur en ABA endogène, suggèrent

leur réponse adaptative aux conditions de disponibilité réduite en eau. Cette réponse moléculaire serait induite par un pic d'ABA endogène.

- Plusieurs marqueurs impliqués dans la biosynthèse de l'amidon suggèrent que son accumulation, constituant principal des réserves carbonées, est un processus important pour l'embryogenèse précoce de l'ES.
- Des marqueurs potentiellement impliqués dans la réorganisation de la chromatine, comme la protéine SUMO ubiquitine, suggèrent la réorganisation intensive de l'expression des gènes vers le développement des ES.
- Plusieurs marqueurs protéiques (protéosome/ubiquitine, chitinase, expansine, disulfure isomérase, tubuline), ainsi que nos observations histologiques, semblent indiquer la transition de la division cellulaire vers le remodelage cellulaire favorable à l'embryogenèse.
- Plusieurs marqueurs protéiques identifiés sont potentiellement impliqués dans la détoxication et la réponse au stress oxydatif (superoxyde dismutase, germin-like protein). Cette capacité des embryons à développer un système d'homéostasie rédox efficace pourrait être cruciale pour réguler et maintenir leur différenciation cellulaire.
- Plusieurs marqueurs sont impliqués dans le métabolisme secondaire et la biosynthèse de molécules protectrices comme les flavonoïdes (Flavonols F3H, Naringerin). Les ES pourraient surmonter la contrainte d'une faible disponibilité en eau, imposée par le milieu de culture, grâce à la synthèse de métabolites secondaires de protection.

Par opposition, en présence d'**une faible teneur en gélifiant** du milieu de maturation, défavorable à la différenciation des ES, ne limite pas la disponibilité en eau et oriente le métabolisme des ES indifférenciés vers la prolifération, processus antagoniste à l'embryogenèse somatique. Les faits suivants corroborent ces conclusions :

- Plusieurs marqueurs protéiques associés à l'anoxie (ADH, PDC, sucrose synthase), le stress osmotique (enolase bifonctionnelle) et le métabolisme du méthane suggèrent que l'environnement de maturation est stressant et défavorable à la différenciation et la maturation des ES.
- L'environnement de maturation entraîne l'émergence rapide de la fermentation alcoolique (ADH) et de la sénescence cellulaire (26S protéasome) confirmée par ailleurs par nos observations cytologiques.

2- Pour atteindre notre deuxième objectif, l'état de maturité de l'ES cotylédonaire a été caractérisé par des analyses biologiques (matière sèche, teneur en eau), biochimiques (réserves glucidiques et protéiques) et protéomiques:

- Cet état de maturité ne varie pas en fonction de la durée de maturation, (10 à 14 semaines). La récolte en routine des ES après 12 semaines de maturation apparait donc fondée sur la base des données biologiques et biochimiques.
- Une **approche innovante** a été utilisée pour définir quel stade de développement zygotique se rapprochait le plus de l'ES de 12 semaines. Une méthode de classification hiérarchique ascendante (HCA), basée sur **neuf variables explicatives** mesurées, nous a permis de démontrer que l'ES cotylédonaire de 12 semaines se rapprochait de l'état de maturité d'un EZ cotylédonaire frais. Ces embryons frais n'avaient pas atteint le niveau de maturité des EZ cotylédonaires de graines déshydratées dans nos conditions expérimentales.
- Les données **quantitatives de protéomique globale**, ont révélé des profils de protéines similaires **entre les ES et EZ frais cotylédonaires**. Ces protéines étaient principalement liées à la réponse au stress, au métabolisme énergétique et aux réserves nutritives. Certaines protéines de réserve (de type légumine) sont apparemment moins accumulées dans les ES contrairement aux EZ, suggérant que l'EZ, contrairement à l'ES, est déjà entré dans une phase de déshydratation. Toutefois, un certain nombre de marqueurs protéiques (oméga-hydroxypalmitate O-féruloyl transférase, la protéine embryonnaire DC-8) suggèrent que les ES se préparent à la déshydratation. Un certain nombre de marqueurs protéiques impliqués dans les voies de la glycolyse (malonate-CoA ligase ; glucose-1-phosphate adénylyltransférase ; NPD -L- rhamnose synthase et l'ADH), suggèrent un métabolisme énergétique plus actif chez l'ES par rapport à l'EZ. Ce type de métabolisme pourrait expliquer pourquoi les ES matures de 12 semaines sont capables d'entrer en germination sans subir le programme la déshydratation, comme observé pour l'EZ.
- Enfin, des **marqueurs protéiques génériques du stade cotylédonaire frais** lors de l'embryogenèse somatique et zygotique du pin maritime ont été mis en évidence pour la première fois au cours de ce travail. En particulier l’apparition simultanée de protéines de choc thermique, de protéines abondantes de l'embryogenèse tardive et de diverses protéines de réserve a été mise en évidence. D'autres marqueurs du stade cotylédonaire frais n'avaient jamais été décrits chez les conifères : par exemple, l'aldose réductase, impliquée dans la régulation du potentiel osmotique de l'embryon au cours de la déshydratation, ou encore la 6-phosphogluconate déshydrogénase, protégeant la cellule contre le stress oxydatif. Enfin, l'adénosine kinase 2 est impliquée dans l'élargissement du pool de nucléotides lors des phases de croissance intensive.

V.2 Perspectives

Les techniques de génomique mises en œuvre au cours de cette thèse ont permis la constitution de jeux de données conséquents, très riches en informations, qui n'ont pas forcément toutes été intégrées et discutées dans cet ouvrage. Nous touchons un peu du doigt la limite actuelle de nombreuses approches moléculaires, qui produisent plus de données que l'on ne peut en exploiter. Cependant, ces données pourront être reprises ultérieurement à la lumière de nouvelles découvertes et restent accessibles en ligne via les bases de données. En effet, les contigs assemblés au cours de l'analyse transcritomique ont été mis à disposition publique dans les archives du NCBI sous le numéro d'accession SRR609713. Nous avons utilisé cette base de données pour annoter les séquences protéiques obtenues par spectrométrie de masse, pour toutes les analyses protéomiques de ce travail de thèse. Cette base de donnée a surtout été utilisée pour réaliser l'assemblage *de novo* du transcriptome du pin maritime, réalisé en collaboration avec toutes les équipes de recherche européennes travaillant sur cette espèce (France, Espagne, Portugal, Canales et al., 2013). Ce travail fondateur pour le pin maritime a fourni un large catalogue de gènes exprimés (27541 transcrits uniques et 9148 ADNc complets) et l'identification de marqueurs polymorphes SSR et SNP robustes pour les applications de génotypage dans les programmes de sélection, mais aussi pour faciliter le séquençage du génome, et les études de génomique fonctionnelle appliquées au pin maritime.

Enfin, en dépit de l'existence de catalogues complets de gènes et de séquences codant pour des protéines de conifères, dont nous nous sommes servi pour comparer toutes les séquences déterminées au cours de ce travail, un grand nombre de transcrits et de protéines n'ont pas été annotés fonctionnellement. Dans un futur proche, lorsque le génome du pin maritime aura entièrement été séquencé et annoté, nous pourrons réhabiliter ces séquences et leur attribuer un rôle fonctionnel, ce qui confèrerait une plus-value importante à ce travail.

L'ensemble des données moléculaires générées au cours de ce projet de thèse nous fournissent énormément de candidats potentiellement marqueurs du développement des ES, que nous pouvons désormais caractériser individuellement. Pour cela, des techniques analytiques récentes existent et ont déjà été mises au point pour suivre spécifiquement certains marqueurs des ES de conifères, par exemple le suivi de l'expression de gènes par qRT-PCR au sein des ES et EZ de pin maritime (de Vega-Bartol et al., 2013a; de Vega-Bartol et al., 2013b). L'objectif serait alors de suivre la dynamique d'expression de marqueurs génétiques

du développement embryonnaire sélectionnés parmi ceux que nous avons isolés par nos approches globales.

Dans notre étude de la régulation de l'embryogenèse précoce des ES, nous avons isolé les ARNm exprimés par les ES. Ceux-ci constituent le niveau de régulation transcriptionnelle de l'embryogenèse, cependant, il existe d'autres niveaux de régulation. Par exemple, les miARN (microARN) sont de petits ARN non-codants impliqués dans la régulation post-transcriptionnelle des gènes (He et Hannon, 2004; Jones-Rhoades et al., 2006). L'action des miARN permet l'élimination d'ARNm cibles et donc la répression de certains gènes. Ils peuvent donc être étudiés au travers de leur propre expression (par RT-PCR) comme de simples transcrits mais plus petits, ou au travers des séquences ARN qu'ils régulent de par leurs motifs spécifiques, par l'étude du dégradome. Ainsi, les miARN sont impliqués dans la régulation de l'embryogenèse. En effet, dans les plantes porteuses de mutations dans les gènes codant des miARN, l'embryogenèse et le développement de méristèmes apparaissent défectueux (Schauer et al., 2002; Bartel, 2004). Chez les gymnospermes, et en particulier chez des espèces dépourvues d'un génome séquencé comme le pin maritime, les informations sur les miARN demeurent limitées. Quelques études sur l'embryogenèse des conifères commencent à s'y intéresser, au traversnotamment de l'analyse de l'expression de plusieurs miARN par qRT-PCR chez le mélèze (Zhang et al., 2012) ou encore chez *Pinus taeda* (Oh et al., 2008). Une autre étude chez le mélèze a montré qu'un miARN (miR171) régule spécifiquement le gène *La*SCL6 dans les cellules embryogènes, suggérant que sa répression permettrait le maintien du potentiel embryogène (Li et al., 2014). Concernant le pin maritime, une analyse récente par puce à ADN du développement de l'EZ a montré que des transcrits codant des petits ARN (sRNA) sont apparus différentiellement régulés à tous les stades de développement de l'EZ avec une prévalence des transcrits codant des miARN du milieu à la fin de l'embryogenèse (de Vega-Bartol et al., 2013b). Cette dynamique d'expression des transcrits (ARNm ou miARN) au cours du développement embryonnaire, peut constituer un autre moyen pour définir les stades de développement de l'ES chez le pin maritime.

Une autre forme d'étude moléculaire de plus en plus répandue, est la métabolomique qui permet de dresser des profils qualitatifs et quantitatifs de tous les métabolites cellulaires (acides aminés libres et protéolytiques, sucres non structuraux, acides organiques, amines et polyamines, phytohormones, acides gras, composés organiques volatils (COV), alcools, polyols, composés phénoliques et polyphénoliques, alcaloïdes, terpènes....). Ces profils sont obtenus à l'aide de techniques analytiques de pointe comme la chromatographie en phase gazeuse couplée à la spectrométrie de masse (GC/MS). Les données de métabolomique

peuvent fournir une foule de renseignements sur l'état biochimique des tissus ; l'interprétation de ces données offre une approche efficace pour la caractérisation fonctionnelle des gènes, et pour la description et l'élucidation des réponses physiologiques des cellules aux conditions environnementales (Roessner et al., 2001a; Roessner et al., 2001b; Fiehn, 2002; Sumner et al., 2003; Weckwerth, 2003). Des études du métabolome de l'embryogenèse chez les conifères ont déjà été entreprises chez *Pinus taeda* (Robinson et al., 2009) ou chez *Picea abies* (Businge et al., 2012).

Un autre type d'approche serait de suivre l'activité enzymatique de certains marqueurs à propriété enzymatique, comme l'adénosine kinase, ou l'aldose réductase que nous avons isolés (Morel et al, en préparation, chapitre III.2). Cela nous renseignerait sur l'état de maturité atteinte par l'ES. C'est une façon destructrice de connaitre l'état de maturité de l'ES à un stade donné, mais qui a fait ses preuves. Par exemple, les activités de trois enzymes, l'adénine phosphoribosyltransferase (APRT), l'adénosine kinase (AK), et l'inosine kinase (IK) associées au recyclage des composés purines ont été mesurées durant le dévelopement de l'ES de *Picea glauca* et augmentent avec le traitement de déshydratation des ES (Stasolla et al., 2001).

Des nouvelles approches sont donc désormais disponibles pour isoler d'autres candidats moléculaires, susceptibles d'apporter des connaissances fondamentales complémentaires à ce travail, aussi bien pour étudier l'embryogenèse précoce, que pour travailler sur les dernières étapes de la maturation.

Concernant cette dernière étape de maturation de l'ES, les conclusions apportées par notre travail laissent entrevoir un certain nombre de possibilités qui pourraient être directement applicables à notre méthodologie de culture afin d'améliorer la qualité des embryons. Une perspective possible à ce travail serait d'étudier des stades d'ES plus tardifs (16, 18 voire 20 semaines). Au-delà de 14 semaines de maturation, des changements de composition moléculaire de l'ES peuvent encore s'effectuer. L'ES peut continuer à compléter sa maturation se rapprochant ainsi d'un EZ déshydraté. Cependant, les ES maintenus sur milieu de maturation peuvent au contraire rentrer en phase de germination précoce, pas forcément visible macroscopiquement mais détectable au niveau des protéines comme déjà observé chez le pin sylvestre (Lelu-Walter et al. 2008).

Beaucoup de travaux ont établi que la vigueur des graines et leur longévité potentielle continuent d'augmenter après la maturité et que les derniers stades de la maturation des graines sont importants pour maximiser leur qualité. Chez certaines espèces de conifères, une

déshydratation doit être appliquée aux ES pour pouvoir germer. Même si pour le pin maritime, elle n'est pas indispensable, elle pourrait contribuer à améliorer la qualité des ES. Des méthodes de déshydratation d'ES cotylédonaire de conifères ont été développées (Roberts et al., 1990). Les paramètres importants à contrôler sont la température (soit température ambiante 25°C, soit froide 4°C), l'hygrométrie (en général élevée, 98 % HR) et la durée de déshydratation (de 1 à 3 semaines). En jouant sur ces trois paramètres, les conditions optimales de déshydratation ont été obtenues avec succès pour de nombreuses espèces de conifères (Attree et Fowke, 1993).

Nous disposons à présent des outils moléculaires pour évaluer la qualité des ES matures ; de plus nous connaissons les marqueurs susceptibles de varier lors de la déshydratation de l'ES à savoir les protéines et les sucres en particuliers ceux de la famille des RFO (raffinose, stacchyose). Nous proposons d'appliquer une déshydratation aux ES cotylédonaires de 12 semaines, d'analyser leurs composés de réserve **(**aux niveaux quantitatif et qualitatif) et de les comparer aux réserves d'un ES non-déshydraté voire à celles d'un EZ. Cette comparaison nous permettra de vérifier l'efficacité du traitement de déshydratation et la similitude recherchée entre ES et EZ mature. En plus des protéines et des sucres d'autres composés de réserve, comme l'amidon et les lipides, pourraient être étudiés pour décrire de façon exhaustive l'ensemble des réserves accumulées par l'ES. Enfin, pour aller au bout de ce raisonnement, il conviendra ultérieurement de vérifier l'impact de la déshydratation sur la germination de l'ES mais surtout *in fine* sur la phénologie, le comportement du plant sur le terrain. En effet, au cours des premiers mois de croissance en forêt, des différences existent entre des plants de pin maritime issus d'ES et des semis (Trontin et al., 2013).

Enfin, la déshydratation des ES présente un avantage appliqué pour une production clonale. Les ES pouvant être produits en très grande quantité, en continu tout au long de l'année, leur déshydratation permettra de les conserver (stade arrêté, « quiescent »). Cela se traduit par une homogénéisation du développement de tous les ES : germination synchrone en quelques jours fournissant *in fine* des plants d'âge et de taille uniforme pour la plantation (Lelu et al., 1995; Dronne et al., 1997). Cette synchronisation du développement facilite donc toutes les manipulations ultérieures.

Bibliographie

— A —

Adams, C. A., Rinne, R. W., et Fjerstad, M. C. (1980). Starch deposition and carbohydrase activities in developing and germinating soya bean seeds. *Annals of Botany* **45**, 577-582.

Alamillo, J. M., et Bartels, D. (1996). Light and stage of development influence the expression of desiccation-induced genes in the resurrection plant *Craterostigma plantagineum*. *Plant, Cell & Environment* **19**, 300-310.

Alazard, P., Canteloup, D., Crémière, L., Daubet, A., et Lesgourgues, T. (2005). Genetic breeding of the maritime pine in Aquitaine: an exemplary success story. (PG_Edition, ed.). By Groupe Pin Maritime du Futur, Cestas, France.

Alazard, P., et Raffin, A. (2002). Les gains génétiques des premières et deuxièmes générations de vergers. *In* "Groupe pin maritime du futur eds. *Le progrès génétique en foret*", pp. 27-40. PG éditions.

Alia, R., et Martin, S. (2003). "EUFORGEN Technical Guidelines for genetic conservation and use for Maritime pine (*Pinus pinaster*)," Bioversity International.

Allona, I., Casado, R., et Aragoncillo, C. (1992). Seed storage protein from *Pinus pinaster* : homology of major components with 11S proteins from angiosperms. *Plant Science* **87**, 9-18.

Allona, I., Quinn, M., Shoop, E., Swope, K., Cyr, S. S., Carlis, J., Riedl, J., Retzel, E., Campbell, M. M., et Sederoff, R. (1998). Analysis of xylem formation in pine by cDNA sequencing. *Proceedings of the National Academy of Sciences* **95**, 9693-9698.

Alvarez, J. M., Cortizo, M., Bueno, N., Rodriguez, A., et Ordas, R. J. (2013). CLAVATA1-LIKE, a leucine-rich-repeat protein receptor kinase gene differentially expressed during adventitious caulogenesis in *Pinus pinaster* and *Pinus pinea*. *Plant Cell, Tissue and Organ Culture (PCTOC)* **112**, 331-342.

Andriotis, V. M. E., Pike, M. J., Kular, B., Rawsthorne, S., et Smith, A. M. (2010). Starch turnover in developing oilseed embryos. *New Phytologist* **187**, 791-804.

Anil, V. S., et Rao, K. S. (2000). Calcium-Mediated Signaling during Sandalwood Somatic Embryogenesis. Role for Exogenous Calcium as Second Messenger. *Plant Physiology* **123**, 1301-1312.

Aquea, F., et Arce-Johnson, P. (2008). Identification of genes expressed during early somatic embryogenesis in *Pinus radiata*. *Plant Physiology and Biochemistry* **46**, 559-568.

Attree, S. M., et Fowke, L. C. (1993). Embryogeny of gymnosperms: advances in synthetic seed technology of conifers. *Plant Cell, Tissue and Organ Culture* **35**, 1-35 (Revue).

— B —

Bailly, C., Audigier, C., Ladonne, F., Wagner, M. H., Coste, F., Corbineau, F., et Côme, D. (2001). Changes in oligosaccharide content and antioxidant enzyme activities in developing bean seeds as related to acquisition of drying tolerance and seed quality. *Journal of Experimental Botany* **52**, 701-708.

Balbuena, T. S., Silveira, V., Junqueira, M., Dias, L. L. C., Santa-Catarina, C., Shevchenko, A., et Floh, E. I. S. (2009). Changes in the 2-DE protein profile during zygotic embryogenesis in the Brazilian Pine (*Araucaria angustifolia*). *Journal of proteomics* **72**, 337-352.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *cell* **116**, 281-297.

Bauer, D., Biehler, K., Fock, H., Carrayol, E., Hirel, B., Migge, A., et Becker, T. W. (1997). A role for cytosolic glutamine synthetase in the remobilization of leaf nitrogen during water stress in tomato. *Physiologia Plantarum* **99**, 241-248.

Belmonte, M. F., Donald, G., Reid, D. M., Yeung, E. C., et Stasolla, C. (2005). Alterations of the glutathione redox state improve apical meristem structure and somatic embryo quality in white spruce (*Picea glauca*). *Journal of experimental botany* **56**, 2355-2364.

Bercetche, J., et Pâques, M. (1995). Somatic embryogenesis in maritime pine (*Pinus pinaster*). *In* "Somatic Embryogenesis in Woody Plants" (G. P. Jain S, Newton RJ, ed.), Vol. 3 Gymnosperms, pp. 269-285. Kluwer Academic Publishers, The Netherlands.

Bernal-Lugo, I., et Leopold, A. C. (1995). Seed stability during storage: Raffinose content and seed glassy state. *Seed Science Research* **5**, 75-80.

Bernier, F., et Berna, A. (2001). Germins and germin-like proteins: plant do-all proteins. But what do they do exactly? *Plant Physiology and Biochemistry* **39**, 545-554.

Bewley, J. D., Bradford, K., et Hilhorst, H. W. M. (2013). Development and maturation. *In* "Seeds. Physiology of development, germination and dormancy" (B. Bewley J.D., K.J, Hilhorst, H.W.M., Nonogaki, H, ed.), pp. 27-83. Springer, New-York.

Bidart, J., et Lacroix, L. (2006). Méthodes d'étude du génome et du transcriptome. *In* "Méthode d'étude en biologie moléculaire", pp. 372-384.

Bishop-Hurley, S. L., Gardner, R. C., et Walter, C. (2003). Isolation and molecular characterization of genes expressed during somatic embryo development in *Pinus radiata*. *Plant Cell, Tissue and Organ Culture* **74**, 267-281.

Black, M., Corbineau, F., Grzesikit, M., Guyi, P., et Côme, D. (1996). Carbohydrate metabolism in the developing and maturing wheat embryo in relation to its desiccation tolerance. *Journal of Experimental Botany* **47**, 161-169.

Blanc, G., Lardet, L., Martin, A., Jacob, J. L., et Carron, M. P. (2002). Differential carbohydrate metabolism conducts morphogenesis in embryogenic callus of *Hevea brasiliensis* (Müll. Arg.). *Journal of Experimental Botany* **53**, 1453-1462.

Blazquez, S., Olmos, E., Hernandez, J. A., Fernandez-Garcia, N., Fernandez, J. A., et Piqueras, A. (2009). Somatic embryogenesis in saffron (*Crocus sativus* L.). Histological differentiation and implication of some components of the antioxidant enzymatic system. *Plant Cell, Tissue and Organ Culture (PCTOC)* **97**, 49-57.

Bomal, C., Le, V. Q., et Tremblay, F. M. (2002). Induction of tolerance to fast desiccation in black spruce (*Picea mariana*) somatic embryos: relationship between partial water loss, sugars, and dehydrins. *Physiologia Plantarum* **115**, 523-530.

Bonga, J. M., Klimaszewska, K. K., et von Aderkas, P. (2010). Recalcitrance in clonal propagation, in particular of conifers. *Plant Cell, Tissue and Organ Culture* **100**, 241-254 (Revue).

Bouffier, L., Raffin, A., et Kremer, A. (2009). La gestion de la variabilité génétique dans le programme d'amélioration du pin maritime (*Pinus pinaster* ait.). *Revue Forestière Française* **61**, 369-388.

Boyer, C., Hilbert, J.-L., et Vasseur, J. (1993). Embryogenesis-related protein synthesis and accumulation during early acquisition of somatic embryogenesis competence in *Cichorium*. *Plant Science* **93**, 41-53.

Bozhkov, P. V., Filonova, L. H., Suarez, M. F., et Gerald, P. S. (2005). Programmed Cell Death in Plant Embryogenesis. *In* "Current Topics in Developmental Biology", Vol. Volume 67, pp. 135-179. Academic Press.

Bray, E. A. (1993). Molecular responses to water deficit. *Plant physiology* **103**, 1035.

Breton, D., Harvengt, L., Trontin, J. F., Bouvet, A., et Favre, J. M. (2005). High subculture frequency, maltose-based and hormone-free medium sustained early development of somatic embryos in Maritime pine. *In Vitro Cellular & Developmental Biology - Plant* **41**, 494-504.

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

Brownfield, D. L., Todd, C. D., Stone, S. L., Deyholos, M. K., et Gifford, D. J. (2007). Patterns of storage protein and triacylglycerol accumulation during loblolly pine somatic embryo maturation. *Plant cell, tissue and organ culture* **88**, 217-223.

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

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

— C —

Cairney, J., Newton, R. J., Funkhouser, E. A., et Chang, S. (1995). Nucleotide sequence of a cDNA from *Atriplex canescens* (Pursh.) Nutt. A homolog of a jasmonate-induced protein from barley. *Plant Physiology* **108**, 1289-1290.

Cairney, J., et Pullman, G. S. (2007). The cellular and molecular biology of conifer embryogenesis. *New Phytologist* **176**, 511-536.

Cairney, J., Zheng, L., Cowels, A., Hsiao, J., Zismann, V., Liu, J., Ouyang, S., Thibaud-Nissen, F., Hamilton, J., Childs, K., Pullman, G. S., Zhang, Y. T., Oh, T., et Buell, C. R. (2006). Expressed Sequence Tags from loblolly pine embryos reveal similarities with angiosperm embryogenesis. *Plant Molecular Biology* **62**, 485-501.

Canales, J., Bautista, R., Label, P., Gomez-Maldonado, J., Lesur, I., Fernández-Pozo, N., Rueda-López, M., Guerrero-fernández, D., Castro-Rodriguez, V., Benzekri, H., Canas, R., Guevara, M.-A., Rodrigues, A., Seoane, P., Teyssier, C., Ehrenmann, F., Morel, A., Le Provost, G., Lalanne, C., Noirot, C., Klopp, C., Raymond, I., Garcia-Gutierrez, A., Trontin, J.-F., Lelu-Walter, M.-A., Miguel, C., Cervera, M. T., canton, F., Plomion, C., Harvengt, L., Avila, C., Claros, M. G., et Canovas, F. (2013). De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. *Plant Biotechnology* (sous presse).

Cella, R., et Carbonera, D. (1997). Peroxidases and morphogenesis. *Plant Peroxidase Newsletter* **10**, 24-29.

Chen, J., Uebbing, S., Gyllenstrand, N., Lagercrantz, U., Lascoux, M., et Kallman, T. (2012). Sequencing of the needle transcriptome from Norway spruce (*Picea abies* Karst L.) reveals lower substitution rates, but similar selective constraints in gymnosperms and angiosperms. *BMC genomics* **13**, 589.

Chiwocha, S., et Aderkas, P. v. (2002). Endogenous levels of free and conjugated forms of auxin, cytokinins and abscisic acid during seed development in Douglas fir. *Plant Growth Regulation* **36**, 191-200.

Ciavatta, V. T., Egertsdotter, U., Clapham, D., Arnold, S. v., et Cairney, J. (2002). A promoter from the loblolly pine *PtNIP1;1* gene directs expression in an early-embryogenesis and suspensor-specific fashion. *Planta* **215**, 694-698.

Ciavatta, V. T., Morillon, R., Pullman, G. S., Chrispeels, M. J., et Cairney, J. (2001). An aquaglyceroporin is abundantly expressed early in the development of the suspensor and the embryo proper of loblolly pine. *Plant Physiology* **127**, 1556-1567.

Colin, A., Meredieu, C., Labbe , T., et Belouard, T. (2010). Etude rétrospective et mise à jour de la ressource en pin maritime du massif des Landes de Gascogne après la tempête Klaus du 24 janvier 2009 *Etat des lieux de la ressource du massif landais après la tempête Klaus de 2009 – IGF rapport final – décembre 2010*

Cordewener, J., Booij, H., van der Zandt, H., van Engelen, F., et de Vries, S. (1991). Tunicamycin-inhibited carrot somatic embryogenesis can be restored by secreted cationic peroxidase isoenzymes. *Planta* **184**, 478-486.

Crowe, J. H., Hoekstra, F. A., et Crowe, L. M. (1992). Anhydrobiosis. *Annual Review of Physiology* **54**, 579-599.

— D —

de Jong, A. J., Schmidt, E. D. L., et de Vries, S. C. (1993). Early events in higher-plant embryogenesis. *Plant Molecular Biology* **22**, 367-377.

de Vega-Bartol, J. J., Santos, R. R., Simoes, M., et Miguel, C. M. (2013a). Normalizing gene expression by quantitative PCR during somatic embryogenesis in two representative conifer species: *Pinus pinaster* and *Picea abies*. *Plant Cell Reports* **32**, 715-729.

de Vega-Bartol, J. J., Santos, R. R., Simoes, M., et Miguel, C. M. (2013b). Normalizing gene expression by quantitative PCR during somatic embryogenesis in two representative conifer species: *Pinus pinaster* and *Picea abies*. *Plant cell reports*, 1-15.

de Vega-Bartol, J. J., Simoes, M., Lorenz, W. W., Rodrigues, A. S., Alba, R., Dean, J. F. D., et Miguel, C. M. (2013c). Transcriptomic analysis highlights epigenetic and transcriptional regulation during zygotic embryo development of *Pinus pinaster*. *BMC plant biology* **13**, 123.

Dembinsky, D., Woll, K., Saleem, M., Liu, Y., Fu, Y., Borsuk, L. A., Lamkemeyer, T., Fladerer, C., Madlung, J., Barbazuk, B., Nordheim, A., Nettleton, D., Schnable, P. S., et Hochholdinger, F. (2007). Transcriptomic and proteomic analyses of pericycle cells of the maize primary root. *Plant Physiology* **145**, 575-588.

Dixon, R., et Paiva, N. (1995). Stress-Induced Phenylpropanoid Metabolism. *Plant Cell.* **7** 1085–1097.

Dodeman, V. L., Ducreux, G., et Kreis, M. (1997). Zygotic embryogenesis versus somatic embryogenesis - a review. *Journal of Experimental Botany* **48**, 1493-1509 (Revue).

Dogra, P. D. (1967). Seed Sterility and disturbances in Embryogeny in conifers with particular reference to seed testing and tree breeding in *Pinaceae*. *In* "Studia Forestia Suecica", Vol. 45, pp. 97, Stockholm.

Dohm, J. C., Lottaz, C., Borodina, T., et Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic acids research* **36**, e105-e105.

Domon, J.-M., Neutelings, G., Roger, D., David, A., et David, H. l. n. (2000). A Basic Chitinase-like Protein Secreted by Embryogenic Tissues of *Pinus caribaea* acts on Arabinogalactan Proteins Extracted from the same Cell Lines. *Journal of Plant Physiology* **156**, 33-39.

Domon, J. M., Dumas, B., Laine, E., Meyer, Y., David, A., et David, H. (1995). Three glycosylated polypeptides secreted by several embryogenic cell cultures of pine show highly specific serological affinity to antibodies directed against the wheat germin apoprotein monomer. *Plant Physiol* **108**, 141-8.

Dong, J., et Dunstan, D. I. (1996). Characterization of three heat-shock-protein genes and their developmental regulation during somatic embryogenesis in white spruce [*Picea glauca* (Moench) Voss]. *Planta* **200**, 85-91.

Dong, J., et Dunstan, D. I. (1997a). Characterization of cDNAs representing five abscisic acid-responsive genes associated with somatic embryogenesis in *Picea glauca*, and their responses to abscisic acid stereostructure. *Planta* **203**, 448-453.

Dong, J., et Dunstan, D. I. (1997b). Endochitinase and beta-1, 3-glucanase genes are developmentally regulated during somatic embryogenesis in *Picea glauca*. *Planta* **201**, 189-194.

Downie, B., et Bewley, J. D. (2000). Soluble sugar content of white spruce (*Picea glauca*) seeds during and after germination. *Physiologia Plantarum* **110**, 1-12.

Dronne, S., Label, P., et Lelu, M.-A. (1997). Desiccation decreases abscisic acid content in hybrid larch (*Larix ×leptoeuropaea*) somatic embryos. *Physiologia Plantarum* **99**, 433-438.

Dunwell, J. M., Gibbings, J. G., Mahmood, T., et Saqlan Naqvi, S. M. (2008). Germin and germin-like proteins: evolution, structure, and function. *Critical Reviews in Plant Sciences* **27**, 342-375 (Revue).

Dunwell, J. M., Khuri, S., et Gane, P. J. (2000). Microbial Relatives of the Seed Storage Proteins of Higher Plants: Conservation of Structure and Diversification of Function during Evolution of the Cupin Superfamily. *Microbiology and Molecular Biology Reviews* **64**, 153-179 (Revue).

Dure, L. (1993). A repeating 11-mer amino acid motif and plant desiccation. *The Plant Journal* **3**, 363-369.

Dure, L., Crouch, M., Harada, J., Ho, T. H. D., Mundy, J., Quatrano, R., Thomas, T., et Sung, Z. R. (1989). Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology* **12**, 475-486.

Dure, L., Greenway, S. C., et Galau, G. A. (1981). Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. *Biochemistry* **20**, 4162-4168.

— E —

Egertsdotter, U., et Arnold, S. v. (1995). Importance of arabinogalactan proteins for the development of somatic embryos of Norway spruce (*Picea abies*). *Physiologia Plantarum* **93**, 334-345.

Egertsdotter, U., et von Arnold, S. (1998). Development of somatic embryos in Norway spruce. *Journal of Experimental Botany* **49**, 155-162.

Engelmann, F. (1991). In vitro conservation of tropical plant germplasm - a review. *Euphytica* **57**, 227-243 (Revue).

Erlich, Y., et Mitra, P. P. (2008). Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nature methods* **5**, 679-682.

— F —

Faurobert, M., Mihr, C., Bertin, N., Pawlowski, T., Negroni, L., Sommerer, N., et Causse, M. (2007). Major proteome variations associated with cherry tomato pericarp development and ripening. *Plant Physiology* **143**, 1327-1346.

Favennec, J. (1999). Aménagement des forêts littorales: cas des forêts dunaires du littoral atlantique français. *Revue forestière française* **51**, 217-229.

Feher, A., Pasternak, T. P., et Dudits, D. (2003). Transition of somatic plant cells to an embryogenic state. *Plant Cell, Tissue and Organ Culture* **74**, 201-228 (Revue).

Fernandez-Pozo, N., Canales, J., Guerrero-Fernandez, D., Villalobos, D. P., Diaz-Moreno, S. M., Bautista, R., Flores-Monterroso, A., Guevara, M. a., Perdiguero, P., et Collada, C. (2011). EuroPineDB: a high-coverage web database for maritime pine transcriptome. *BMC genomics* **12**, 366.

Fiehn, O. (2002). Metabolomics : the link between genotypes and phenotypes. *Plant Molecular Biology* **48**, 155-171.

Filonova, L. H., Bozhkov, P. V., et Arnold, S. v. (2000). Developmental pathway of somatic embryogenesis in *Picea abies* as revealed by time-lapse tracking. *Journal of Experimental Botany* **51**, 249-264.

Finkelstein, R. R., Gampala, S. S. L., et Rock, C. D. (2002). Abscisic acid signaling in seeds and seedlings. *The Plant Cell Online* **14**, S15-S45.

Finnie, C., Bak-Jensen, K. S., Laugesen, S., Roepstorff, P., et Svensson, B. (2006). Differential appearance of isoforms and cultivar variation in protein temporal profiles revealed in the maturing barley grain proteome. *Plant Science* **170**, 808-821.

Fischerova, L., Fischer, L., Vondrakova, Z., et Vagner, M. (2008). Expression of the gene encoding transcription factor PaVP1 differs in *Picea abies* embryogenic lines depending on their ability to develop somatic embryos. *Plant cell reports* **27**, 435-441.

Flinn, B. S., Roberts, D. R., et Taylor, I. E. P. (1991). Evaluation of somatic embryos of interior spruce. Characterization and developmental regulation of storage proteins. *Physiologia Plantarum* **82**, 624-632.

Footitt, S., Ingouff, M., Clapham, D., et von Arnold, S. (2003). Expression of the viviparous 1 (*Pavp1*) and p34cdc2 protein kinase (*cdc2Pa*) genes during somatic embryogenesis in Norway spruce (*Picea abies* [L.] Karst). *Journal of experimental botany* **54**, 1711-1719.

— G —

Gaj, M. D., Zhang, S., Harada, J. J., et Lemaux, P. G. (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of *Arabidopsis*. *Planta* **222**, 977-988.

Galau, G. A., Hughes, D. W., et Dure, L. (1986). Abscisic acid induction of cloned cotton late embryogenesis-abundant (LEA) mRNAs. *Plant Molecular Biology* **7**, 155-170.

Gallardo, K., Le Signor, C., Vandekerckhove, J., Thompson, R. D., et Burstin, J. (2003). Proteomics of Medicago truncatula seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol* **133**, 664-82.

Gallardo, K., Firnhaber, C., Zuber, H., Hericher, D., Belghazi, M., Henry, C., Kuster, H., et Thompson, R. (2007). A combined proteome and transcriptome analysis of developing *Medicago truncatula* seeds: evidence for metabolic specialization of maternal and filial tissues. *Mol Cell Proteomics* **6**, 2165-79.

Gallois, J.-L., Nora, F. R., Mizukami, Y., et Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes & development* **18**, 375-380.

Gibson, S. I. (2004). Sugar and phytohormone response pathways: navigating a signalling network. *Journal of Experimental Botany* **55**, 253-264.

Gifford, D. J. (1988). An electrophoretic analysis of the seed proteins from *Pinus monticola* and eight other species of pine. *Canadian journal of botany* **66**, 1808-1812.

Gion, J. M., Lalanne, C., Le Provost, G., Ferry-Dumazet, H., Paiva, J., Chaumeil, P., Frigerio, J. M., Brach, J., Barré, A., de Daruvar, A., Claverol, S., Bonneu, M., Sommerer, N., Negroni, L., et Plomion, C. (2005). The proteome of maritime pine wood forming tissue. *PROTEOMICS* **5**, 3731-3751.

Goldberg, R. B., Barker, S. J., et Perez-Grau, L. (1989). Regulation of gene expression during plant embryogenesis. *Cell* **56**, 149-160.

Gomez-Garay, A., Lopez, J. A., Camafeita, E., Bueno, M. A., et Pintos, B. (2013). Proteomic perspective of *Quercus suber* somatic embryogenesis. *Journal of Proteomics* **93**, 314-325.

Gosslova, Monika, Svobodova, Hana, Lipavska, Helena, Albrechtova, Jana, Vreugdenhil, et Dick (2001). Comparing carbohydrate status during norway spruce seed development and somatic embryo formation. *In Vitro Cellular & Developmental Biology-Plant* **37**, 24-28.

Graham, I. A. (1996). Carbohydrate control of gene expression in higher plants. *Research in microbiology* **147**, 572-580.

Green, M. J., McLeod, J. K., et Misra, S. (1991). Characterization of Douglas fir protein body composition by SDS-PAGE and electron microscopy. *Plant physiology and biochemistry* **29**, 49-55.

Grzelczak, Z. F., et Lane, B. G. (1984). Signal resistance of a soluble protein to enzymic proteolysis. An unorthodox approach to the isolation and purification of germin, a rare growth-related protein. *Canadian journal of biochemistry and cell biology* **62**, 1351-1353.

Gupta, P. K., et Pullman, G. S. (1990). Method for reproducing coniferous plants by somatic embryogenesis. US Patent No. 4,957,866.

Gupta, P. K., et Pullman, G. S. (1991). Method for reproducing coniferous plants by somatic embryogenesis using abscisic acid and osmotic potential variation., US Patent No. 5036007.

Gygi, S. P., Rochon, Y., Franza, B. R., et Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Molecular and cellular biology* **19**, 1720-1730.

— H —

Hakman, I. (1993a). Embryology in Norway spruce (*Picea abies*). An analysis of the composition of seed storage proteins and deposition of storage reserves during seed development and somatic embryogenesis. *Physiologia Plantarum* **87**, 148-159.

Hakman, I. (1993b). Embryology in Norway spruce (*Picea abies*). Immunochemical studies on transport of a seed storage protein. *Physiologia Plantarum* **88**, 427-433.

Hakman, I., Fowke, L. C., Arnold, S. v., et Eriksson, T. (1985). The development of somatic embryos in tissue cultures initiated from immature embryos of *Picea abies* (Norway spruce). *Plant Science, Irish Republic* **38**, 53-59.

Hakman, I., Hallberg, H., et Palovaara, J. (2009). The polar auxin transport inhibitor NPA impairs embryo morphology and increases the expression of an auxin efflux facilitator protein PIN during *Picea abies* somatic embryo development. *Tree physiology* **29**, 483-496.

Hara, M., Terashima, S., et Kuboi, T. (2001). Characterization and cryoprotective activity of cold-responsive dehydrin from *Citrus unshiu*. *Journal of Plant Physiology* **158**, 1333-1339.

He, L., et Hannon, G. J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nature Reviews Genetics* **5**, 522-531.

Helal, N. A. S. (2011). The green revolution via synthetic (artificial) seeds: a review. *Research Journal of Agriculture and Biological Sciences* **7**, 464-477 (Revue).

Hetherington, A. M., et Quatrano, R. S. (1991). Mechanisms of action of abscisic acid at the cellular level. *New Phytologist* **119**, 9-12.

Higgins, T. J. V. (1984). Synthesis and regulation of major proteins in seeds. *Annual Review of Plant Physiology* **35**, 191-221 (Revue).

Hilbert, J.-L., Dubois, T., et Vasseur, J. (1992). Detection of embryogenesis-related proteins during somatic embryo formation in *Cichorium*. *Plant physiology and biochemistry* **30**, 733-741.

Hjortswang, H. I., Filonova, L. H., Vahala, T., et Arnold, S. v. (2002). Modified expression of the *Pa18* gene interferes with somatic embryo development in Norway spruce. *Plant Growth Regulation* **38**, 75-82.

Hoekstra, F. A., Golovina, E. A., Tetteroo, F. A. A., et Wolkers, W. F. (2001). Induction of Desiccation Tolerance in Plant Somatic Embryos: How Exclusive Is the Protective Role of Sugars? *Cryobiology* **43**, 140-150.

Hohmann, S., Nielsen, S. r., Weig, A., et Eisenbarth, D. (2000). Role of Aquaporins During Elongation Growth of Castor Bean Seedlings. *In* "Molecular Biology and Physiology of Water and Solute Transport", pp. 357-364. Springer US.

Hughes, D. W., et Galau, G. A. (1991). Developmental and environmental induction of *Lea* and *LeaA* mRNAs and the postabscission program during embryo culture. *Plant Cell* **3**, 605-618.

Humanez, A., Blasco, M., Brisa, C., Segura, J., et Arrillaga, I. (2012). Somatic embryogenesis from different tissues of Spanish populations of maritime pine. *Plant Cell, Tissue and Organ Culture (PCTOC)* **111**, 373-383.

— I —

Ikeda, M., et Kamada, H. (2006). Comparison of molecular mechanisms of somatic and zygotic embryogenesis. *In* "Somatic embryogenesis" (A. Mujib, Samaj, J, ed.), pp. 51-68. Springer-Verlag Berlin Heidelberg.

Imin, N., Nizamidin, M., Daniher, D., Nolan, K. E., Rose, R. J., et Rolfe, B. G. (2005). Proteomic Analysis of Somatic Embryogenesis in *Medicago truncatula*. Explant Cultures Grown under 6-Benzylaminopurine and 1-Naphthaleneacetic Acid Treatments. *Plant Physiology* **137**, 1250-1260.

Ingouff, M., Farbos, I., Lagercrantz, U., et von Arnold, S. (2001). PaHB1 is an evolutionary conserved HD-GL2 homeobox gene expressed in the protoderm during *Norway spruce* embryo development. *Genesis* **30**, 220-230.

Ingram, J., et Bartels, D. (1996). The molecular basis of dehydration tolerance in plants. *Annual review of plant biology* **47**, 377-403 (Revue).

Iraqi, D., et Tremblay, F. M. (2001a). Analysis of carbohydrate metabolism enzymes and cellular contents of sugars and proteins during spruce somatic embryogenesis suggests a regulatory role of exogenous sucrose in embryo development. *Journal of Experimental Botany* **52**, 2301-2311.

Iraqi, D., et Tremblay, F. M. (2001b). The role of sucrose during maturation of black spruce (*Picea mariana*) and white spruce (*Picea glauca*) somatic embryos. *Physiologia Plantarum* **111**, 381-388.

Isik, F., Li, B. L., et Frampton, J. (2003). Estimates of additive, dominance and epistatic genetic variances from a clonally replicated test of loblolly pine. *Forest Science* **49**, 77-88.

— J —

Jaikaran, A. S., Kennedy, T. D., Dratewka-Kos, E., et Lane, B. G. (1990). Covalently bonded and adventitious glycans in germin. *Journal of Biological Chemistry* **265**, 12503-12512.

Jarlet-Hugues, E. (1989). Recherches sur l'aptitude à l'embryogenèse somatique de matériel juvénile et de matériel issu d'arbres adultes de *Pinus pinaster* Sol., Doctorat d'Université Pierre et Marie Curie, Paris VI, 135 pp.

Jenik, P. D., et Barton, M. K. (2005). Surge and destroy: the role of auxin in plant embryogenesis. *Development* **132**, 3577-3585.

Jensen, U., et Berthold, H. (1989). Legumin-like proteins in gymnosperms. *Phytochemistry* **28**, 1389-1394.

Jensen, U., et Lixue, C. (1991). Abies seed protein profile divergent from other *Pinaceae*. *Taxon*, 435-440.

Jimenez, V. M., et Bangerth, F. (2001). *In vitro* culture and endogenous hormone levels in immature zygotic embryos, endosperm and callus cultures of normal and high-lysine barley genotypes. *Journal of Applied Botany* **75**, 1-7.

Johnson, K. L., Jones, B. J., Bacic, A., et Schultz, C. J. (2003). The fasciclin-like arabinogalactan proteins of *Arabidopsis*. A multigene family of putative cell adhesion molecules. *Plant Physiology* **133**, 1911-1925.

Jones-Rhoades, M. W., Bartel, D. P., et Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.* **57**, 19-53.

Jordy, M. N., et Favre, J. M. (2003). Spatio-Temporal Variations in Starch Accumulation During Germination and Post-Germinative Growth of Zygotic and Somatic Embryos of *Pinus pinaster*. *Biologia Plantarum* **46**, 507-512.

Jordy, M. N. L. (2004). Seasonal variation of organogenetic activity and reserves allocation in the shoot apex of *Pinus pinaster* Ait. *Annals of botany* **93**, 25-37.

Jorge, I., Navarro, R. M., Lenz, C., Ariza, D., et Jorran, J. (2006). Variation in the holm oak leaf proteome at different plant developmental stages, between provenances and in response to drought stress. *Proteomics* **6**, S207-S214.

– K –

Kader, J.-C. (1996). Lipid-transfer proteins in plants. *Annual review of plant biology* **47**, 627-654 (Revue).

Kairong, C., Ji, L., Gengmei, X., Jianlong, L., Lihong, W., et Yafu, W. (2002). Effect of hydrogen peroxide on synthesis of proteins during somatic embryogenesis in *Lycium barbarum*. *Plant Cell, Tissue and Organ Culture* **68**, 187-193.

Kapros, T., Bögre, L., Németh, K., Bako, L., Györgyey, J., Wu, S., et Dudits, D. (1992). Differential Expression of Histone H3 Gene Variants during Cell Cycle and Somatic Embryogenesis in Alfalfa. *Plant Physiology* **98**, 621-625.

Kawahara, R., Sunabori, S., Fukuda, H., et Komamine, A. (1992). A gene expressed preferentially in the globular stage of somatic embryogenesis encodes elongation-factor 1α in carrot. *European Journal of Biochemistry* **209**, 157-162.

Kermode, A. R. (1990). Regulatory mechanisms involved in the transition from seed development to germination. *Critical Reviews in Plant Sciences* **9**, 155-195 (Revue).

Kermode, A. R. (1997). Approaches to elucidate the basis of desiccation-tolerance in seeds. *Seed Science Research* **7**, 75-95.

Kermode, A. R., et Finch-Savage, B. E. (2002). Desiccation sensitivity in orthodox and recalcitrant seeds in relation to development. *In* "Desiccation and survival in plants: drying without dying", pp. 149-184. Cabi, Wallingford, UK.

Kingston-Smith, A. H., Walker, R. P., et Pollock, C. J. (1999). Invertase in leaves: Conundrum or control point ? *Journal of Experimental Botany* **50**, 735-743.

Klimaszewska, K., Bernier-Cardou, M., Cyr, D., et Sutton, B. (2000). Influence of gelling agents on culture medium gel strength, water availability, tissue water potential, and maturation response in embryogenic cultures of *Pinus strobus* L. *In Vitro Cellular & Developmental Biology - Plant* **36**, 279-286.

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

Klimaszewska, K., Noceda, C., Pelletier, G., Label, P., Rodriguez, R., et Lelu-Walter, M. (2009). Biological characterization of young and aged embryogenic cultures of *Pinus pinaster* (Ait.). *In Vitro Cellular & Developmental Biology - Plant* **45**, 20-33.

Klimaszewska, K., Park, Y. S., Overton, C., Maceacheron, I., et Bonga, J. M. (2001). Optimized somatic embryogenesis in *Pinus strobus* L. *In Vitro Cellular & Developmental Biology - Plant* **37**, 392-399.

Klimaszewska, K., Trontin, J.-F., Becwar, M., Devillard, C., Park, Y.-S., et Lelu-Walter, M.-A. (2007). Recent progress in somatic embryogenesis of four *Pinus spp*. *Tree For. Sci. Biotech.* **1**, 11-25 (Revue).

Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. *Humangenetik* **26**, 231-243.

Kong, L., et Aderkas, P. v. (2007). Genotype effects on ABA consumption and somatic embryo maturation in interior spruce (*Picea glauca* * engelmanni). *Journal of Experimental Botany* **58**, 1525-1531.

Konradova, H., Gricova, M., et Lipavska, H. (2003). Cold-induced accumulation of raffinose family oligosaccharides in somatic embryos of Norway spruce (*Picea abies*). *In Vitro Cellular & Developmental Biology - Plant* **39**, 425-427.

Konradova, H., Lipavska, H., Albrechtova, J., et Vreugdenhil, D. (2002). Sucrose metabolism during somatic and zygotic embryogeneses in Norway spruce: content of

soluble saccharides and localisation of key enzyme activities. *Journal of Plant Physiology* **159**, 387-396.

Kremer, A. (1992). Predictions of age-age correlations of total height based on serial correlations between height increments in Maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics* **85**, 152-158.

Kreuger, M., et Holst, G.-J. (1993). Arabinogalactan proteins are essential in somatic embryogenesis of *Daucus carota* L. *Planta* **189**, 243-248.

Kwong, R. W., Bui, A. Q., Lee, H., Kwong, L. W., Fischer, R. L., Goldberg, R. B., et Harada, J. J. (2003). LEAFY COTYLEDON1-LIKE Defines a Class of Regulators Essential for Embryo Development. *The Plant Cell Online* **15**, 5-18.

— L —

Label, P., et Lelu, M.-A. (2000). Exogenous abscisic acid fate during maturation of hybrid larch (*Larix* x *leptoeuropaea*) somatic embryos. *Physiologia Plantarum* **109**, 456-462.

Lane, B. G. (2000). Oxalate oxidases and differentiating surface structure in wheat: germins. *The Biochemical journal* **349**, 309-321.

Lane, B. G. (2002). Oxalate, Germins, and High Plant Pathogens. *IUBMB life* **53**, 67-75.

Lara-Chavez, A., Egertsdotter, U., et Flinn, B. S. (2012). Comparison of gene expression markers during zygotic and somatic embryogenesis in pine. *In Vitro Cellular & Developmental Biology-Plant* **48**, 341-354.

Larkins, B., Vasil, I., Schwartz, B., Vernon, D., et Meinke, D. (1997). Development of the Suspensor: Differentiation, Communication, and Programmed Cell Death During Plant Embryogenesis. *In* "Cellular and Molecular Biology of Plant Seed Development", Vol. 4, pp. 53-72. Springer Netherlands.

Le, B. H., Cheng, C., Bui, A. Q., Wagmaister, J. A., Henry, K. F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R., et Horvath, S. (2010). Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences* **107**, 8063-8070.

Leal, I., et Misra, S. (1993). Molecular cloning and characterization of a legumin-like storage protein cDNA of Douglas fir seeds. *Plant molecular biology* **21**, 709-715.

Lee, E. K., Cho, D. Y., et Soh, W. Y. (2001). Enhanced production and germination of somatic embryos by temporary starvation in tissue cultures of *Daucus carota*. *Plant Cell Reports* **20**, 408-415.

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

Lelu-Walter, M.-A., Thompson, D., Harvengt, L., Sanchez, L., Toribio, M., et Pâques, L. (2013). Somatic embryogenesis in forestry with a focus on Europe: state-of-the-art, benefits, challenges and future direction. *Tree Genetics & Genomes* **9**, 883-899 (Revue).

Lelu-Walter, M. A., Bernier-Cardou, M., et Klimaszewska, K. (2006). Simplified and improved somatic embryogenesis for clonal propagation of *Pinus pinaster* (Ait.). *Plant Cell Rep* **25**, 767-76.

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

Lelu, M. A., et Label, P. (1994). Changes in the levels of abscisic acid and its glucose ester conjugate during maturation of hybrid larch (*Larix* × *leptoeuropaea*) somatic embryos, in relation to germination and plantlet recovery. *Physiologia Plantarum* **92**, 53-60.

Li, W.-F., Zhang, S.-G., Han, S.-Y., Wu, T., Zhang, J.-H., et Qi, L.-W. (2014). The post-transcriptional regulation of *La*SCL6 by miR171 during maintenance of embryogenic potential in *Larix kaempferi* (Lamb.) Carr. *Tree Genetics & Genomes* **10**, 223-229.

Li, X., Wu, H. X., Dillon, S. K., et Southerton, S. G. (2009). Generation and analysis of expressed sequence tags from six developing xylem libraries in *Pinus radiata* D. Don. *BMC genomics* **10**, 41.

Linkies, A., Graeber, K., Knight, C., et Leubne-Metzger, G. (2010). The evolution of seeds. *New Phytologist* **186**, 817-831.

Lipavskà, H., et Konradova, H. (2004). Somatic embryogenesis in conifers: the role of carbohydrate metabolism. *In Vitro Cellular & Developmental Biology-Plant* **40**, 23-30 (Revue).

Lipavskà, H., Svobodová, H., Albrechtová, J., Kumstýřová, L., Vágner, M., et Vondráková, Z. (2000). Carbohydrate status during somatic embryo maturation in Norway spruce. *In Vitro Cellular & Developmental Biology - Plant* **36**, 260-267.

Lippert, D., Zhuang, J., Ralph, S., Ellis, D. E., Gilbert, M., Olafson, R., Ritland, K., Ellis, B., Douglas, C. J., et Bohlmann, J. (2005). Proteome analysis of early somatic embryogenesis in *Picea glauca*. *Proteomics* **5**, 461-73.

Lu, J., Vahala, J., et Pappinen, A. (2011). Involvement of ethylene in somatic embryogenesis in Scots pine (*Pinus sylvestris* L.). *Plant Cell, Tissue and Organ Culture (PCTOC)* **107**, 25-33.

— M —

Manfre, A. J., LaHatte, G. A., Climer, C. R., et Marcotte, W. R. (2009). Seed dehydration and the establishment of desiccation tolerance during seed maturation is altered in the *Arabidopsis thaliana* mutant atem6-1. *Plant and cell physiology* **50**, 243-253.

Mathieu, M., Lelu-Walter, M. A., Blervacq, A. S., David, H., Hawkins, S., et Neutelings, G. (2006). Germin-like genes are expressed during somatic embryogenesis and early development of conifers. *Plant Molecular Biology* **61**, 615-627.

Mathieu, M., Neutelings, G., Hawkins, S., Grenier, E., et David, H. (2003). Cloning of a pine germin-like protein (GLP) gene promoter and analysis of its activity in transgenic tobacco Bright Yellow 2 cells. *Physiologia Plantarum* **117**, 425-434.

Mayer, K. F. X., Schoof, H., Haecker, A., Lenhard, M., Jurgens, G., et Laux, T. (1998). Role of WUSCHEL in Regulating Stem Cell Fate in the *Arabidopsis* Shoot Meristem. *Cell* **95**, 805-815.

Merkle, S. A., Parrott, W. A., et Flinn, B. S. (1995). Morphogenic aspects of somatic embryogenesis. *In* "In vitro embryogenesis in plants", pp. 155-203. Springer.

Merzeau, D., Canteloup, D., et Raffin, A. (2002). Les gains génétiques des premières et deuxième générations de vergers. Le progrès génétique en forêt. *In* "Groupe Pin Maritime du Futur", pp. 49-56. PG Edition.

Metzker, M. L. (2010). Sequencing technologie : the next generation. *Nature Reviews Genetics* **11**, 31-46 (Revue).

Miguel, C., Goncalves, S., Tereso, S., Marum, L., Maroco, J., et Oliveira, M. M. (2004). Somatic embryogenesis from 20 open-pollinated families of Portuguese plus trees of maritime pine. *Plant Cell, Tissue and Organ Culture* **76**, 121-130.

Misra, S., et Green, M. J. (1990a). Developmental gene expression in conifer embryogenesis and germination. I, Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* [moench] voss.). *Plant Science* **68**, 163-173.

Misra, S., et Green, M. J. (1990b). Developmental gene expression in conifer embryogenesis and germination. I. Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* [Moench] Voss.). *Plant Science* **68**, 163-173.

Misra, S., Kermode, A., et Bewley, J. D. (1985). Maturation drying as the" switch" that terminates seed development and promotes germination. *Molecular form and function of the plant genome*, 113-128.

Molina, A., Segura, A., et Garcia-Olmedo, F. (1993). Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. *FEBS letters* **316**, 119-122.

Mondal, T. K., Bhattacharya, A., Sood, A., et Ahuja, P. S. (2002). Factors affecting germination and conversion frequency of somatic embryos of Tea [*Camellia sinensis* (L.) O. Kuntze]. *Journal of plant physiology* **159**, 1317-1321.

Murray, B. G. (1998). Nuclear DNA amounts in gymnosperms. *Annals of Botany* **82**, 3-15.

Muzquiz, M., Burbano, C., Pedrosa, M. M., Folkman, W., et Gulewicz, K. (1999). Lupins as a potential source of raffinose family oligosaccharides: Preparative method for their isolation and purification. *Industrial Crops and Products* **9**, 183-188.

— N —

Nagmani, R., Diner, A., Garton, S., et Zipf, A. (1995). Anatomical comparison of somatic and zygotic embryogeny in conifers. *In* "Jain S, Gupta P, Newton R, eds. Somatic embryogenesis in woody plants" (K. Academic, ed.), Vol. 1, pp. 23–48, Dordrecht, the Netherlands.

Neutelings, G., Domon, J., Membré, N., Bernier, F., Meyer, Y., David, A., et David, H. (1998). Characterization of a germin-like protein gene expressed in somatic and zygotic embryos of pine (*Pinus caribaea* Morelet). *Plant Molecular Biology* **38**, 1179-1190.

Newton, C. H., Flinn, B. S., et Sutton, B. C. S. (1992). Vicilin-like seed storage proteins in the gymnosperm interior spruce (*Picea glauca/engelmanii*). *Plant molecular biology* **20**, 315-322.

Nolan, K. E., Kurdyukov, S., et Rose, R. J. (2009). Expression of the SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE1 (SERK1) gene is associated with developmental change in the life cycle of the model legume *Medicago truncatula*. *Journal of experimental botany* **60**, 1759-1771.

— O —

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *Journal of biological chemistry* **250**, 4007-4021.

Obendorf, R. L. (1997). Oligosaccharides and galactosyl cyclitols in seed desiccation tolerance. *Seed Science Research* **7**, 63-74.

Oh, T. J., Wartell, R. M., Cairney, J., et Pullman, G. S. (2008). Evidence for stage-pecific modulation of specific microRNAs (miRNAs) and miRNA processing components in zygotic embryo and female gametophyte of loblolly pine (*Pinus taeda*). *New Phytologist* **179**, 67-80.

Oliver, A. E., Crowe, L. M., et Crowe, J. H. (1998). Methods for dehydration-tolerance: depression of the phase transition temperature in dry membranes and carbohydrate vitrification. *Seed Science Research* **8**, 211-221.

Overvoorde, P. J., et Grimes, H. D. (1994). The role of calcium and calmodulin in carrot somatic embryogenesis. *Plant and cell physiology* **35**, 135-144.

— P —

Palmisano, A. M., Ambrogioni, L., et Caroppo, S. (1994). First record of a Bursaphelenchus species from *Pinus pinaster* in Italy1. *EPPO Bulletin* **24**, 467-474.

Palovaara, J., et Hakman, I. (2008). Conifer *WOX*-related homeodomain transcription factors, developmental consideration and expression dynamic of *WOX2* during *Picea abies* somatic embryogenesis. *Plant Molecular Biology* **66**, 533-549.

Palovaara, J., Hallberg, H., Stasolla, C., et Hakman, I. (2010). Comparative expression pattern analysis of *WUSCHEL-related homeobox 2* (*WOX2*) and *WOX8/9* in developing seeds and somatic embryos of the gymnosperm *Picea abies*. *New Phytologist* **188**, 122-135.

Pan, Z., Zeng, Y., An, J., Ye, J., Xu, Q., et Deng, X. (2012). An integrative analysis of transcriptome and proteome provides new insights into carotenoid biosynthesis and regulation in sweet orange fruits. *Journal of Proteomics* **75**, 2670-2684.

Park, S., et Harada, J. J. (2008). *Arabidopsis* embryogenesis. *In* "Plant Embryogenesis", pp. 3-16. Springer.

Park, Y. S., Lelu-Walter, M. A., Harvengt, L., Trontin, J. F., MacEacheron, I., Klimaszewska, K., et Bonga, J. M. (2006). Initiation of somatic embryogenesis in *Pinus banksiana*, *P. strobus*, *P. pinaster*, and *P. sylvestris* at three laboratories in Canada and France. *Plant Cell, Tissue and Organ Culture* **86**, 87-101.

Pasternak, T., Potters, G., Caubergs, R., et Jansen, M. A. K. (2005). Complementary interactions between oxidative stress and auxins control plant growth responses at plant, organ, and cellular level. *Journal of Experimental Botany* **56**, 1991-2001.

Pavy, N., Paule, C., Parsons, L., Crow, J., Morency, M.-J., Cooke, J., Johnson, J., Noumen, E., Guillet-Claude, C., et Butterfield, Y. (2005). Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC genomics* **6**, 144.

Peace, E. A., Lea, P. J., et Darrall, N. M. (1995). The effect of open-air fumigation with SO2 and O3 on carbohydrate metabolism in Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*). *Plant, Cell & Environment* **18**, 277-283.

Pédezert, G. (2009). La Gironde en bref - Raport INSEE - octobre 2009.

Perez Rodriguez, M. J., Suarez, M. F., Heredia, R., Avila, C., Breton, D., Trontin, J. F., Filonova, L., Bozhkov, P., Arnold, S. v., Harvengt, L., et Canovas, F. M. (2006). Expression patterns of two glutamine synthetase genes in zygotic and somatic pine embryos support specific roles in nitrogen metabolism during embryogenesis. *New Phytologist* **169**, 35-44.

Pescador, R., Kerbauy, G. B., dos Santos, D. Y. A. C., Dal Vesco, L. L., de Freitas Fraga, H. P., et Guerra, M. P. (2012). Comparative study of reserve lipid accumulation

during somatic and zygotic *Acca sellowiana* (O. Berg.) Burret embryogenesis. *Acta Physiologiae Plantarum* **34**, 771-778.

Pignard, G., Bélouard, T., et Stach, N. (2001). Suivi de la ressource pin maritime en Aquitaine, Union Européenne PDZR, Derf, Conseil Régional d'Aquitaine. *Inventaire Français National*.

Plomion, C., Chagne, D., Pot, D., Kumar, S., Wilcox, P. L., Burdon, R. D., Prat, D., Peterson, D. G., Paiva, J., Chaumeil, P., Vendramin, G. G., Sebastiani, F., Nelson, C. D., Echt, C. S., Savolainen, O., Kubisiak, T. L., Cervera, M. T., Maria, N. d., et Islam-Faridi, M. N. (2007). Pines. *In* "Forest trees", pp. 29-92. Springer-Verlag GmbH, Heidelberg, Germany.

Pot, D., Chantre, G., Rozenberg, P., Rodrigues, J. C., Jones, G. L., Pereira, H., Hannrup, B., Cahalan, C., et Plomion, C. (2002). Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). *Annals of Forest Science* **59**, 563-575.

Pullman, G. S., et Buchanan, M. (2008). Identification and quantitative analysis of stage-specific carbohydrates in loblolly pine (*Pinus taeda*) zygotic embryo and female gametophyte tissues. *Tree Physiology* **28**, 985-96.

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

— Q —

Quiroz-Figueroa, F., Fuentes-Cerda, C., Rojas-Herrera, R., et Loyola-Vargas, V. (2002). Histological studies on the developmental stages and differentiation of two different somatic embryogenesis systems of *Coffea arabica*. *Plant Cell Reports* **20**, 1141-1149.

— R —

Raghavan, V. (2006). Can carrot and *Arabidopsis* serve as model systems to study the molecular biology of somatic embryogenesis? *CURRENT SCIENCE-BANGALORE* **90**, 1336.

Rai, M. K., Shekhawat, N. S., Gupta, A. K., Phulwaria, M., Ram, K., et Jaiswal, U. (2011a). The role of abscisic acid in plant tissue culture: a review of recent progress. *Plant Cell, Tissue and Organ Culture (PCTOC)* **106**, 179-190.

Rai, M. K., Shekhawat, N. S., Harish, Gupta, A. K., Phulwaria, M., Ram, K., et Jaiswal, U. (2011b). The role of abscisic acid in plant tissue culture: a review of recent progress. *Plant Cell, Tissue and Organ Culture* **106**, 179-190 (Revue).

Ralph, S. G., Chun, H. J. E., Kolosova, N., Cooper, D., Oddy, C., Ritland, C. E., Kirkpatrick, R., Moore, R., Barber, S., et Holt, R. A. (2008). A conifer genomics resource of 200,000 spruce (*Picea* spp.) ESTs and 6,464 high-quality, sequence-finished full-length cDNAs for Sitka spruce (*Picea sitchensis*). *BMC genomics* **9**, 484.

Ramarosandratana, A., Harvengt, L., Bouvet, A., Calvayrac, R., et Pâques, M. (2001a). Effects of carbohydrate source, polyethylene glycol and gellan gum concentration of embryonal-suspensor mass (ESM) proliferation and maturation of maritime pine somatic embryos. *In Vitro Cellular & Developmental Biology - Plant* **37**, 29-34.

Ramarosandratana, A., Harvengt, L., Bouvet, A., Calvayrac, R., et Pâques, M. (2001b). Influence of the embryonal-suspensor mass (ESM) sampling on development and proliferation of maritime pine somatic embryos. *Plant Science* **160**, 473-479.

Renauld, F. (1995). "La fructification du pin maritime : influence du climat, du mode de conduite en verger et outils de prévision de production des cônes." Rapport de fin d'étude ISAB.

Rigault, P., Boyle, B., Lepage, P., Cooke, J. E. K., Bousquet, J., et MacKay, J. J. (2011). A white spruce gene catalog for conifer genome analyses. *Plant physiology* **157**, 14-28.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L., et Pachter, L. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* **12**, R22.

Roberts, D. R., Flinn, B. S., Webb, D. T., Webster, F. B., et Sutton, B. (1990a). Abscisic acid and indole-3-butyric acid regulation of maturation and accumulation of storage proteins in somatic embryos of interior spruce. *Physiologia Plantarum* **78**, 355-360.

Roberts, D. R., Sutton, B. C. S., et Flinn, B. S. (1990b). Synchronous and high frequency germination of interior spruce somatic embryos following partial drying at high relative humidity. *Canadian Journal of Botany* **68**, 1086-1090.

Robinson, A. R., Dauwe, R., Ukrainetz, N. K., Cullis, I. F., White, R., et Mansfield, S. D. (2009). Predicting the regenerative capacity of conifer somatic embryogenic cultures by metabolomics. *Plant biotechnology journal* **7**, 952-963.

Rodriguez-Serrano, M., Barany, I., Prem, D., Coronado, M.-J., Risueno, M. C., et Testillano, P. S. (2012). NO, ROS, and cell death associated with caspase-like activity increase in stress-induced microspore embryogenesis of barley. *Journal of experimental botany* **63**, 2007-2024.

Roessner, U., Luedemann, A., Brust, D., Fiehn, O., Linke, T., Willmitzer, L., et Fernie, A. R. (2001a). Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *The Plant Cell Online* **13**, 11-29.

Roessner, U., Willmitzer, L., et Fernie, A. R. (2001b). High-resolution metabolic phenotyping of genetically and environmentally diverse potato tuber systems. Identification of phenocopies. *Plant Physiology* **127**, 749-764.

— S —

Sabala, I., Elfstrand, M., Farbos, I., Clapham, D., et von Arnold, S. (2000). Tissue-specific expression of *Pa18*, a putative lipid transfer protein gene, during embryo development in Norway spruce (*Picea abies*). *Plant molecular biology* **42**, 461-478.

Sales, M. c. P., Gerhardt, I. R., Grossi-de-Sá, M. F., et Xavier-Filho, J. (2000). Do legume storage proteins play a role in defending seeds against bruchids? *Plant Physiology* **124**, 515-522.

Sanchez-Ballesta, M. T., Rodrigo, M. J., Lafuente, M. T., Granell, A., et Zacarias, L. (2004). Dehydrin from citrus, which confers in vitro dehydration and freezing protection activity, is constitutive and highly expressed in the flavedo of fruit but responsive to cold and water stress in leaves. *Journal of agricultural and food chemistry* **52**, 1950-1957.

Satoh, S., Kamada, H., Harada, H., et Fujii, T. (1986). Auxin-controlled glycoprotein release into the medium of embryogenic carrot cells. *Plant physiology* **81**, 931-933.

Schauer, S. E., Jacobsen, S. E., Meinke, D. W., et Ray, A. (2002). DICER-LIKE1 : blind men and elephants in *Arabidopsis* development. *Trends in plant science* **7**, 487-491.

Schlögl, P. S., dos Santos, A. L. W., do Nascimento Vieira, L., Floh, E. I. S., et Guerra, M. P. (2012). Gene expression during early somatic embryogenesis in Brazilian pine (Araucaria angustifolia (Bert) O. Ktze). *Plant Cell, Tissue and Organ Culture (PCTOC)* **108**, 173-180.

Schmidt, E. D., Guzzo, F., Toonen, M. A., et de Vries, S. C. (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* **124**, 2049-2062.

Schwarzerova, K., Vondrakova, Z., Fischer, L., Borikova, P., Bellinvia, E., Eliasova, K., Havelkova, L., Jindriska, F., Vagner, M., et Opatrny, Z. (2010). The role of actin isoforms in somatic embryogenesis in Norway spruce. *BMC plant biology* **10**, 89.

Sghaier-Hammami, B., Drira, N., et Jorrin-Novo, J. V. (2009a). Comparative 2-DE proteomic analysis of date palm (*Phoenix dactylifera* L.) somatic and zygotic embryos. *J Proteomics* **73**, 161-77.

Sghaier-Hammami, B., Valledor, L., Drira, N., et Jorrin-Novo, J. V. (2009b). Proteomic analysis of the development and germination of date palm (*Phoenix dactylifera* L.) zygotic embryos. *Proteomics* **9**, 2543-2554.

Shah, K., Gadella Jr, T. W. J., van Erp, H., Hecht, V., et de Vries, S. C. (2001). Subcellular Localization and Oligomerization of the *Arabidopsis thaliana* Somatic Embryogenesis Receptor Kinase 1 Protein. *Journal of Molecular Biology* **309**, 641-655.

Shelbourne, C. (1997). Genetics of adding value to the end-products of radiata pine. *In* "IUFRO 97 Genetics of Radiata Pine, Proceedings of NZ FRI-IUFRO Conference 1-5 December" (M. J. Burdon RD, ed.), Vol. 203, pp. 129-141. FRI Bull, Rotorua, New Zealand.

Shi, J., Zhen, Y., et Zheng, R.-H. (2010). Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. *Journal of Experimental Botany* **61**, 2367-2381.

Shih, M.-D., Hoekstra, F. A., et Hsing, Y.-I. C. (2008). Late embryogenesis abundant proteins. *Advances in Botanical Research* **48**, 211-255.

Shotwell, M. A., et Larkins, B. A. (1989). The biochemistry and molecular biology of seed storage proteins. *The biochemistry of plants: a comprehensive treatise*.

Silveira, V., Balbuena, T. S., Santa-Catarina, C., Floh, E. I. S., Guerra, M. P., et Handro, W. (2004). Biochemical changes during seed development in *Pinus taeda* L. *Plant Growth Regulation* **44**, 147-156.

Singh, H. (1978). Embryology of gymnosperms. *In* "Handbuch der Pflanzenanatomie (Encyclopedia of Plant Anatomy)", Vol. 10, Part 2. Berlin, Germany Gebruder Borntraeger.

Skriver, K., et Mundy, J. (1990). Gene expression in response to abscisic acid and osmotic stress. *The Plant Cell* **2**, 503.

Stasolla, C., Belmonte, M. F., Zyl, L. v., Craig, D. L., Liu, W. B., Yeung, E. C., et Sederoff, R. R. (2004). The effect of reduced glutathione on morphology and gene expression of white spruce (*Picea glauca*) somatic embryos. *Journal of Experimental Botany* **55**, 695-709.

Stasolla, C., Loukanina, N., Ashihara, H., Yeung, E. C., et Thorpe, T. A. (2001). Purine and pyrimidine metabolism during the partial drying treatment of white spruce (*Picea glauca*) somatic embryos. *Physiologia Plantarum* **111**, 93-101.

Stasolla, C., et Yeung, E. C. (2003). Recent advances in conifer somatic embryogenesis: improving somatic embryo quality. *Plant Cell, Tissue and Organ Culture* **74**, 15-35 (Revue).

Steiner, N., Santa-Catarina, C., Guerra, M. P., Cutri, L., Dornelas, M. C., et Floh, E. I. S. (2012). A gymnosperm homolog of SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE-1 (*SERK1*) is expressed during somatic embryogenesis. *Plant Cell, Tissue and Organ Culture* **109**, 41-50.

Sterk, P., Booij, H., Schellekens, G. A., Van Kammen, A., et de Vries, S. C. (1991). Cell-specific expression of the carrot EP2 lipid transfer protein gene. *The Plant Cell Online* **3**, 907-21.

Steward, F. C., Marion, O., Smith, M., et Smith, J. (1958). Growth and Organized Development of Cultured Cells. I. Growth and Division of Freely Suspended Cells. *American Journal of Botany* **45**, 693-703.

Stone, S. L., Braybrook, S. A., Paula, S. L., Kwong, L. W., Meuser, J., Pelletier, J., Hsieh, T.-F., Fischer, R. L., Goldberg, R. B., et Harada, J. J. (2008). *Arabidopsis* LEAFY COTYLEDON2 induces maturation traits and auxin activity: implications for somatic embryogenesis. *Proceedings of the National Academy of Sciences* **105**, 3151-3156.

Stone, S. L., Kwong, L. W., Yee, K. M., Pelletier, J., Lepiniec, L. c., Fischer, R. L., Goldberg, R. B., et Harada, J. J. (2001). LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proceedings of the National Academy of Sciences* **98**, 11806-11811.

Stroeher, V. L., Maclagan, J. L., et Good, A. G. (1997). Molecular cloning of a *Brassica napus* cysteine protease gene inducible by drought and low temperature stress. *Physiologia Plantarum* **101**, 389-397.

Suarez, M. F., Filonova, L. H., Smertenko, A., Savenkov, E. I., Clapham, D. H., Arnold, S. v., Zhivotovsky, B., et Bozhkov, P. V. (2004). Metacaspase-dependent programmed cell death is essential for plant embryogenesis. *Current Biology* **14**, R339-R340.

Sumner, L. W., Mendes, P., et Dixon, R. A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* **62**, 817-836.

Sunderlikova, V., Salaj, J., Matusikova, I., et Wilhelm, E. (2009). Isolation and characterization of an embryo-specific *Em*-like gene of pedunculate oak (*Quercus robur* L.) and its temporal and spatial expression patterns during somatic and zygotic embryo development. *Trees* **23**, 135-144.

Sung, D. Y., Kaplan, F., et Guy, C. L. (2001). Plant Hsp70 molecular chaperones: protein structure, gene family, expression and function. *Physiologia plantarum* **113**, 443-451.

Sutton, B. (2002). Commercial delivery of genetic improvement to conifer plantations using somatic embryogenesis. *Ann. For. Sci.* **59**, 657-661.

— T —

Tahir, M., Law, D. A., et Stasolla, C. (2006). Molecular characterization of *PgAGO*, a novel conifer gene of the ARGONAUTE family expressed in apical cells and required for somatic embryo development in spruce. *Tree physiology* **26**, 1257-1270.

Tang, G.-Q., Lüscher, M., et Sturm, A. (1999). Antisense Repression of Vacuolar and Cell Wall Invertase in Transgenic Carrot Alters Early Plant Development and Sucrose Partitioning. *The Plant Cell Online* **11**, 177-190.

Tautorus, T. E., Fowke, L. C., et Dunstan, D. I. (1991). Somatic embryogenesis in conifers. *Canadian Journal of Botany* **69**, 1873-1899 (Revue).

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

Terskikh, V. V., Feurtado, J. A., Borchardt, S., Giblin, M., Abrams, S. R., et Kermode, A. R. (2005). In vivo13C NMR metabolite profiling: potential for understanding and assessing conifer seed quality. *Journal of experimental botany* **56**, 2253-2265.

Teyssier, C., Grondin, C., Bonhomme, L., Lomenech, A.-M., Vallance, M., Morabito, D., Label, P., et Lelu-Walter, M.-A. (2011). Increased gelling agent concentration promotes somatic embryo maturation in hybrid larch (*Larix×eurolepsis*): a 2-DE proteomic analysis. *Physiologia Plantarum* **141**, 152-165.

Teyssier, C., Maury, S., Beaufour, M., Grondin, C., Delaunay, A., Le Metté, C., Ader, K., Cadene, M., Label, P., et Lelu-Walter, M.-A. (2014). In search of markers for somatic embryo maturation in hybrid larch (*Larix— eurolepis*): global DNA methylation and proteomic analyses. *Physiologia Plantarum* **150**, 271-291.

Thibaud-Nissen, F., Shealy, R. T., Khanna, A., et Vodkin, L. O. (2003). Clustering of Microarray Data Reveals Transcript Patterns Associated with Somatic Embryogenesis in Soybean. *Plant Physiology* **132**, 118-136.

Thivolle-Cazat, A., Colin, A., Meredie, C., Drouineau, S., Cavaignac, S., Bailly, A., Labbe, T., et Lambert, P. (2013). Analyse prospective de la ressource forestière et des disponibilités en bois de la région Aquitaine à l'horizon 2025. *IGF - Rapport final - juillet 2013*.

Thoma, S., Hecht, U., Kippers, A., Botella, J., de Vries, S., et Somerville, C. (1994). Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from *Arabidopsis*. *Plant Physiology* **105**, 35-45.

Thomashow, M. F. (1999). Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annual review of plant biology* **50**, 571-599 (Revue).

Thompson, E. W., et Lane, B. G. (1980). Relation of protein synthesis in imbibing wheat embryos to the cell-free translational capacities of bulk mRNA from dry and imbibing embryos. *Journal of Biological Chemistry* **255**, 5965-5970.

Trontin, J. F., Canlet, F., Reymond, I., Debille, S., Durandeau, K., Harvengt, L., Rousseau, J., de Boisseson, J., Fraysse, J., Alazard, P., Bailly, A., Teyssier, C., Le Metté, C., Morel, A., Label, P., et Lelu-Walter, M. (2013). L'embryogenèse somatique : une méthode de multiplication végétative du pin maritime pour demain ? *FCBA - INFO - Août 2013*.

Tunnacliffe, A., et Wise, M. J. (2007). The continuing conundrum of the LEA proteins. *Naturwissenschaften* **94**, 791-812.

Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T. C., McElver, J., Aux, G., et Patton, D. (2004). Identification of genes required for embryo development in *Arabidopsis*. *Plant Physiology* **135**, 1206-1220.

— U —

Uddenberg, D., Valladares, S., Abrahamsson, M., Sundstrom, J., Sundas-Larsson, A., et von Arnold, S. (2011). Embryogenic potential and expression of embryogenesis-related genes in conifers are affected by treatment with a histone deacetylase inhibitor. *In* "BMC Proceedings", Vol. 5, pp. P151. BioMed Central Ltd.

Uesono, Y., et Toh-e, A. (2002). Transient inhibition of translation initiation by osmotic stress. *Journal of Biological Chemistry* **277**, 13848-13855.

Unlu, M., Morgan, M. E., et Minden, J. S. (1997). Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071-2077.

— V —

Vales, T., Feng, X., Ge, L., Xu, N., Cairney, J., Pullman, G. S., et Peter, G. F. (2007). Improved somatic embryo maturation in loblolly pine by monitoring ABA-responsive gene expression. *Plant cell reports* **26**, 133-143.

Valledor, L., Castillejo, M. A., Lenz, C., Rodriguez, R., Canal, M. J., et Jorrin, J. (2008). Proteomic analysis of *Pinus radiata* needles: 2-DE map and protein identification by LC/MS/MS and substitution-tolerant database searching. *Journal of Proteome Research* **7**, 2616-2631.

van Engelen, F. A., et de Vries, S. C. (1992). Extracellular proteins in plant embryogenesis. *Trends in Genetics* **8**, 66-70.

van Hengel, A. J., Guzzo, F., van Kammen, A., et de Vries, S. C. (1998). Expression Pattern of the Carrot EP3 Endochitinase Genes in Suspension Cultures and in Developing Seeds. *Plant Physiology* **117**, 43-53.

van Hengel, A. J., van Kammen, A., et de Vries, S. C. (2002). A relationship between seed development, Arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis. *Physiologia Plantarum* **114**, 637-644.

Vestman, D., Larsson, E., Uddenberg, D., Cairney, J., Clapham, D., Sundberg, E., et von Arnold, S. (2011). Important processes during differentiation and early development of somatic embryos of Norway spruce as revealed by changes in global gene expression. *In* "BMC Proceedings", Vol. 5, pp. P78. BioMed Central Ltd.

von Aderkas, P., Bonga, J., Klimaszewska, K., et Owens, J. (1991). Comparison of Larch Embryogeny *In Vivo* and *In Vitro*. *In* "Woody Plant Biotechnology", Vol. 210, pp. 139-155. Springer US.

von Aderkas, P., et Bonga, J. M. (2000). Influencing micropropagation and somatic embryogenesis in mature trees by manipulation of phase change, stress and culture environment. *Tree Physiology* **20**, 921-928 (Revue).

von Aderkas, P., Lelu, M. A., et Label, P. (2001). Plant growth regulator levels during maturation of larch somatic embryos. *Plant Physiology and Biochemistry* **39**, 495-502.

von Arnold, S., Bozhkov, P., Clapham, D., Dyachok, J., Filonova, L., Högberg, K.-A., Ingouff, M., et Wiweger, M. (2005). Propagation of Norway spruce via somatic embryogenesis. *Plant Cell, Tissue and Organ Culture* **81**, 323-329.

von Arnold, S., Sabala, I., Bozhkov, P., Dyachok, J., et Filonova, L. (2002). Developmental pathways of somatic embryogenesis. *Plant Cell, Tissue and Organ Culture* **69**, 233-249.

— W —

Wang, W., Vinocur, B., Shoseyov, O., et Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in plant science* **9**, 244-252.

Washburn, M. P., Wolters, D., et Yates, J. R. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature biotechnology* **19**, 242-247.

Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., et Humphery-Smith, I. (1995). Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* **16**, 1090-1094.

Weckwerth, W. (2003). Metabolomics in systems biology. *Annual review of plant biology* **54**, 669-689.

Winter, H., et Huber, S. C. (2000). Regulation of Sucrose Metabolism in Higher Plants : Localization and Regulation of Activity of Key Enzymes. *Critical Reviews in Biochemistry and Molecular Biology* **35**, 253-289 (Revue).

Wise, M. J. (2003). LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *Bmc Bioinformatics* **4**, 52.

Wise, M. J., et Tunnacliffe, A. (2004). POPP the question: what *do* LEA proteins do? *Trends in plant science* **9**, 13-17.

Wiweger, M., Farbos, I., Ingouff, M., Lagercrantz, U., et Von Arnold, S. (2003). Expression of *Chia4-Pa* chitinase genes during somatic and zygotic embryo development in Norway spruce (*Picea abies*): similarities and differences between gymnosperm and angiosperm class IV chitinases. *Journal of experimental botany* **54**, 2691-2699.

Wu, X. L., Chory, J., et Weigel, D. (2007). Combinations of *WOX* activities regulate tissue proliferation during *Arabidopsis* embryonic development. *Developmental Biology* **309**, 306-316.

— Y —

Yan, G., Xi, M., Wang, G., Yang, L., et Shi, J. (2010). Molecular characteristics and expression analysis of *PmSERK1* during somatic embryogenesis in Masson pine. *Molecular Plant Breeding* **8**, 53-58.

Yang, X., et Zhang, X. (2010). Regulation of Somatic Embryogenesis in Higher Plants. *Critical Reviews in Plant Sciences* **29**, 36-57 (Revue).

— Z —

Zavattieri, M. A., Frederico, A. M., Lima, M., Sabino, R., et Arnholdt-Schmitt, B. (2010). Induction of somatic embryogenesis as an example of stress-related plant reactions. *Electronic Journal of Biotechnology* **13**, 1-9 (Revue).

Zeevaart, J. A. D., et Creelman, R. A. (1988). Metabolism and physiology of abscisic acid. *Annual review of plant physiology and plant molecular biology* **39**, 439-473.

Zeng, F., Zhang, X., Zhu, L., Tu, L., Guo, X., et Nie, Y. (2006). Isolation and Characterization of Genes Associated to Cotton Somatic Embryogenesis by Suppression Subtractive Hybridization and Macroarray. *Plant Molecular Biology* **60**, 167-183.

Zhang, J., Zhang, S., Han, S., Wu, T., Li, X., Li, W., et Qi, L. (2012). Genome-wide identification of microRNAs in larch and stage-specific modulation of 11 conserved microRNAs and their targets during somatic embryogenesis. *Planta* **236**, 647-657.

Zhen, Y., Zhao, Z.-Z., Zheng, R.-H., et Shi, J. (2012). Proteomic analysis of early seed development in *Pinus massoniana* L. *Plant Physiology and Biochemistry* **54**, 97-104.

Zimmerman, J. L. (1993). Somatic embryogenesis: a model for early development in higher plants. *Plant Cell* **5**, 1411-1423 (Revue).

Zivy, M., et de Vienne, D. (2000). Proteomics: a link between genomics, genetics and physiology. *Plant molecular biology* **44**, 575-580.

Zuo, J., Niu, Q.-W., Frugis, G., et Chua, N.-H. (2002). The WUSCHEL gene promotes vegetative-to-embryonic transition in *Arabidopsis*. *The Plant Journal* **30**, 349-359.

Annexes

1. Somatic embryo maturation in maritime pine (*Pinus pinaster*): contribution of a 2-DE proteomic analysis for a better understanding

Extended abstract

Morel A, Teyssier C, Beaufour M, Reymond I, Le Metté C, Cadene M, Harvengt L, Trontin JF, Label P, Lelu-Walter MA 2013. Somatic embryo maturation in maritime pine (Pinus pinaster): contribution of a 2-DE proteomic analysis for a better understanding. Dans: Park, Y.S. and Bonga, J.M. (eds). Proceedings of the IUFRO Working Party 2.09.02 conference on "Integrating vegetative propagation, biotechnologies and genetic improvement for tree production and sustainable forest management" Juin 25-28, 2012, Brno Czech Republic. Published online:http://www.iufro20902.org/ ; pp166-169.

2. *De novo* assembly of maritime pine transcriptome: implications for forest breeding and biotechnology

Publication avec comité de lecture

Canales, Javier; Bautista, Rocio; Label, Philippe; Gomez-Maldonado, Josefa; Lesur, Isabelle; Fernández-Pozo, Noe; Rueda-López, Marina; Guerrero-fernández, Dario; Castro-Rodriguez, Vanessa; Benzekri, Hicham; Canas, Rafael; Guevara, Maria-Angeles; Rodrigues, Andreia; Teyssier, Caroline; **Morel Alexandre**; Ehrenmann, François; Le Provost, Gregoire; Lalanne, Celine; Noirot, Celine; Klopp, Christophe; Garcia-Gutierrez, Angel; Trontin, Jean-François; Lelu-Walter, Marie-Anne; Miguel, Celia; Cervera, Maria Teresa; canton, Francisco; Plomion, Christophe; Harvengt, Luc; Avila, Concepcion; Claros, M. Gonzalo; Canovas, Francisco. *De novo* assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. Plant Biotechnology Journal (sous presse).

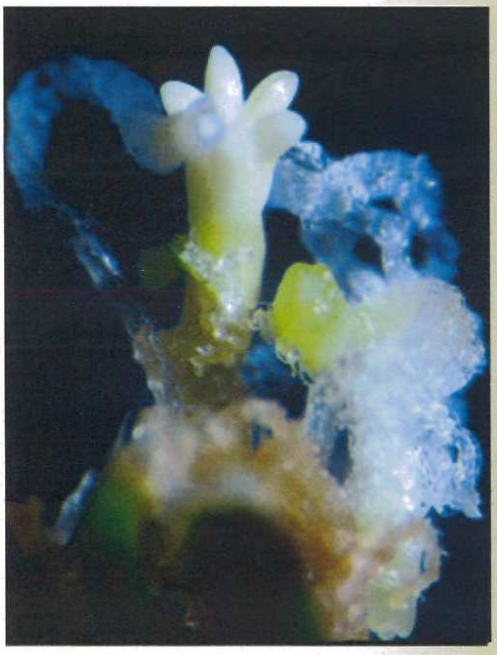
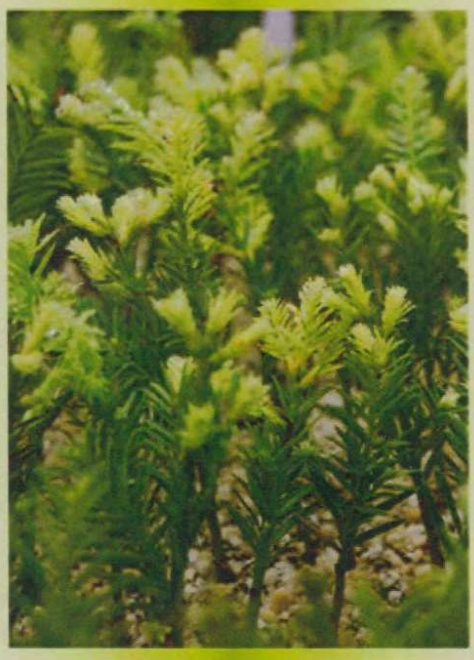
3. L'embryogenèse somatique : une méthode de multiplication végétative du pin maritime pour demain ?

Publication sans comité lecture

Trontin JF, Canlet F, Reymond I, Debille S, Durandeau K, Harvengt L, Rousseau JP, de Boissesson JM, Fraysse JY, Alazard P, Bailly A, Teyssier C, Le Metté C, **Morel A,** Label P, Lelu-Walter MA 2013. L'embryogenèse somatique : une méthode de multiplication végétative du pin maritime pour demain ? FCBA Info, septembre 2013, 9 pp. Publié en ligne le 05/09/2013 (http://www.fcbainfo.fr).

IUFRO

Integrating vegetative propagation, biotechnologies and genetic improvement for tree production and sustainable forest management


Proceedings

Second International Conference of the IUFRO Working Party 2.09.02

June 25 – 28, 2012 • Brno, Czech Republic

IUFRO Working Party 2.09.02:
Somatic Embryogenesis and Other Vegetative Propagation Technologies

Conference

PROCEEDINGS

Integrating vegetative propagation, biotechnologies and genetic improvement for tree production and sustainable forest management

June 25-28, 2012
Brno, Czech Republic

COMMITTEES

Organizing committee

YILL-SUNG PARK, Natural Resources Canada – Canadian Forest Service, Fredericton, New Brunswick, Canada. *ypark@nrcan.gc.ca*

JANA KRAJŇÁKOVÁ, Mendel University in Brno, Faculty of Forestry and Wood Technology, Brno, Czech Republic. *jana.krajnakova@uniud.it*

MARIANO TORIBIO, Instituto Madrileño de Investigación y Desarrollo Rural, Agrario y Alimentario (IMIDRA), Alcalá de Henares (Madrid), Spain. *mariano.toribio@madrid.org*

JEAN-FRANÇOIS TRONTIN, FCBA Technological Institute, Biotechnology and Advanced Forestry department, Biotechnology Lab, Bordeaux, France. *jean-francois.trontin@fcba.fr*

SO-YOUNG PARK, Korea Forest Research Institute (KFRI), Department of Forest Genetic Resources, Suwon, Korea. *soypark7@korea.kr*

Conference Executive Committee and Local Arrangements

PETR HORÁČEK (Dean, Faculty of Forestry and Wood Science, Mendel University)

LIBOR JANKOVSKÝ (Vice Dean, Faculty of Forestry and Wood Science, Mendel University)

JANA KRAJŇÁKOVÁ (visiting scientist, Mendel University)

BŘETISLAV BRZOBOHATÝ (Head, Molecular Biology and Radiobiology, Mendel University)

VILÉM REINÖHL (Project Manger, Central European Institute of Technology, Mendel University)

MARTIN VÁGNER (Institute Experimental Botany (Prague), Academy Science Czech Rep)

JANA LEVOVÁ (event manager, Barceló Brno Palace Hotel)

MARTINA ŠATINSKÁ (The South Moravian Centre for International Mobility)

Scientific Committee

ANTONIO BALLESTER (Spain)
JAN BONGA (Canada)
JORGE CANHOTO (Portugal)
CARMEN DIAZ-SALA (Spain)
DON DURZAN (USA)
YOUSRY EL-KASSABY (Canada)
PRAMOD GUPTA (USA)
HELY HÄGMAN (Finland)
CATHY HARGREAVES (New Zealand)
LADISLAV HAVEL (Czech Republic)
PETR HORÁČEK (Czech Republic)
KRYSTYNA KLIMASZEWSKA (Canada)
MARIE-ANNE LELU-WALTER (France)
MILAN LSTIBŮREK (Czech Republic)
JANA MALÁ (Czech Republic)
SCOTT MERKLE (USA)
OLIVIER MONTEUUIS (France)
HEUNG-KYU MOON (Korea)
TERÉZIA SALAJ (Slovak Republic)
ANA VIEITEZ (Spain)
PATRICK VON ADERKAS (Canada)
SARA VON ARNOLD (Sweden)

Preface

Vegetative propagation of woody plants by *in vitro* culture has been attempted since the1930s; however, it was not until about 20 years later that the first reports appeared of successful plantlet formation in some hardwood species. With conifers this was not achieved until 1975. Since that time, steady progress has been made in propagating woody plants using tissues, organs, cells and protoplasts, which finally led to the discovery of somatic embryogenesis (SE) in conifers in 1985. The successful SE and other clonal propagation methods have, for some tree species, arrived at the point where they can be implemented industrially. This success in vegetative propagation technologies with trees offers new opportunities for tree breeding, genetic conservation and restoration, and sustainable forest management. However, for many important tree species, recalcitrance is still an issue, requiring further research and development.

The IUFRO Working Party 2.09.02 on "Somatic Embryogenesis and Other Vegetative Propagation Technologies" was organized in 2008 with its main objective fostering the development and application of SE and other vegetative propagation technologies in woody plant propagation. The inaugural conference of this Working Party was held in Suwon, Korea in 2010, which brought leading scientists in the field together who enthusiastically embraced the development and the applications of SE and other vegetative propagation technologies. In June 2012, the second conference was hosted by the Mendel University in Brno, Czech Republic. The conference was attended by about 100 scientists and practitioners from the around the world and 38 oral and 54 poster presentations were made. A highlight of this conference was the recognition and awarding of the Mendel Medal (by the Mendel University) to the scientists who contributed significantly to the development of vegetative propagation technology. The honourees were: Drs. Antonio Ballester (Spain), Jan Bonga (Canada), Vladimir Chalupa (Czech Republic), Don Durzan (USA), Inger Hakman (Sweden), Pramod Gupta (USA), Ana Vieitez (Spain), and Sara von Arnold (Sweden). We certainly appreciate their pioneering contributions.

In addition to the scientific program, the participants enjoyed field tours to the Botanical garden and greenhouse in Brno, the Arboretum in Krtiny, caves in the Moravian Karst, a visit to the Austerlitz battle field memorial and the Forestry and Game Management Research Institute's Kunovice Station. As well, the participants enjoyed the social events at Cerna Hora and Uherske Hradiste!

These proceedings contain the papers and abstracts that were submitted for this proceeding. Abstracts that were not submitted for editing are found in their original format in the "Book of Abstracts" published during the conference. We are grateful to all the contributors to the conference. We would like to acknowledge the contributions of the Executive and Scientific committees that have made this conference such a great experience. In particular, we are grateful to Drs. Petr Horacek (Dean, Faculty of Forestry and Wood Technology) and Jana Krajnakova for making this conference a huge success!

Fredericton, Canada

Yill Sung Park and Jan Bonga

CONTENTS

Full Article

Interpolated apomictic somatic embryogenesis, androsporogenesis, asexual heterospory, mitosporogenesis and genomic silencing in a gymnosperm artificial sporangium
Don J Durzan 3

Recalcitrance in the *in vitro* propagation of trees
Jan M. Bonga 37

Partial rejuvenation of mature hardwood trees through somatic embryogenesis: The example of pedunculate oak
Antonio Ballester, Ana M. Vieitez 47

Initiation of embryogenic suspensor masses (ESM) and somatic embryogenesis in Japanese red pine (*Pinus densiflora*)
Y.W. Kim, H.K. Moon, S.Y. Park 56

A combined method to increase somatic embryogenesis efficiency in valuable cell lines of *Pinus* spp.
I.A. Montalbán, P. Moncaleán 63

The Embryogenic Lines and Somatic Embryogenesis of Coniferous Species in Siberia
I. Tretiakova, E. Voroshilova, A. Ivanitskya, D. Shuvaev, M. Park 73

Towards scaling-up the micropropagation of *Juglans major* (Torrey) Heller var. 209 x *J. regia* L., a hybrid walnut of commercial interest
Ricardo Julián Licea Moreno, Ana Valeria Morales, Marcos Daquinta Gradaille, Luis Gómez 80

Building the "pipeline": Applying somatic embryogenesis, bioreactors and transgenic technology to restore the American chestnut
Scott A. Merkle, Lisheng Kong, Ryan A. Tull, Campbell J. Nairn 92

New ornamental conifers for harsh northern conditions through cutting propagation of special forms of Norway spruce
Teijo Nikkanen, Susanne Heiska, Tuija Aronen 98

Second International Conference of the IUFRO Working Party 2.09.02
June 25-28, 2012 • Brno, Czech Republic

Effect of auxins on the induction of somatic embryos from immature zygotic embryonic axes of *Ocotea porosa* (Nees ex Mart.) Liberato Barroso
Marguerite Quoirin, Luciana L. Pelegrini, Luciana L. F. Ribas 110

Effect of picloram and silver nitrate on embryogenic callus induction in *Acrocomia aculeata* (Jacq.) Lodd. ex Mart. by thin cell layer culture
Marguerite Quoirin, João Henrique Delfrate Padilha, Luciana Lopes Fortes Ribas 116

Somatic Embryogenesis in Brazilian pine: establishment of biochemical markers for selection of cell lines with high embryogenic potential
André Luis Wendt dos Santos, Leonardo Jo, Claudete Santa-Catarina, Miguel Pedro Guerra, Eny Iochevet Segal Floh 123

The prospects for using somatic embryogenesis to propagate Sitka spruce in the UK
Trevor Fenning, Yill-Sung Park 129

Extended Abstract

Patterning during somatic embryogenesis in conifers
Sara von Arnold, Daniel Uddenberg, Emma Larsson 141

Somatic embryogenesis induction on male catkins of Holm oak (*Quercus ilex* L.)
Miquel Blasco, Azahara Barra, Carmen Brisa, Juan Segura, Mariano Toribio, Isabel Arrillaga 143

Scaled-up egg suspension cultures of *Taxus brevifolia* biosynthesize taxol in an artificial sporangium
Don J Durzan 146

Towards mass-propagation of Norway spruce in Finland
Susanne Heiska, Laura Jaakola, Saila Varis, and Tuija Aronen 148

SE propagation and genetic diversity – example from a practical case
Karl-Anders Högberg 151

In vitro propagation of *Uncaria rhynchophylla* – a medicinal woody plant
Katsuaki Ishii, Naoki Takata, Toru Taniguchi 153

Somatic embryogenesis of black saxaul (*Haloxylon aphyllum*), a plant for saline soil reclamation of the dry aral seabed
V.Yu. Kirillov, B.M. Mukanov, T.N. Stikhareva, A.U.Manabayeva, N.K.Chebotko, N.B. Kazangapova 155

Secondary phenolic compounds in somatic embryogenesis of *Pinus sylvestris* L.: a preliminary study
Jinrong Lu, Katri Kostiainen[2], Laura Jaakola, Susanne Heiska, Anni Harju, Riitta Julkunen-Tiitto, Martti Venäläinen, Tuija Aronen 160

Somatic embryogenesis from Spanish provenances of maritime pine
Angeles Morcillo, Miquel Blasco, Laura Lorenzo, Vicente Almazán, Carmen Brisa, Juan Segura, Isabel Arrillaga 163

Somatic embryo maturation in maritime pine (*Pinus pinaster*): contribution of a 2-DE proteomic analysis for a better understanding
A. Morel, C. Teyssier, M. Beaufour, I. Reymond, C. Le Metté, M. Caden, L. Harvengt, J.-F. Trontin, P. Label, M.A. Lelu Walter 166

Native state of extracellular matrix of early conifer embryogenic tissue imaged by environmental scanning electron microscope
V. Neděla, J. Hřib , B. Vooková, J. Runštuk , S. Svidenská 170

Incorporating wood quality traits in multi-varietal forestry of white spruce
Yill-Sung Park, Yuhui Weng, Shawn Mansfield 173

The biochemical characteristics of the physiological activity of beech and spruce embryos
B. Pešek, K. Eliášová, M. Vágner, Z. Vondráková 175

Micropropagation of a recalcitrant pine (*Pinus pinea* L.): An overview of the effects of ectomycorrhizal inoculation
C. Ragonezi, A. T. Caldeira, M. do Rosário Martins, D. Teixeira, L. S. Dias, O. Miralto, E. Ganhão, K. Klimaszewska, A. Zavattieri 180

Somatic embryogenesis as an effective regeneration support for reverse genetics in maritime pine: the Sustainpine collaborative project as a case study
Jean-François Trontin, Sandrine Debille, Francis Canlet, Luc Harvengt, Marie-Anne Lelu-Walter, Philippe Label, Caroline Teyssier, Marie-Claude Lesage-Descauses, Claire Le Metté, Célia Miguel, José de Vega-Bartol, Mariagrazia Tonelli, Raissa Santos, Andrea Rupps, Seyedeh Batool Hassani, Kurt Zoglauer, Elena Carneros, Carmen Díaz-Sala, Dolores Abarca, Isabel Arrillaga, Isabel Mendoza-Poudereux, Juan Segura, Concepción Avila, Marina Rueda, Javier Canales, Francisco M. Cánovas 184

Short Abstract

Culturing nucellus as a way to produce proteins involved in conifer reproduction
Patrick von Aderkas, Andrea Coulter, Darryl Hardie, Natalie Prior 191

Somatic embryogenesis from clonal sources of *E. camaldulensis*
Mara Dekel, Yaarit Wainberg, Ziv Shani 191

Analysis of different promoters and reporter genes in somatic embryos of *Pinus pinaster* Ait. and *Larix decidua* Mill.
S.B. Hassani, T. Benneckenstein, A. Rupps, G. Hensel, S. Broeders, J.-F. Trontin, K. Zoglauer 192

From angiosperm models to forest trees: A study on expression behavior of *Arabidopsis* homologous genes during somatic embryogenesis of *Larix decidua*
A. Rupps, J. Raschke, K. Zoglauer 193

Propagation of American chestnut using non-germinable somatic embryos
Lisheng Kong, Scott A. Merkle, Campbell J. Nairn 194

Climatic adaptation in Norway spruce - Molecular dissection of a novel epigenetic memory mechanism
Y.K. Lee, I. Yakovlev, C.G. Fossdal, Ø. Johnsen, J.E. Olsen 194

The role of polyamines in the embryogenic and organogenic capacity of yellow polar (*Liriodendron tulipifera*)
So-Young Park, Heung-Kyu Moon, Yong-Wook Kim 195

Somatic embryogenesis in *Abies nordmanniana*: present status and future application
Kurt Zoglauer, Claudia Aurich, Peter Uehre, Sepp Herrmann 196

Photo Gallery
. 197

Conference Participants
. 205

Proceedings of the IUFRO Working Party 2.09.02 conference
"Integrating vegetative propagation, biotechnologies and genetic improvement for tree production and sustainable forest management"
June 25-28, 2012. Brno, Czech Republic

Somatic embryo maturation in maritime pine (*Pinus pinaster*): contribution of a 2-DE proteomic analysis for a better understanding

A. Morel[1], C. Teyssier[1], M. Beaufour[2], I. Reymond[3], C. Le Metté[1], M. Caden[2], L. Harvengt[3], J.-F. Trontin[3], P. Label[1], M.A. Lelu-Walter [1*]

[1]: INRA, UR 588, Unité Amélioration, Génétique et Physiologie Forestières, 2163 Av. de la Pomme de pin, CS 4001 Ardon, 45075 Orléans cedex 2, France.
[2] : CNRS, UPR4301, Centre de Biophysique Moléculaire, Rue Charles Sadron, 45071 Orléans cedex 2, France.
[3]: FCBA, Biotechnology and advanced Forestry Department, Genetics & Biotechnology group, 71, Route d'Arcachon, Pierroton, 33610 Cestas, France.
**Corresponding author: marie-anne.lelu-walter@orleans.inra.fr*

Keywords: maritime pine, somatic embryogenesis, proteomic, maturation

Abstract: Somatic embryogenesis has reached application stage for *Picea* and *Larix* species; however it remains to be optimized for pine species. Thus maritime pine maturation needs improvement. To optimize the maturation step, it is necessary to develop markers that can be used to verify or monitor the quality of somatic embryos (SE). Therefore, a proteomic study was performed of two maturation stages to identify protein markers. There were about 100 significantly and differentially expressed proteins (Student's test, $p<0.001$). They were mainly involved in carbohydrate or lipid metabolism and genetic information processing. In addition, many storage proteins were identified (vicillin-like, legumin-like, LEA proteins); some of them were, surprisingly, observed from the beginning of maturation. Our ultimate goal is to have a better understanding of SE maturation in *Pinus pinaster*. It is expected that this information will help to optimize the process and *in vitro* plant production.

Introduction

Maritime pine is extensively used in south-western Europe for lignocellulose production. Increasing demand for wood from this pine species requires efficient production of improved varieties and this is currently achieved only through seed orchard management. As in most long-lived tree species, variety design and deployment would greatly benefit from setting up an efficient vegetative propagation system of selected seed resources. Somatic embryogenesis is a promising technology for most conifers (Klimaszewska et al., 2007) and significant progress has been made in recent years for maritime pine (Lelu-Walter et al., 2006). However, whether « true » maturity of harvested cotyledonary SE has been achieved is still a debatable question. Cotyledonary SE are currently selected at the end of the maturation phase based on morphological criteria. Embryo maturation is a crucial step, giving rise to the establishment of reserve compounds including specific storage proteins that will ensure optimal germination and subsequent *in situ* plantlet development. A 2D-PAGE proteomic approach has been developed at INRA of Orléans (Teyssier et al., 2011) to define the protein status of developing SE. We aim to achieve a better understanding of the processes acting during the maturation phase of SE.

As a first step towards this goal this work presents our first results in proteomic analysis during SE development. This new knowledge is expected to refine our current methodology for practical application of somatic embryogenesis in maritime pine.

Materials and methods

Experiments were conducted with the embryogenic line AAY06006 initiated in July 2006 by FCBA from immature zygotic embryos harvested from mother clone G1.2631 (Landes origin). Trees were pollinated with a mixture of pollen from Morocco genotypes. Maturation was performed according to Lelu-Walter et al., 2006 with

the following modifications. Basal maturation medium consisted in mLV basal medium (Litvay et al., 1985) that contained 9 gl^{-1} gellan gum (Phytagel™, SIGMA), 80 µM *cis-trans* (±) abscisic acid (ABA), and 0.2M sucrose. Cotyledonary SE were harvested after 1 (immature stage) or 12 weeks maturation (cotyledonary stage) for 2DE-proteomic analysis according to Teyssier et al. (2011). Briefly, total soluble proteins were extracted in liquid nitrogen from 400 mg ES (fresh mass) and precipitated with phenol. In the first dimensional separation (IEF), samples containing 450 µg protein were loaded and equilibrated onto 24-cm IPG strips, pH 4–7 (Protean IEF Cell system, Biorad, Marnes-La-Coquette, France). The second dimensional separation was performed in 2-D PAGE. Five biological replicates were analyzed for each sample. Gels were stained with colloidal CBB-G according to Gion et al., (2005), then images were scanned and analyzed using Progenesis software (Nonlinear Dynamics, Newcastle upon Tyne, United Kingdom). The volume of each spot detected was normalized relative to the total volume of the spots on the gel. Every spot automatically detected was manually checked. Selected spots were analyzed for protein identification by in-gel proteolysis followed by peptide extraction and nanoLC-MS/MS (nano high performance liquid chromatography on Dionex Ultimate coupled to tandem mass spectrometry on a Bruker Esquire HCT Ion Trap) and interrogation of *Pinus* TC databases.

Results and discussion

Statistical analysis and mass spectrometry identification

More than 1000 spots were reproducibly defined from 2D gels. Student's t test was performed on each of the normalized volume spots and more than 10% of all spots have shown a significant volume difference between the 2 tested stages ($P < 0.001$). Figure 1 shows their location on a 2D gel sample after 12 weeks of maturation. The proteins inside spots are defined according to their pI and MW.

The proteins were identified by in-gel proteolysis followed by nanoLC-MS/MS, using current TC DNA sequence databases. Whenever possible, the identified proteins were classified according to their biological function.

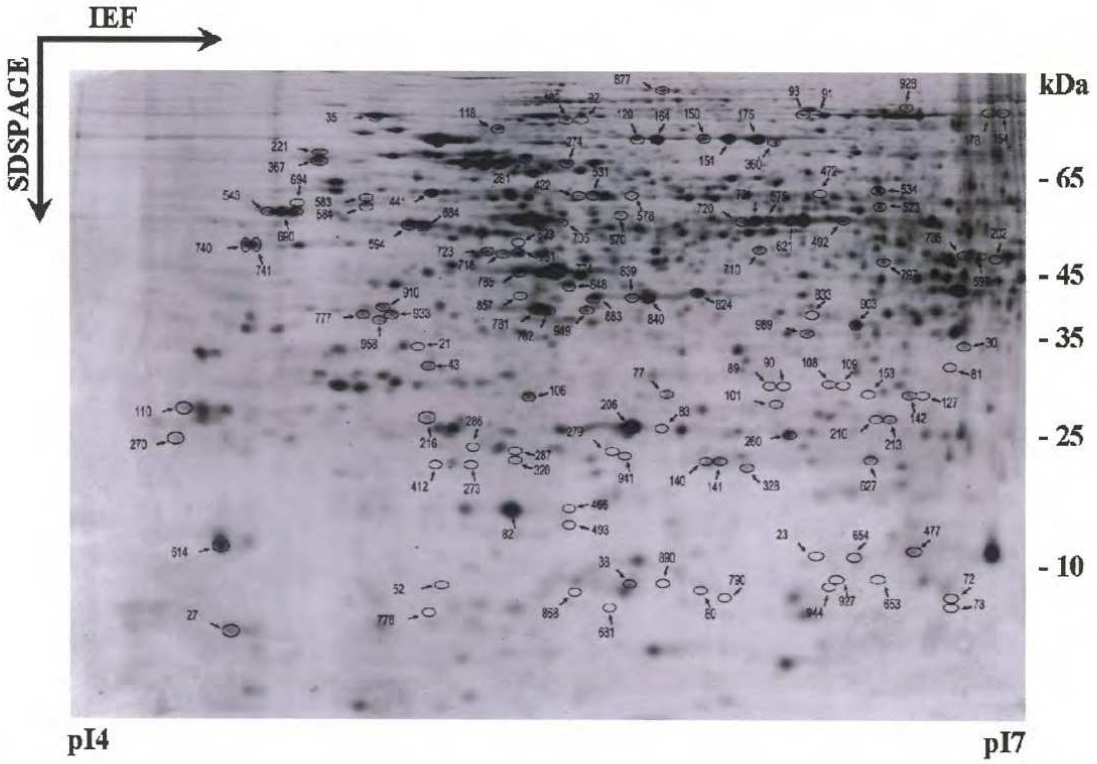


Figure 1. *Representative 2-DE map obtained for maritime pine SE after 12 weeks of maturation. Marked spots displayed significant differences in their abundance ($P < 0.001$) between 1 week (immature SE) and 12 weeks of maturation (cotyledonary SE).*

Biological interpretation

Based on the comparison of major functional classes (Fig. 2), we found that more proteins involved in metabolism are overexpressed in cotyledonary SE (12 weeks) than in immature ones (50% vs. 41% of total protein extract, respectively). Overexpressed proteins in cotyledonary SE are involved in carbon metabolism (20%), lipid synthesis (15%) and energy metabolism (15%); considering immature SE, over-expressed proteins are involved in

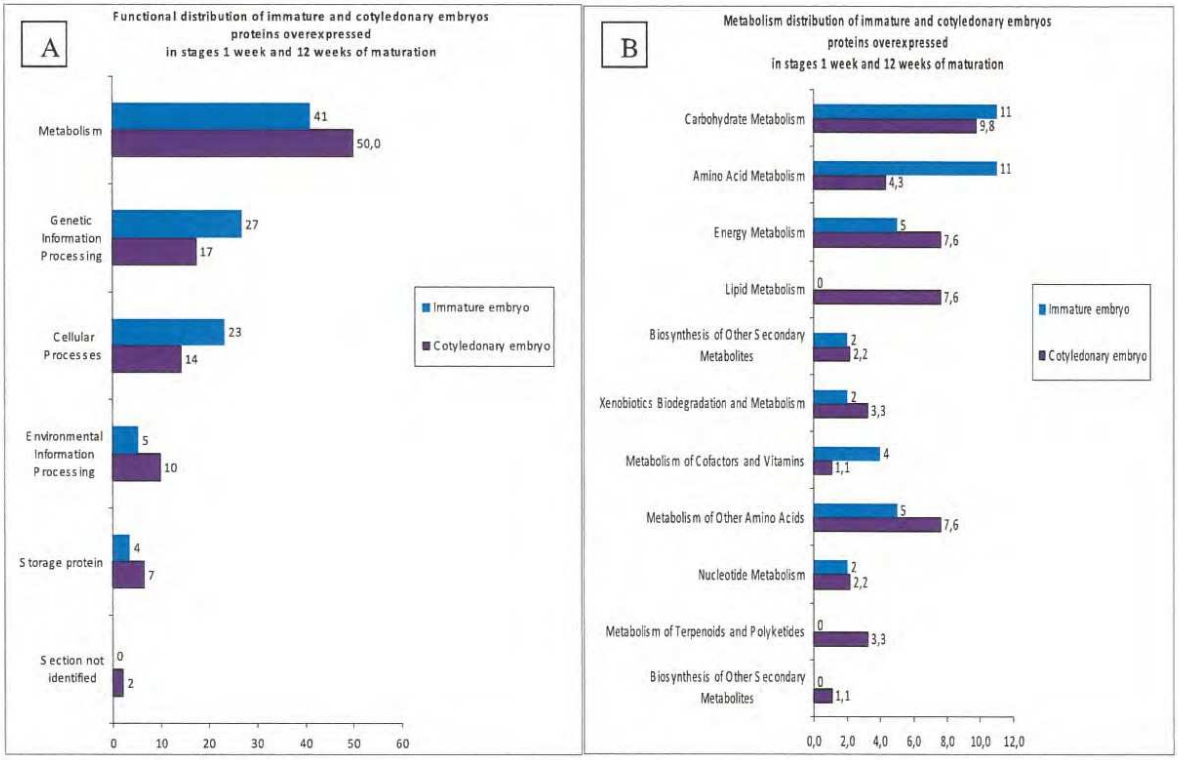


Figure 2. *Distributions of the identified putative proteins among functional class (A) and metabolism class (B) according to two culture stages (immature and cotyledonary SE). Functional classification catalog Kegg: Kyoto Encyclopedia of Genes and Genomes,* ***http://www.genome.jp/kegg/.*** *Numbers represent percentage of total proteins extracted for a stage of culture. Proteins involved in more than one process were assigned to more than one categorical group. Hence the sum of proteins in the categories exceeds the total number of proteins.*

amino acid synthesis (26%) and carbon metabolism (25%). Among the protein enzyme group, we found that ascorbate peroxidase involved in amino acid metabolism was significantly overexpressed (x 2.1) in immature SE when compared to cotyledonary SE. A similar observation was reported for Chinese fir (Shi et al., 2010) and soybean (Bailly et al., 2001) for zygotic embryos. Thus, ascorbate peroxidase could characterize the immature state of maritime pine SE. On the opposite, glutathione peroxidase, an enzyme involved in lipid metabolism, was 5.1 times more expressed in cotyledonary SE than in immature SE. An increased expression of this enzyme has been shown during SE development of *Eleutherococcus senticosus* (Shohael et al., 2007). We concluded that glutathione peroxidase is a good candidate marker of the cotyledonary state of maritime pine SE.

Conclusion and Perspectives

This preliminary work allowed us to identify putative protein markers involved in the main metabolic pathways such as energy and lipid metabolism. Metabolism of the amino acids seemed to be low in cotyledonary SE. Even if some markers have been identified, they could be related to embryogenesis and not to maturity. To answer this question it would be necessary to study the mature zygotic embryo by proteomic.

References

Bailly C, Audigier C, Ladonne F, Wagner MH, Coste F, Corbineau F, Côme D (2001). Changes in oligosaccharide content and antioxidant enzyme activities in developing bean seeds as related to acquisition of drying tolerance and seed quality. J Exp Bot 52: 701-708.

Gion, JM, Lalanne, C, Le Provost, G, Ferry-Dumazet, H, Paiva, J, Chaumeil, P, Frigerio, JM, Brach, J, Barré, A, de Daruvar, A, Claverol, S, Bonneu, M, Sommerer, N, Negroni, L, Plomion, C (2005). The proteome of maritime pine wood forming tissue. PROTEOMICS 5: 3731-3751.

Klimaszewska K, Trontin JF, Becwar M, Devillard C, Park YS, Lelu-Walter M-A (2007). Recent progress on somatic embryogenesis of four *Pinus* spp. Tree For Sci Biotech 1: 11-25.

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

Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2006). Simplified and improved somatic embryogenesis for clonal propagation of *Pinus pinaster* (Ait.). Plant Cell Rep 25: 767-76.

Shi J, Zhen Y, Zheng R-H (2010). Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. J Exp Bot 61: 2367-2381.

Shohael A, Ali M, Hahn E, Paek K (2007). Glutathione metabolism and antioxidant responses during *Eleutherococcus senticosus* somatic embryo development in a bioreactor. Plant Cell Tissue Organ Cult 89: 121-129.

Teyssier C, Grondin C, Bonhomme L, Lomenech A-M, Vallance M, Morabito D, Label P, Lelu-Walter M-A (2011). Increased gelling agent concentration promotes somatic embryo maturation in hybrid larch (*Larix×eurolepis*): a 2-DE proteomic analysis. Physiol Plant 141: 152-165.

Plant Biotechnology Journal

Plant Biotechnology Journal (2013), pp. 1–14 doi: 10.1111/pbi.12136

Research Article

De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology

Javier Canales[1†], Rocio Bautista[2†], Philippe Label[3†], Josefa Gómez-Maldonado[1], Isabelle Lesur[4,5,6], Noe Fernández-Pozo[2], Marina Rueda-López[1], Dario Guerrero-Fernández[2], Vanessa Castro-Rodríguez[1], Hicham Benzekri[2], Rafael A. Cañas[1], María-Angeles Guevara[7], Andreia Rodrigues[8], Pedro Seoane[2], Caroline Teyssier[9], Alexandre Morel[9], François Ehrenmann[4,5], Grégoire Le Provost[4,5], Céline Lalanne[4,5], Céline Noirot[10], Christophe Klopp[10], Isabelle Reymond[11], Angel García-Gutiérrez[1], Jean-François Trontin[11], Marie-Anne Lelu-Walter[9], Celia Miguel[8], María Teresa Cervera[7], Francisco R. Cantón[1], Christophe Plomion[4,5], Luc Harvengt[11], Concepción Avila[1,2], M. Gonzalo Claros[1,2] and Francisco M. Cánovas[1,2,*]

[1]*Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, Universidad de Málaga, Málaga, Spain*

[2]*Plataforma Andaluza de Bioinformática, Edificio de Bioinnovación, Parque Tecnológico de Andalucía, Málaga, Spain* [3]*INRA, Université Blaise Pascal, Aubière Cedex, France*

[4]*INRA, Cestas, France*

[5]*Université de Bordeaux, Talence, France*

[6]*HelixVenture, Mérignac, France*

[7]*Departamento de Ecología y Genética Forestal, INIA-CIFOR, Madrid, Spain*

[8]*Forest Biotech Lab, IBET/ITQB, Oeiras, Portugal*

[9]*INRA, Unité Amélioration, Génétique et Physiologie Forestières, Orléans Cedex 2, France*

[10]*INRA de Toulouse Midi-Pyrénées, Auzeville, Castanet Tolosan cedex, France*

[11]*FCBA, Pôle Biotechnologie et Sylviculture, Cestas, France*

Received 20 July 2013;
revised 24 September 2013;
accepted 26 September 2013.
*Correspondence (Tel: +34 952131942;
fax: +34 952132376;
email: canovas@uma.es)
†These authors contributed equally to work.

Keywords: conifers, transcriptome sequencing, next-generation sequencing, full-length cDNA, transcription factors, single nucleotide polymorphism.

Summary

Maritime pine (*Pinus pinaster* Ait.) is a widely distributed conifer species in Southwestern Europe and one of the most advanced models for conifer research. In the current work, comprehensive characterization of the maritime pine transcriptome was performed using a combination of two different next-generation sequencing platforms, 454 and Illumina. *De novo* assembly of the transcriptome provided a catalogue of 26 020 unique transcripts in maritime pine trees and a collection of 9641 full-length cDNAs. Quality of the transcriptome assembly was validated by RT-PCR amplification of selected transcripts for structural and regulatory genes. Transcription factors and enzyme-encoding transcripts were annotated. Furthermore, the available sequencing data permitted the identification of polymorphisms and the establishment of robust single nucleotide polymorphism (SNP) and simple-sequence repeat (SSR) databases for genotyping applications and integration of translational genomics in maritime pine breeding programmes. All our data are freely available at *SustainpineDB*, the *P. pinaster* expressional database. Results reported here on the maritime pine transcriptome represent a valuable resource for future basic and applied studies on this ecological and economically important pine species.

Introduction

Forests are essential components of the ecosystems covering approximately one-third of the Earth's land area and playing a fundamental role in the regulation of terrestrial carbon sinks. Trees represent nearly 80% of the plant biomass (Olson *et al.*, 1983) and 50%–60% of annual net primary production in terrestrial ecosystems (Field *et al.*, 1998).

Conifers are the most important group of gymnosperms. Having diverged from a common ancestor more than 300 million years ago (Bowe *et al.*, 2000), gymnosperms and angiosperms have evolved very efficient and distinct physiological adaptations (Leitch and Leitch, 2012). Coniferous forests dominate large ecosystems in the Northern Hemisphere and include a broad variety of woody plant species, some of which are the largest, tallest and longest living organisms on Earth (Farjon, 2010).

Conifer trees are also of great economic importance, as they are the primary source for timber and paper production worldwide. Total timber production in the European Union in 2011 was 427 million m^3 (UNECE, 2013). Approximately 22% was used to produce energy, while the rest was used to supply industrial demands. A study of United Nations Economic Commission for Europe/Food and Agriculture Organization (UNECE/FAO) point out that the future needs in forest biomass to meet the demands of industrial wood as an energy source will exceed production by 2020. The development of a more productive and sustainable forest plantation is essential to meet the increasing demand of wood worldwide together with minimizing environmental impacts (e.g. decreasing pressure on natural forests).

The extant conifers comprise 615 species classified into eight families within the division *Pinophyta* (Farjon, 2010). Some of the most important conifer trees such as pines, spruces and firs are included in the family *Pinaceae*. The genus *Pinus* comprises the largest number of diversified species (113). Maritime pine (*Pinus pinaster* Aiton) is a broadly planted species (4.2 million hectares) in the southwestern part of the Mediterranean Basin, especially along the Atlantic coast in France, Spain and Portugal where it is the dominant species on more than 2.3 million hectares (Sanz *et al.*, 2006). The maritime pine is particularly tolerant to abiotic stresses showing relatively high levels of intraspecific variability (Aranda *et al.*, 2010). The maritime pine is also one of the most genetically studied conifer species for genomic research in Europe (Mackay *et al.*, 2012; Neale and Kremer, 2011), and a large number of genomic resources and phenotypic data have been generated in the last few years and are available for the conifer research community (http://www.scbi.uma.es/sustainpine; https://w3.pierroton.inra.fr/PinusPortal). Furthermore, knowledge gained from studying this conifer species will potentially help to better understand gene function and diversity in closely related, economically significant species but also in other noneconomic but environmentally important gymnosperm species (Neale and Kremer, 2011).

Until recently, advances in the genomics of conifers were hampered by the large size of their genomes ranging from 20 to 40 Gb, which is more than 200-fold the *Arabidopsis* genome and roughly sevenfold the human genome (Mackay *et al.*, 2012; Ritland, 2012). As the conifer genome is extremely large, major research efforts were concentrated on transcriptomic analysis. The first large-scale EST (Expressed Sequence Tag) projects based on Sanger sequencing revealed that transcriptomes of pines and spruces are highly diverse and complex (Allona *et al.*, 1998; Cairney *et al.*, 2006; Li *et al.*, 2009; Pavy *et al.*, 2005; Ralph *et al.*, 2008).

The emergence of the next-generation sequencing (NGS) has profoundly transformed the landscape of genome analysis. Research efforts in conifers have provided growing catalogues of ESTs for several species of economic and ecological importance, including *Picea sitchensis* (Ralph *et al.*, 2008), *Picea glauca* (Rigault *et al.*, 2011), *P. pinaster* (Fernàndez-Pozo *et al.*, 2011) and *Picea abies* (Chen *et al.*, 2012). Recent landmark studies in conifer genomics reported the draft assembly of the first genome for a gymnosperm species. Specifically, massive parallel sequencing data were used for assembling the 20 Gb genomes of white (Birol *et al.*, 2013) and Norway (Nystedt *et al.*, 2013) spruces.

A full catalogue of genes and protein coding sequences in maritime pine is necessary to improve our understanding of biological functions and evolutionary relationships with other conifer trees and angiosperms. Comparative genomics provide a powerful mean to study gene structure and the evolution of gene function and regulation. Analysis of key genes and pathways allows scientists to better understand how complex biological processes are regulated and evolve (Koonin *et al.*, 2004; Soltis and Soltis, 2003).

In this study, we present results from deep sequencing of the *P. pinaster* transcriptome. The data were used to create a reference transcriptome in this pine species, containing an extensive set of genes expressed in a variety of actively growing tissues/organs (http://www.scbi.uma.es/sustainpinedb/sessions/new). Comparison to available public databases indicates that the transcriptome of *P. pinaster* is similar in size to that of previously characterized *P. glauca* (Rigault *et al.*, 2011) and *P. abies* (Nystedt *et al.*, 2013).

Results

Sequencing strategy and *de novo* assembly

Massive scale cDNA sequencing was performed to define a reference transcriptome for maritime pine. Long reads obtained by the 454 sequencing platform and short reads obtained by the Illumina platform were used for assembly of 210 513 putative unigenes from 18 cDNA libraries constructed from a variety of tree sources, tissues and experimental conditions (Table 1). Figure 1 shows a schematic representation of the workflow followed for preprocessing and assembling, as well as the combination of software algorithms utilized to obtain a final catalogue of unigenes. A total of 6 381 011 long reads provided 4 098 257 useful long reads (64.2%, mean length: 282 bp), and 591 174 069 short reads provided 292 596 546 useful short reads (49.5%, mean length: 86 bp). As shown in Figure 2, the recovery of useful 454 long reads was variable among the different libraries (37%–92%) with high recoveries (88%–92%) in EuroPineDB and UAGPF. Overall, recovery was around 70% of the initial reads. Regarding short reads, only 32% of single-end short reads were kept after data cleaning. The primary reason of out-filtering these single-end sequences was redundancy.

The assembly strategy used a combination of two completely different algorithms with the hypothesis that it may provide superior results than each algorithm by separate. MIRA3 and Euler-SR were selected for long-read assembly because the former is based on overlap-layout-consensus algorithm, while Euler-SR is based on de Bruijn graphs resolved by means of an Eulerian path. Furthermore, each algorithm produced the better results with simulated data (H Benzekri and MG Claros, unpublished results). Regarding short reads, we have used de Bruijn-based algorithms. ABySS was selected for two main reasons: (i) it can efficiently assemble using different k-mers (Lin *et al.*, 2011) and (ii) it is able to assemble vertebrate-sized genomes and transcriptomes (Li *et al.*, 2010). Finally, contigs separately obtained from short read and long read (355 483 in total) were reconciled using CAP3, resulting in 210 513 contigs to describe a *P. pinaster* transcriptome. Notably, 1241 unigenes were longer than 3000 nucleotides. Unigene length ranged from 40 to 7876 bp, with an average of 495 bp and a median of 361 bp.

The complete set of unigenes was compared using blast with the previous *P. pinaster* transcriptome available at EuroPineDB (Fernàndez-Pozo *et al.*, 2011), and 128 294 unigenes (61%) were not described in this earlier database, while 935 EuroPineDB unigenes were not present in the current unigene collection. When the *P. pinaster* transcriptome was compared with the *P. glauca* and *P. abies* draft genomes, 87.5% and 99.3% of

Table 1 Description of samples used for DNA sequencing

Gene library	Sequencing platform	Sampled plant material	Experimental conditions	SRA code
EuroPineDB	Sanger/454	Bud, xylem, phloem, stem, needles, roots, stem, embryos, callus, cone, male and female strobili	ESTs and SSH libraries from different tissues and conditions as described by Fernández-Pozo et al., 2011	SRS479769
Biogeco1	454	Xylem, bud and needle	ESTs from differentiating xylem, swelling bud and young needles	SRX032960, SRX032961, SRX032962, SRX032963
Biogeco2	454	Bud	EST from quiescent buds harvested on 2-year-old maritime pine (low growing family) in well-watered or drought-stress conditions	SRX031546
Biogeco3	454	Bud	EST from quiescent buds harvested on 2-year-old maritime pine (fast growing family) in well-watered or drought-stress conditions	SRX031589
UAGPF1	454	Embryome	ESTs from developing, immature embryos (1-week maturation)	SRX022618
INIA_PPIN	454	Bud	ESTs from buds	PRJNA221139
U_root	454	Root	ESTs from roots (1-month-old seedlings)	SRS480239
U_tip	454	Root tips	ESTs from root tips (1-month-old seedlings)	SRS480265
U_H	454	Hypocotyl	ESTS from hypocotyl (1-month-old seedlings)	SRS480236
U_N	454	Needle	ESTs from needles (1-month-old seedlings)	SRS480237
U_Cot_Os	454	Cotyledon	ESTs from cotyledons grown under dark conditions	SRS479771
U_H_Os	454	Hypocotyl	ESTs from hypocotyl grown under dark conditions	SRS480236
U_R_6	454	Roots	ESTs from roots (6-month-old seedlings)	SRS480238
U_S_8	454	Stem	ESTs from stem (8-month-old seedlings)	SRS480261
UAGPF2	Illumina	Somatic embryo	Paired-end ESTs from developing, immature embryos (1 week maturation)	SRR609713
BIOGECO4	Illumina	Bud	ESTs from young and aged buds	SRX031587
BIOGECO5	Illumina	Root	ESTs from drought-stressed and control roots in hydropony	SRX031592, SRX031590
BIOGECO6	Illumina	Bud	ESTs from young and aged buds	SRX031594
IBET	Illumina	Zygotic embryo	Paired-end ESTs from embryos	SRS481044

homology was found, respectively, confirming that most assembled unigenes were pine transcripts.

Annotation of unigenes

Unigene annotation was achieved by combining the results of several annotation processes. Each annotation is associated with an *E*-value to enable the empirical assessment of annotation quality. A preliminary analysis of the collection of unigenes using Full-LengtherNext (FLN) revealed that, from the 181 100 unigenes annotated (46.6%), 26 020 were nonredundant transcripts based on orthologue ID. It is also remarkable that 18 667 full-length (FL) unigenes were reconstructed with a mean length of 1495 nucleotides, representing 19% of the total annotated unigenes (Table 2). Of these, 9641 FL unigenes were different, unique genes (9.8% of the total annotated unigenes, and 37.0% of unique unigenes). The frequency distribution of FL unigenes (Figure 3) indicated a high proportion of unigenes ranging from 500 to 1500 nucleotides with the longest transcript being 7876 nucleotides.

Preliminary analysis using FLN also revealed that 111 577 unigenes (53.0%) did not possess significant homology to any other plant gene. This number includes new conifer genes as well as artefactual assemblies. To distinguish between both possibilities, FLN includes a TestCode analysis (Fickett, 1982) and a comparison with the noncoding RNA database (http://www.mirbase.org). As a result, at least 9799 nonredundant coding unigenes can worth a consideration of putative new conifer gene. In fact, 4608 unigenes had a homologue EST in the Pine Gene Index 9.0 database (http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=pine). However, only 176 unigenes from this study were determined to be candidate noncoding RNAs. Therefore, the minimal *P. pinaster* transcriptome can be calculated as 26 020 unigenes having unique ID. In addition, it could also be considered the 9799 unigenes without homology but having coding characteristics, plus the 176 noncoding RNAs, that is 35 995 unigenes.

Because this transcriptome was deemed satisfactory, genes annotated as described in Experimental procedures (GO term, a definition or a KEGG code) were subjected to statistical analyses. Of total unigenes, 62.2% (130 845) were annotated, which indicated that the level of annotation was similar to previously published results (Fernández-Pozo *et al.*, 2011). Furthermore, the distribution of GO terms at level 2 of biological process and at level 3 of molecular function (Figure S1) shows that the putative transcriptome covers most important cell functions. A total of 58 296 unigenes possessed unknown sequences that could not be found in existing databases. The annotated transcriptome can be browsed, downloaded and queried at http://www.scbi.uma.es/sustainpinedb/.

Validation of full-length cDNA sequences in the transcriptome database

Full-length cDNA (FLcDNAs) are essential for gene annotation, unambiguous determination of intron–exon boundaries and gene

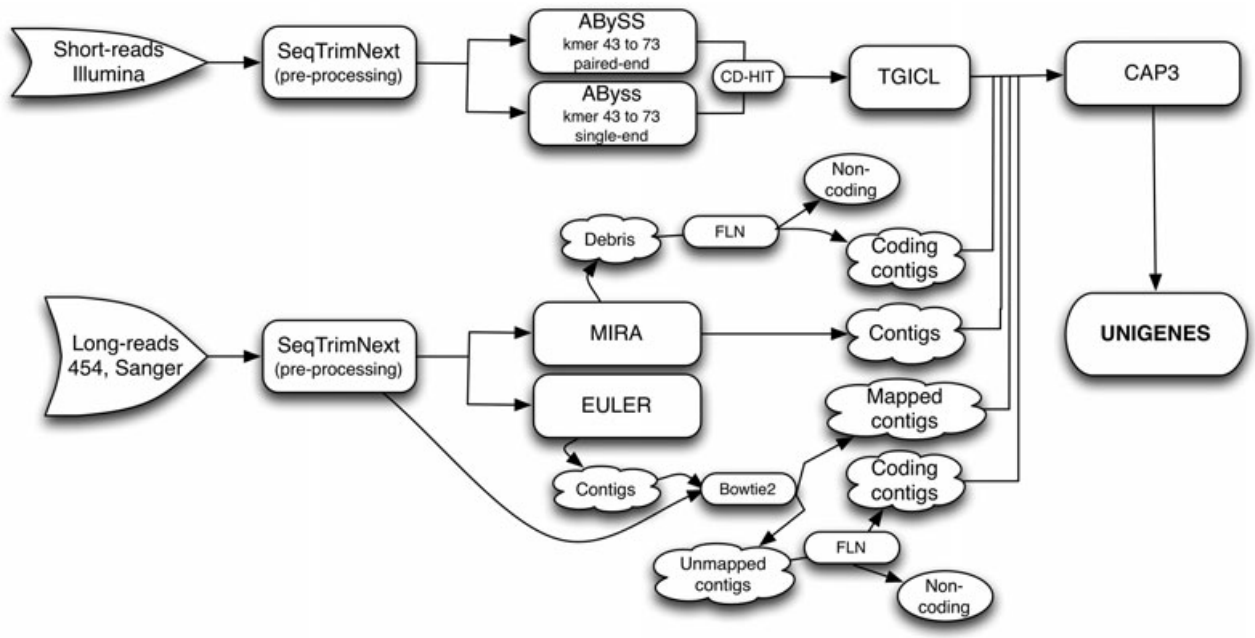


Fig. 1 Schematic representation of 454 and Illumina sequencing, trimming and assembling into unigenes of *Pinus pinaster* transcriptome short read and long read.

functional analysis. To examine the quality of the FLcDNA collection established in the maritime pine transcriptome database, the validation of a number of selected genes coding for proteins of a variety of sizes was advisable. Sequences encoding structural and regulatory genes were selected, appropriate primers based on available sequences were designed, and the corresponding FLcDNAs were amplified by RT-PCR from RNA samples extracted from a variety of maritime pine tissues. Table 3 shows a summary of this work. All selected genes were successfully amplified from cDNA samples using the sequence information available in the database. PCR products fit the theoretical, predicted size in most studied unigenes (Canales *et al.*, 2012; Cánovas *et al.*, 2007; Rueda-López *et al.*, 2013; Villalobos *et al.*, 2012).

Maritime pine regulatory genes

Transcription factors (TF) were specifically searched for in the Sustainpine database. The unique transcripts containing domains of plant TF in maritime pine using Pfam motifs were 877 distributed in 30 families (Table S1). Comparative analysis was performed with other woody plants including white spruce (*P. glauca*), poplar (*Populus trichocarpa*) and grapevine (*Vitis vinifera*), the herbaceous models *Arabidopsis* (*Arabidopsis thaliana)* and rice (*Oryza sativa*), and the model moss *Physcomitrella patens*. The total number of TF in maritime pine was similar to the number previously reported for white spruce (Rigault *et al.*, 2011), suggesting that the information quoted here could be close to the full representation of the maritime pine transcriptome. In addition, the TF gene number in *Physcomitrella* (802) and conifer species (877–892) is smaller than in angiosperms (Table S1), either woody (1337–2499) or herbaceous (1407–1828).

Overall, the relative representation of each TF domain among total TF domains was largely conserved in maritime pine and spruces (Figure 4 and Table S1) with only a few exceptions. For example, the SRF family is under-represented with only four members in maritime pine relative to the 45–71 reported in spruces or the 42–109 members reported in angiosperms. The AP2 family with 106 members is over-represented with regard to spruces (24–63), but contains a similar number of TF than in angiosperms (93–209). In contrast, the histone-like TF (CBF/NF-Y) family is under-represented in maritime pine (26) and Norway spruce (20) when compared to white spruce (52), but exhibits similar numbers to those found in angiosperm species.

When we compared this general distribution in maritime pine to six plant models, we observed conservation of TF categories that can be classified establishing a hierarchical clustering (Figure S2) where a closer conservation in members of each family is observed for *Physcomitrella*, spruce and pine with a second branch for *Arabidopsis*, rice and grapevine and a separate branch for poplar with a greater number of TF. Likely, the recent genome duplication increased the number of TF in each family of poplar (Tuskan *et al.*, 2006). Families not found in maritime pine and previously reported in white spruce are as follows: RNA polymerase II TF SIII, RNA polymerase III, TF IIIC subunit and alcohol dehydrogenase TF Myb/SANT-like (Rigault *et al.*, 2011).

Enzyme-encoding genes

The representation of gene family members involved in various metabolic pathways was determined, and data retrieved from the SustainpineDB were compared with those found in Norway spruce (*P. abies)*, *P. trichocarpa* and *A. thaliana* (Figure 5). A first set of genes studied were those participating in nitrogen acquisition and assimilation. The number of genes encoding nitrate reductase and nitrite reductase was similar in the three species (1–2). For genes involved in inorganic nitrogen transport, some important differences were observed among species. For instance, *Arabidopsis* contains almost twice genes encoding nitrate transporters that maritime pine, Norway spruce and poplar. For ammonium transporters, only the poplar genome presents an expanded gene family of 14 members.

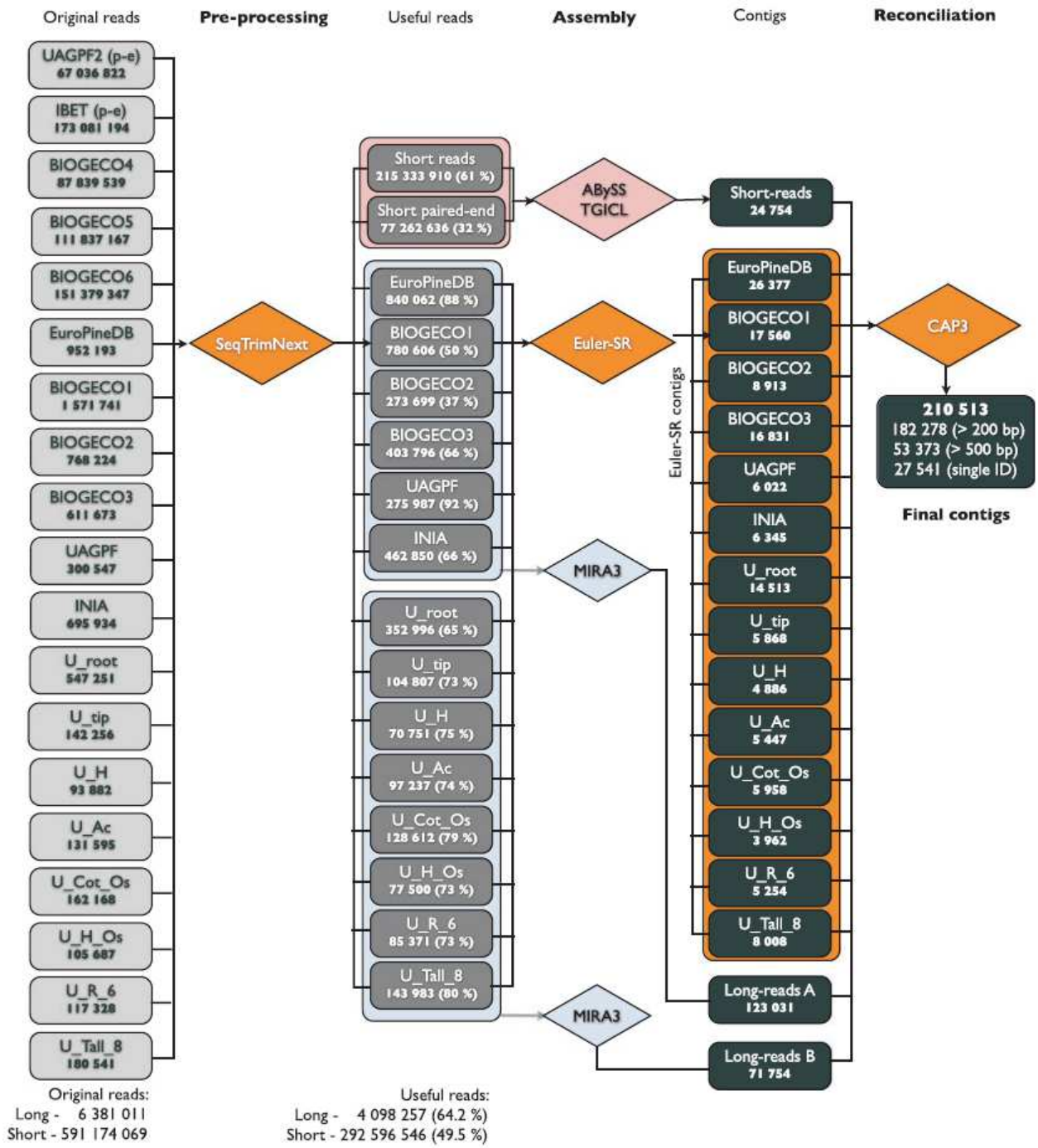


Fig. 2 Flow chart showing preprocessing into useful reads, assembly into contigs and overlap-based reconciliation into final unigenes of sequenced data from 5 (591 174 069 short reads, Illumina) or 14 (6 381 011 long reads, 454) cDNA libraries in maritime pine.

Fewer transcripts for genes encoding enzymes of ammonium assimilation were found in conifer species. Only 2–3 transcripts for glutamine synthetase (GS) were identified in gymnosperms (*P. pinaster* and *P. abies*), in accordance with previous results (Cánovas *et al.*, 2007). In contrast, genomes of angiosperm species are endowed with GS families with a higher number of members, eight in *Populus* and six in *Arabidopsis*. A single expressed gene was found for ferredoxin-glutamate synthase (Fd-GOGAT) and NADH-GOGAT in maritime pine and Norway spruce. In contrast, two genes encode Fd-GOGAT and NADH-GOGAT in poplar.

A second group of genes where those encoding enzymes involved in synthesis of methionine and S-adenosylmethionine (SAM), the activated form of methionine, which participate in a number of essential metabolic pathways in plants. In particular, we focused on three genes involved in the synthesis and recycling of SAM, a methyl donor in multiple cellular transmethylation reactions (Figure 5). The number of genes encoding cobalamin-independent

Table 2 Summary of final data for the *P. pinaster* transcriptome

	Absolute number	%
Unigenes	210 513	100.0
Artefacts	2010	0.95
Unigenes after resolving artefacts	209 928	99.7
Unigenes > 200 nt	181 100	86.0
Unigenes > 500 nt	52 550	25.0
Unigenes > 3000 nt	1241	0.59
Unigenes with orthologue*	98 175	46.6
Different orthologue ID	26 020	26.5
Complete transcripts (full length)	18 667	19.0
Different full-length transcripts	9641	9.8
Putative ncRNAs	176	0.08
Unigenes without orthologue*	111 577	53.0
Coding	29 736	26.6
Putative coding	23 545	21.1
Nonredundant coding	9799	8.8

Numbers in bold can be considered the representative amount of unigenes of the category.

*Percentages for this classification are calculated using this file as 100.0%.

methionine synthase and SAM synthase was identical in conifers and *Arabidopsis* but slightly higher in *P. trichocarpa*, whereas an additional S-adenosyl-L-homocysteine hydrolase gene was found in pine and Norway spruce in regard to angiosperms.

Cellulose, hemicellulose and lignin are major components of the secondary cell wall, a well-developed structure in woody perennials compared with *Arabidopsis,* particularly in trees such as maritime pine and poplar. A similar number of genes encoding the cellulose synthase catalytic subunit (CesA) were found in conifers compared with *Arabidopsis*, whereas the number increased to 18 in poplar. The number of genes encoding sucrose synthase and UDP-glucose 6-dehydrogenase in the maritime pine transcriptome was slightly lower compared with the other three plant species.

We also compared the number of genes encoding three key enzymes of the phenylpropanoid pathway that direct the carbon flow from the shikimate pathway to monolignol biosynthesis: phenylalanine ammonia-lyase, cinnamoyl-CoA reductase and cinnamyl alcohol dehydrogenase (Figure 5). The number of genes encoding the three enzymes was similar in maritime pine, Norway spruce and *Arabidopsis* and also in poplar in the case of phenylalanine ammonia-lyase. In contrast, the members of the cinnamoyl-CoA reductase and cinnamyl alcohol dehydrogenase families are considerably higher in poplar (40 and 21, respectively).

Single nucleotide polymorphism and simple-sequence repeat identification

Using CLCBio with stringent criteria both in terms of minimum allele frequency (MAF = 20%) and contig depth (minimum of 10 reads), we identified 55 607 in silico Single nucleotide polymorphism (SNPs) in 13 892 unigenes (Table S2), corresponding to an average of 1 SNPs per 2 kb surveyed in the assembled transcriptome. Setting the contig depth parameter to 4 or 10, yielded 542 830 or 20 208 SNPs, respectively, with GigaBayes. Suprisingly, only 10 829 and 2106 SNPs were simultaneously found between CLCBio and GigaBayes with relaxed (minimum contig depth, CRL = 4) and more stringent (CRL=10) detection criteria, respectively; corresponding to 2% and 10.4% of common SNPs with CLCBio. Despite a fivefold increase in commonly detected SNPs using the more stringent detection criteria in GigaBayes, both software detected highly divergent sets of SNPs. Given this inconsistency, we used a set of 5138 SNPs validated in natural populations and mapping pedigrees of maritime pine (Lepoittevin *et al.*, 2010; Chancerel *et al.*, 2011; Chancerel *et al.,* 2013) to evaluate the ability of both software to identify true SNPs. Prior to this analysis, we checked that the flanking sequences associated with the true SNPs were found in the present assembly using BLASTn analysis. Using CLCBio, 36.26% (1863 SNPs) of true SNPs were detected, while this rate dropped to 9.17% (471 SNPs) and 1.36% (70 SNPs) using relaxed and stringent GigaBayes criteria. By normalizing

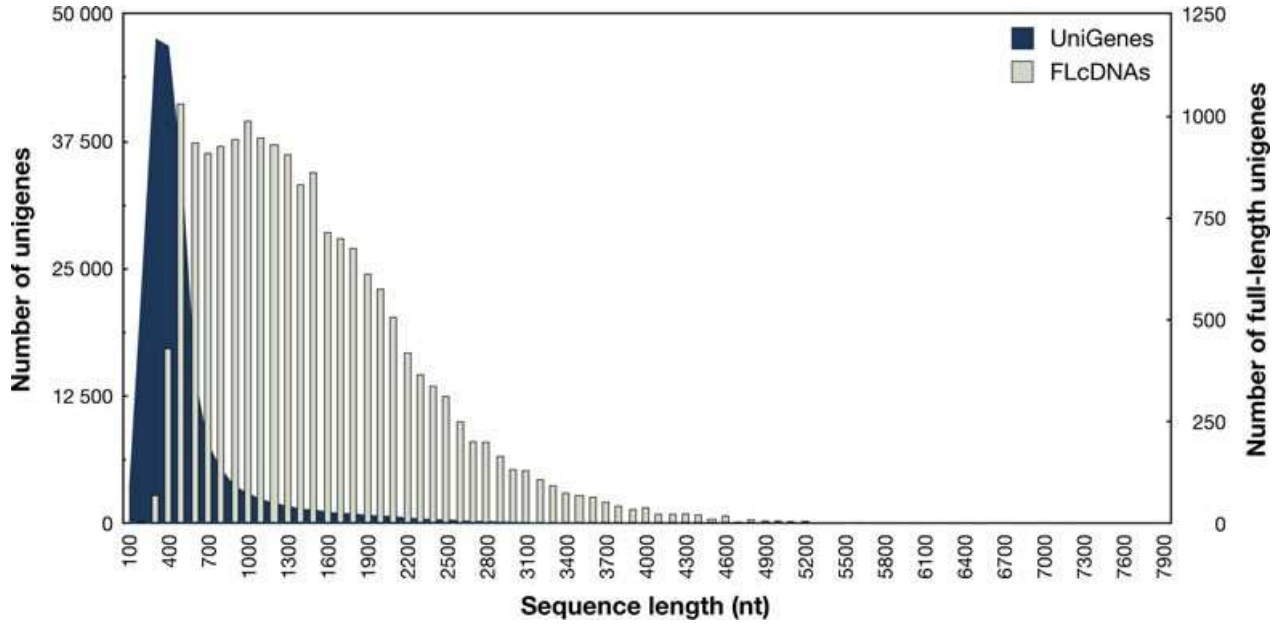


Fig. 3 Frequency distribution of total unigenes and full-length unigenes.

Table 3 Summary of FLcDNA used for sequence assembly validation by PCR amplification

Gene name	Theoretical size of ORF from assembly (bp)	Experimental size of ORF (bp)	Accession number*
Arginase	1026	1026	sp_v3.0_unigene23824
Xyloglucan endotransglycosylase	860	860	sp_v3.0_unigene29476
Phenylalanine ammonia-lyase	2265	2265	sp_v3.0_unigene17298 AY641535
PII-protein	714	732	sp_v3.0_unigene23578 AJ489604
Asparagine synthetase 1	1782	1782	sp_v3.0_unigene14147 HQ625490
Asparagine synthetase 2	1770	1773	sp_v3.0_unigene18231 HM222940
Glutamate decarboxylase	1530	1530	sp_v3.0_unigene11755 DQ125946
Glutamate dehydrogenase	1236	1236	sp_v3.0_unigene15901 HM222941
Sucrose synthase	1914	1914	sp_v3.0_unigene34880
Ammonium transporter 1.1	1584	1584	KC807907
Ammonium transporter 1.2	1464	1464	KC807908
Ammonium transporter 1.3	1539	1539	KC807909
Ammonium transporter 2.1	1461	1461	KC807910
Ammonium transporter 2.3	1446	1446	KC807911
Glutamine dumper 1	345	345	sp_v3.0_unigene97635
Glutamine dumper 2	384	384	sp_v3.0_unigene20421
†MYB1	1023	1023	sp_v3.0_unigene29297 EU482890
†MYB4	946	946	sp_v3.0_unigene127348
†MYB 8	1605	1605	FN868598
‡Dof2	1065	1065	KC688677
‡Dof3	1047	1047	KC688678
‡Dof4	1053	1053	KC688679
‡Dof6	907	907	KC688680
‡Dof7	1404	1404	KC688681
‡Dof8	822	822	KC688682
‡Dof9	897	897	KC688683
§NAC1	1559	1545	sp_v3.0_unigene7635
§NAC2	1476	1398	sp_v3.0_unigene14173
§NAC3	1457	1388	sp_v3.0_unigene18613
§NAC4	1630	1278	sp_v3.0_unigene20354
§NAC5	1824	1180	sp_v3.0_unigene1398
Methionine synthase	2301	2301	sp_v3.0_unigene34452 HE566045
S-adenosylmethionine synthase	1176	1176	sp_v3.0_unigene3027 HE574556
S-adenosylhomocysteine hydrolase	1458	1458	sp_v3.0_unigene19530 HE574555
Methylenetetrahydrofolate reductase	1785	1785	sp_v3.0_unigene15993 HE574560
Caffeate O-methyltransferase	1161	1095	sp_v3.0_unigene17184 HE574557
Hydroxycinnamoyl-CoA: shikimate hydr oxycinnamoyl transferase	1302	1302	sp_v3.0_unigene8683 HE574565
Glycine decarboxilase complex H-protein a	504	504	sp_v3.0_unigene30488 HE574553
Glycine decarboxilase complex H-protein b	516	516	sp_v3.0_unigene126824 HE574563
Mitochondrial serine hydroxymethyltransferase	1572	1572	sp_v3.0_unigene439 HE574554

Table 3 Continued

Gene name	Theoretical size of ORF from assembly (bp)	Experimental size of ORF (bp)	Accession number*
Cytosolic serine hydroxymethyltransferase	1413	1413	sp_v3.0_unigene17057 HE574564
D-3-Phosphoglycerate dehydrogenase	1947	1947	sp_v3.0_unigene543 HE574561
3-Phosphoserine aminotransferase	1302	1302	sp_v3.0_unigene37851 HE574562
Pinoresinol-lariciresinol reductase	939	939	sp_v3.0_unigene17681 HE574558
Phenylcoumaran benzylic ether reductase	927	927	sp_v3.0_unigene31659 HE574559
Phenylpropenal double-bond reductase	1056	1056	sp_v3.0_unigene22698 HE575885

*Accession number of unigene in Sustainpine and GenBank.

†MYB family of TF.

‡Dof family of TF.

§NAC family of TF.

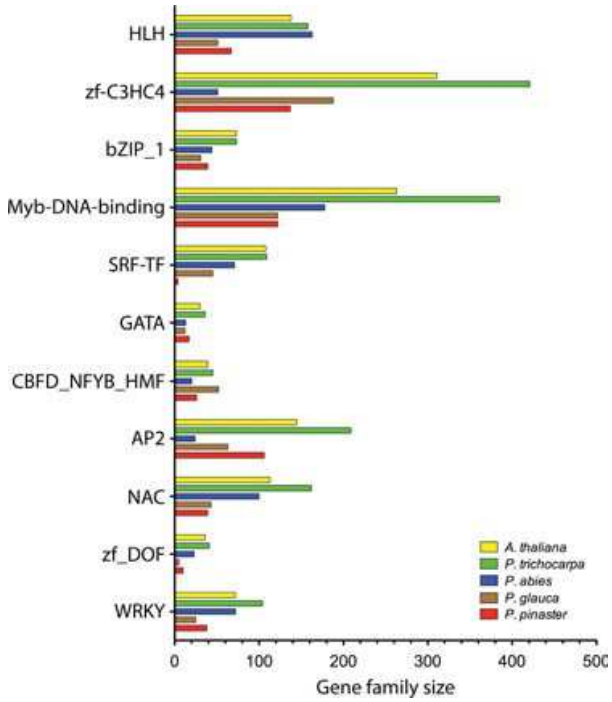


Fig. 4 Distribution of unique transcripts corresponding to TF gene families in *Pinus pinaster* and comparison to other plant transcriptomes. The number of different encoded transcripts with the conserved DNA-binding domain of each family is represented. The distribution of TF gene families in *P. pinaster*, *Picea glauca, Picea abies, Populus trichocarpa and Arabidopsis thaliana* is compared.

Gygabayses detection rates to the 55 607 putative SNPs identified by CLCBio, the relaxed and stringent GigaBayes detection procedures led to a success rate of 0.94% (48 true SNPs detected) and 3.75% (193 SNPs), respectively. In conclusion, CLCBio was found to perform nearly 10 times better than the most stringent GigaBayes SNP detection procedure with our data set.

A total of 5974 putative simple-sequence repeat (SSRs) were found, with trinucleotide repeats (3309) being the most common, and dinucleotide repeats (479) the less abundant. This is in agreement to previously published *P. pinaster* SSR abundance (Fernàndez-Pozo *et al.*, 2011).

Discussion

Maritime pine transcriptome assembly

Long-read sequence data sets are required for transcriptome assembly in nonmodel species for which a reference genome is not available. In conifers, 454 sequencing has been recently used to generate well-defined transcriptomes in several species of ecological and economic interest, that is, *Pinus contorta* (Parchman *et al.*, 2010), *P. glauca* (Rigault *et al.*, 2011), *P. pinaster* (Fernàndez-Pozo *et al.*, 2011), *Pinus taeda* and 11 other conifers (Lorenz *et al.*, 2012). In the present work, we used a combination of 454 and Illumina sequencing to define a minimal reference transcriptome for maritime pine (*P. pinaster*). A similar approach was recently used to characterize, for example, the globe artichoke transcriptome (Scaglione *et al.*, 2012). The nonredundant transcriptome resulting from the assembly contains 26 020 unique transcripts with orthologue ID in public databases, a number very close to the 27 720 unique cDNA clusters reported for the *P. glauca* transcriptome (Rigault *et al.*, 2011) and higher than the 17 000 unique coding genes obtained in the assembly of *P. contorta* transcriptome (Parchman *et al.*, 2010). The number of unique transcripts in maritime pine is also close to the number of genes (28 354) resulting from the draft assembly of the 20-gigabase genome of *P. abies* (Nystedt *et al.*, 2013). Considering all the available data, an elevated coverage of the maritime pine transcriptome is estimated.

FLcDNAs catalogues as genomic resources

The availability of large collections of FLcDNAs in several conifers, such as Sitka (Ralph *et al.*, 2008) and white spruces (Rigault *et al.*, 2011), as well as *Cryptomeria* (Futamura *et al.*, 2008), has greatly facilitated the assembly and annotation of FLcDNAs in maritime pine. FLcDNAs are crucial for accurate

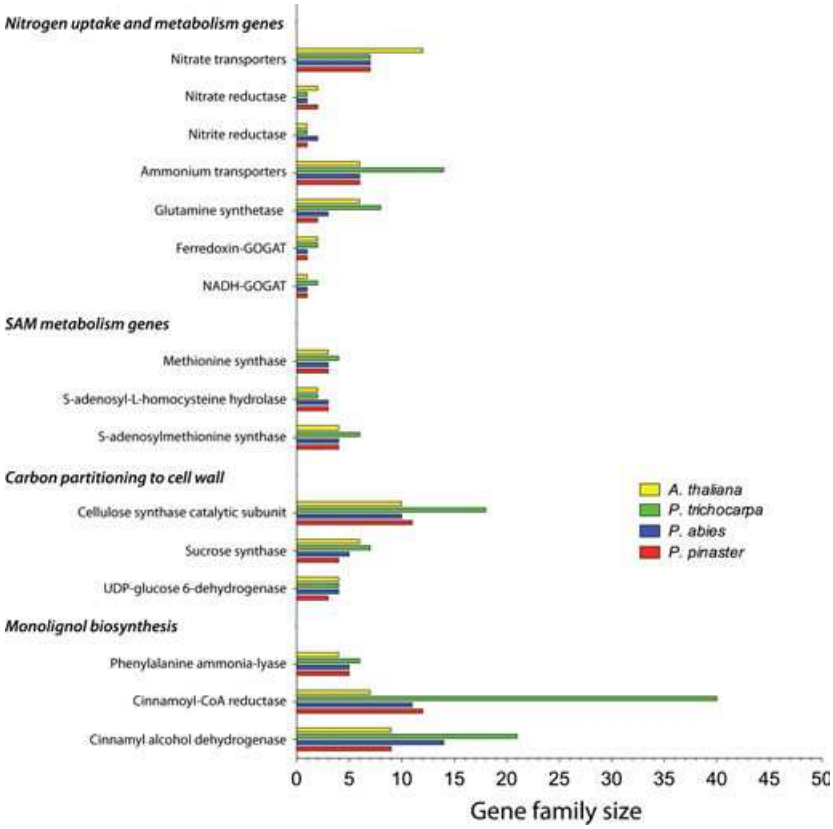


Fig. 5 Comparison of gene families for relevant enzymes in *Pinus pinaster, Picea abies, Populus trichocarpa* and *Arabidopsis thaliana.* The following databases were used in addition to *Sustainpine*DB*: P. abies* v1.0, *P. trichocarpa* v3.0, *A. thaliana* TAIR 10.

annotation, comparative analysis with other conifer species and also for functional analysis of relevant genes associated to maritime pine growth, development and response to environmental changes. Furthermore, this genomic resource will greatly facilitate protein identification as well as protein–protein interaction studies through proteomics approaches (Cánovas *et al.*, 2004). For all these reasons, it was of paramount importance to validate the assembly of the FLcDNA collection (9641 different transcripts).

Over the last few years, refined protocols have been developed for *Agrobacterium*-mediated genetic transformation of maritime pine embryogenic tissue, cryopreservation of transgenic lines and efficient transgenic plant regeneration through somatic embryogenesis (reviewed in Trontin *et al.*, 2013). These new developments and the availability of a large collection of FLcDNAs have paved the way for the application of reverse genetics towards functional dissection of traits of economic and ecological interest in maritime pine trees. Thus, the availability of both standardized transformation methods and FLcDNA catalogues is expected to significantly increase throughput in candidate gene analysis together with facilitating comparison across laboratories interested in maritime pine genomics. The functional analysis of key genes is crucial for future applications in tree improvement, new variety design and sustainable forest management (e.g. development of marker-assisted selection).

Maritime pine gene families and genome size

It has been suggested that the increased size and complexity of conifer genomes relative to angiosperms may be explained by the existence of large gene families (Kinlaw and Neale, 1997). However, this assumption is not fully supported by available data as most TF (Figure 4 and Table S1) or other gene families (Figure 5) present in maritime pine (this work) or spruce genomes (Birol *et al.*, 2013; Nystedt *et al.*, 2013; Rigault *et al.*, 2011) were of similar or even lower size compared with angiosperm species (*P. trichocarpa*, *A. thaliana* and *V. vinifera*). Meanwhile, the existence of large gene families in conifers coding for enzymes of secondary metabolism has been reported (Martin *et al.*, 2004), there are other families in primary metabolism that contain similar, or even shorter, number of functional members in conifers than in angiosperms (Cánovas *et al.*, 2007). High genome size and complexity in conifers may be more readily explained by divergence and accumulation of retrotransposons and pseudogenes (Morse *et al.*, 2009; Nystedt *et al.*, 2013). Retrotransposons and pseudogenes can be expressed and contribute in some extent to the collection of unigenes without orthologue in the maritime pine transcriptome. Accumulation of pseudogenes may have functional advantages in the regulation of gene expression if they are expressed. Recently, Poliseno *et al.* (2010) reported that expressed pseudogenes compete with authentic target transcripts for miRNA binding and, as such, modulate expression levels of their cognate genes.

Transcription factors

The identification of transcription factors (TF) and subsequent analysis of the composition and organization of TF families are necessary steps to understand the regulatory networks associated with key processes in conifer trees. The number of TF in *P. pinaster* appears to be similar to that of *P. glauca* (Rigault *et al.*, 2011), but considerably lower compared with those found in the genomes of several angiosperms. This fact is confirmed by studies carried out in specific families. For example, the Dof family has only ten members in maritime and loblolly pines (Rueda-López *et al.*, 2013); this is twofold to eightfold lower than gene

number in angiosperms, including 36 in *A. thaliana*, 30 in rice (Lijavetzky *et al.*, 2003), 22 in grape or 46 in poplar (Pérez-Rodríguez *et al.*, 2010). Moreover, it has also been argued that some subfamilies such as Knox I involved in meristem cell identity have pursued a distinct path of evolution in conifers. Three of the four groups delineated by phylogenetic analyses in angiosperms have no conifer sister groups (Guillet-Claude *et al.*, 2004), suggesting that conifers evolved in a nonlinear fashion compared with angiosperms. Recently, it has been proposed that expansion of the VASCULAR NAC DOMAIN (VND) gene family might be related to xylem vessel complexity in angiosperms (Nystedt *et al.*, 2013). Other example of this divergent pattern of evolution for TF is found in the HD-Zip III gene family involved in regulation of cambium, and primary and secondary vascular differentiation whose structure also has diverged considerably between angiosperms and gymnosperms (Côté *et al.*, 2010). Furthermore, distinct evolutionary trends in angiosperms and gymnosperms are evident by differential family gene expansion of subgroup 4 R2R3-Myb with more recent duplications in *P. glauca* (Bedon *et al.*, 2010). In this particular case, the higher number of members in conifers appears to be related with isoprenoid- and flavonoid-oriented stress responses.

Most of the existing knowledge of plant TF genes was obtained from studies in *Arabidopsis.* While *Arabidopsis* is a useful model for many developmental and environmentally regulated processes in higher plants, it lacks certain traits that are of immense value to agriculture or forestry. To better understand the evolutionary relationship and dependence of transcriptional regulation and morphological complexity in *Viridiplantae* is important to analyse how particular families of transcription factors did expand in correlation with the general increase in morphological complexity. In this sense, the availability of full-length sequences from gymnosperms come to close a huge gap and will help to understand how the land plants managed, in terms of transcriptional regulation, to become multicellular.

Gene families of metabolic pathways

A major contrast was observed for genes involved in nitrogen metabolism likely related with adaptations to nitrogen availability and use. For example, poplar, Norway spruce and pine possess almost half the number (7) of nitrate transporters encoded in the *Arabidopsis* genome (12). This difference may be due to differences among species on soil nitrate uptake and transport during active growth, compared with trees in which active growth is primarily supported by internal nitrogen storage and mobilization (Cantón *et al.*, 2005). In contrast, the family of ammonium transporters is expanded in *P. trichocarpa* (14) compared with pine, spruce and *Arabidopsis* (6). It has been suggested that this specific feature may be related to the peculiar physiology of a perennial and mycorrhizal tree (Couturier *et al.*, 2007). However, the observation of similar number of ammonium transporters in conifer trees and *Arabidopsis* does not support the above assumption anymore. Another major difference is the reduced number of the GS gene family in pine (2) and spruce (3) compared with *Populus* (8) and *Arabidopsis* (6). Deep transcriptome analysis in the current study supports previous reports describing only two genes for cytosolic GS pine and the lack of a plastidic isoform (Cánovas *et al.*, 2007).

In trees, consumption of methyl units during lignification implies the existence of an important carbon sink, and the S-adenosylmethionine availability may affect wood quality through alterations in lignin content and composition (Villalobos *et al.*, 2012). Nevertheless, no major differences were found in the number of genes in poplar and conifers with respect to *Arabidopsis*.

The number of genes related to carbon partitioning to cellulose and hemicellulose seems to be similar in the four species, with the exception of CesA that is expanded in *P. trichocarpa*. The *CesA* gene family in poplar includes 18 members that form nine phylogenetic groups, eight of which contain a pair of *CesA* genes with nearly identical sequence as a likely result of recent genome duplication. Some of them showed redundant expression xylem-specific, and it was suggested that more *CesA* genes could be required for the massive synthesis of cellulose in trees (Suzuki *et al.*, 2006). However, the number of genes encoding *CesA* subunits in *P. pinaster, P. abies* and *A. thaliana* seems to be similar.

Monolignol biosynthesis is a complex branch of the general phenylpropanoid pathway. The number of genes encoding phenylalanine ammonia-lyase was similar in all species (4–6). However, an increased number of genes encoding cinnamoyl-CoA reductase were found in *P. trichocarpa*, *P. pinaster and P. abies* compared with *Arabidopsis*. Previous work in angiosperms showed that cinnamoyl-CoA reductase family represents the largest lignin biosynthesis gene family in several species (Xu *et al.*, 2009). In conifers, transcriptomic data reported here suggest that the cinnamoyl-CoA reductase family is also expanded. In contrast, it has been previously reported that conifers may have a single copy of the cinnamyl alcohol dehydrogenase gene (Mackay *et al.*, 1997). However, different transcripts were found in maritime pine and Norway spruce.

Overall, gene families involved in metabolic pathways were of similar size in *P. pinaster* and *P. abies*. The number of genes in *P. pinaster* was either similar or lower to those existing in *Populus* or *Arabidopsis*. These findings should be confirmed when a reference genome for pine is available.

SNP identification

Based on previous unigene sets constructed with significantly fewer sequences obtained essentially from one ecotype (that of the Aquitaine provenance), highly multiplexed SNP arrays were constructed in maritime pine for linkage applications (Chancerel *et al.*, 2011, 2013) and association mapping (Lepoittevin *et al.*, 2012). The present study provides the most comprehensive SNP catalogue for maritime pine bringing together sequences from a diverse range of ecotypes (France, Spain, Portugal) and therefore allowing the use of high throughput genotyping technologies for applications in: i) genetic diversity and population structure analysis, with less ascertainment bias than was observed from the analysis of previous SNP data sets (Chancerel *et al.,* 2011) thank to a better coverage of the species diversity; and, ii) genomic selection, because about 14k different gene-loci are now covered by at least one SNP. Our study reveals a strong influence of the SNP calling algorithm on the number of detected SNPs. Such variability in the number of called SNPs between SNP detection software has been already reported in another conifer species (Muller *et al.*, 2012) and indicates that results should be interpreted with caution, especially if based on a single detection approach. In our case, using a set of already validated SNPs, we show that CLCBio was able to detect true SNPs at a rate that was ten times higher than the most stringent criteria implemented in GigaBayes. Trees have long generation times, and breeding (especially for pine) is a slow process (Mullin *et al.*, 2011). Genomic selection offers the possibility to increase genetic gain per time unit

in these long-lived organisms as illustrated recently for *P. taeda* (Resende *et al.*, 2012) and accelerate their domestication (Harfouche *et al.*, 2012). Genomic selection will also permit to control more precisely that sufficient level of diversity is maintained in future varieties to allow forest trees to cope with major biotic and abiotic constraints in rapidly evolving environment.

Concluding remarks

In this work, comprehensive characterization of the *P. pinaster* transcriptome was performed using a combination of two different next-generation sequencing platforms, 454 and Illumina. The *de novo* assembly of the maritime pine transcriptome provides a large catalogue of expressed genes (26 020 unigenes with orthologue, 9799 unigenes with coding characteristics but without orthologue, 176 ncRNA) and a relevant collection of FLcDNAs (9641). In addition, the sequencing data permitted the identification and establishment of robust SSR- and SNP-based databases for genotyping applications and translational integration in maritime pine breeding programmes. These genomic resources will facilitate genome sequencing, functional genomics and applied studies in maritime pine trees.

Experimental procedures

Tree source, tissues and experimental conditions

Pinus pinaster samples of developing xylem were collected from different genotypes of a Corsican clonal population planted in 1986 at the forestry station of INRA-Pierroton (Aquitaine, France), as described (Villalobos *et al.*, 2012). Miscellanea of maritime pine tissues (cones, male and female strobili, buds, xylem and phloem) were collected from adult trees located at Sierra Bermeja (Málaga). Roots, stems and needles of 2-week-old pine seedlings and zygotic embryos were also sampled as well as maritime pine embryonal masses. Somatic embryos were produced from embryonal masses line (AAY06006) of *P. pinaster* originating from a Landes 9 Morocco polycross. Embryonal masses were induced from immature zygotic embryo (Park *et al.*, 2006). Embryonal masses and somatic embryos were cultured according to Lelu-Walter *et al.* (2006). Zygotic embryos at several developmental stages isolated from immature seeds were also used. All samples were frozen in liquid nitrogen and stored at -80 °C untill further use.

RNA isolation, cDNA synthesis and construction of libraries

Total RNA was isolated as described by Canales *et al.* (2012). Double-stranded cDNA libraries were constructed using the MINT kit (Evrogen) with incorporated modifications proposed by Babik *et al.* (2010). The quality of cDNA libraries was checked on 1.5% agarose gels and Agilent 2100 Bioanalyzer. cDNA libraries were also constructed from poly(A) enriched RNA. Briefly, equal amounts of total RNA (50 lg) from different tissues were combined prior to mRNA purification. Poly(A)+ mRNA was double purified using the Oligotex mRNA Mini Kit (Qiagen), according to the manufacturer's instructions. First and second strand cDNA synthesis was conducted from 10 lg of mRNA following the protocol described by Durban *et al.* (2011). Five cDNA synthesis reactions were performed to obtain sufficient cDNA quantity for 454 pyrosequencing. RNA isolation, purification and library construction from somatic embryos were carried out by GATC company (Constanz, Germany) using Smart cDNA construction kit and standard procedures for Illumina GA IIx sequencing and using standard kit for construction of normalized cDNA library dedicated to Roche/454 GS FLX sequencing.

DNA sequencing and preprocessing analysis

Libraries were sequenced using different sequencing platforms, including Roche/454 Titanium, Roche/454 GS FLX, Illumina GAIIx and Illumina HiSeq 2000. Chloroplast, mitochondrial and ribosomal sequences in short reads were filtered out by reference mapping using BWA (Li and Durbin, 2009). All reads were preprocessed using the SeqTrimNext pipeline (http://www.scbi.uma.es/seqtrimnext; Falgueras *et al.,* 2010) available at the Genotoul's computer facilities (INRA Toulouse, France) and at the Plataforma Andaluza de Bioinformática (University of Málaga, Spain). Low quality, ambiguous and low complexity stretches, linkers, adaptors, vector fragments, organelle DNA, polyA/polyT tails, and contaminated sequences were removed while keeping the longest informative part of the read, discarding sequences below 20 (short reads) or 40 bp (long reads). The command line for long reads was *seqtrimnext -t transcriptomics_plants.txt -Q input_reads.fastq > output.txt*, where the configuration file *transcriptomics_plants.txt* is provided by default with SeqTrimNext and contains the configuration parameters. The command line for short reads was the same than above, but configuration file also indicates to filter out sequences shorter than 20 pb, skip preliminary statistic calculations and skip the read clonality analysis.

Transcriptome assembly

The bioinformatics process for obtaining the putative unigenes using the SeqTrimNext preprocessed reads is detailed in Figure 1. Short reads and long reads were assembled using well-established protocols (see Methods S1). The complexity and redundancy of short read and long read assemblies was reduced using CAP3 assembler (Liang *et al.*, 2000) to obtain the unigenes. A 95% cut-off for overlap per cent identity was applied to cope with both sequence variation (high heterozygosity of pine genes) and genome heterogeneity between samples as previously described (Fernández-Pozo *et al.*, 2011).

Annotation analysis

Unigenes were analysed using Sma3s (Pérez-Pulido, A., Muñoz-Mérida, A., Claros, M.G., Trelles, O., submitted; http://www.bioinfocabd.upo.es/node/11) with the command line *sma3s_v2.pl -a 123 -d uniprot_plants.dat -i Unigenes.fasta -b Unigenes.blast -p F > output* using as input the unigenes (.fasta file), the blast result of unigenes against the plant UniProtKB database (.blast file) and the metadata of the plant UniProtKB database (.dat file) to provide a gene description, GO terms, EC keys, KEGG maps and InterPro codes for every sequence. They were also analysed with FLN to provide gene description, identify, which unigenes corresponded to FLcDNAs, detect putative start and stop codons as well as the putative protein sequence, extract, which unigenes could be sRNAs, and obtain a quick preview of the pine unigene content. AutoFact (Koski *et al.*, 2005) was also used to provide a third gene description.

Identification of polymorphisms

Because nucleotide variation is considered frequent in plant genes, we screened the unigenes for SNPs and SSRs variation. SSRs were screened using MREPS (http://bioinfo.lifl.fr/mreps/; Kolpakov *et al.*, 2003) with default parameters. SNP detection

was performed based on 4 098 953 Roche/454 and Sanger reads using either CLC-Bio Workbench, v6.0 (CLC Bio, Aarhus, Denmark) or Gygabayes (http://bioinformatics.bc.edu/marthlab/wiki/index.php/Software), an implementation and expansion of the PolyBayes SNP detection algorithm from Marth *et al.* (1999). With respect to the former, we first used the reference mapping function to map the reads onto the 210 513 unigenes generated in the present study. The following parameters were used: similarity score = 80%, minimum length fraction = 90% and maximum number of hits for reads = 10. Then, we used the SNP detection function, based on the neighbourhood quality standard (NQS) algorithm, using the following parameters: minimum coverage = 10, minimum central base quality = 20, minimum neighbourhood quality over a window length of 11 nucleotides = 15, maximum gap and mismatch count = 2 and minimum allele frequency = 20%. With respect to the latter, we used the following parameters: (ploidy = diploid; CRL = 4 or 10; CAL2 = 2; PSL=0.9; O = 1; D = 0.001). In both software, no insertion/deletion variants (InDels) were considered.

Acknowledgements

We are indebted to the anonymous reviewers for their thorough evaluation and constructive recommendations that helped to improve this manuscript. We would like to thank Aranzazu Flores Monterroso, Remedios Crespillo, Jose Vega Bartol and Marta Simões for help in the preparation of samples. This work was supported by the SUSTAINPINE Project funded by the Plant KBBE programme, Scientific and Technological Cooperation in Plant Genome Research (PLE2009-0016). Zygotic embryo data production was supported by FCT (Portugal) through projects PTDC/AGRGPL/102877/2008 and P-KBBE/AGR-GPL/0001/2009, and grants PEst-OE/EQB/LA0004/2011 and SFRH/BD/79779/2011 (AR). Alexandre Aguiar and Isabel Carrasquinho from Instituto Nacional de Investigação Agrária e Veterinária (INIAV) are acknowledged for making zygotic embryos available. Somatic embryo data production was supported by the EMBRYOME project, contract number 33639, funded by the french 'Conseil Regional de la Region Centre'.

References

Allona, I., Quinn, M., Shoop, E., Swope, K., St Cyr, S., Carlis, J., Riedl, J., Retzel, E., Campbell, M.M., Sederoff, R. and Wetten, R.W. (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proc. Natl Acad. Sci. USA,* 95, 9693–9698.

Aranda, I., Alia, R., Ortega, U., Dantas, A.K. and Majada, J. (2010) Intra-specific variability in biomass partitioning and carbon isotopic discrimination under moderate drought stress in seedlings from four *Pinus pinaster* populations. *Tree. Genet. Genome*, 6, 169–178.

Babik, W., Stuglik, M., Qi, W., Kuenzli, M., Kuduk, K., Koteja, P. and Radwan, J. (2010) Heart transcriptome of the bank vole (*Myodes glareolus*): towards understanding the evolutionary variation in metabolic rate. *BMC Genomics,* 11, 390.

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S.D., Schmidt, A., Gershenzon, J., Grima-Pettenati, J., Seguin, A. and MacKay, J. (2010) Subgroup 4 R2R3-MYBs in conifer trees: gene family expansion and contribution to the isoprenoid-oriented responses. *J. Exp. Bot.* 61, 3847–3864.

Birol, I., Raymond, A., Jackman, S.D., Pleasance, S., Coope, R., Taylor, G.A., Yuen, M.M., Keeling, C.I., Brand, D., Vandervalk, B.P., Kirk, H., Pandoh, P., Moore, R.A., Zhao, Y., Mungall, A.J., Jaquish, B., Yanchuk., A., Ritland, C., Boyle, B., Bousquet, J., Ritland, K., Mackay, J., Bohlmann, J. and Jones, S.J. (2013) Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics,* 29, 1492–1497 doi:10.1093/bioinformatics/btt178.

Bowe, L.M., Coat, G. and dePamphilis, C.W. (2000) Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proc. Natl Acad. Sci. USA,* 97, 4092–4097.

Cairney, J., Zheng, L., Cowels, A., Hsiao, J., Zismann, V., Liu, J., Ouyang, S., Thibaud-Nissen, F., Hamilton, J., Childs, K., Pullman, G.S., Zhang, Y., Oh, T. and Buell, C.R. (2006) Expressed sequence tags from loblolly pine embryos reveal similarities with angiosperm embryogenesis. *Plant Mol. Biol.* 62, 485–501.

Canales, J., Rueda-Lopez, M., Craven-Bartle, B., Avila, C. and Canovas, F.M. (2012) Novel insights into regulation of asparagine synthetase in conifers. *Front. Plant Sci.* 3, 100.

Canovas, F.M., Dumas-Gaudot, E., Recorbet, G., Jorrin, J., Mock, H.-P. and Rossignol, M. (2004) Plant proteome analysis. *Proteomics*, 4, 285–298.

Canovas, F.M., Avila, C., Canton, F.R., Canas, R.A. and de la Torre, F. (2007) Ammonium assimilation and amino acid metabolism in conifers. *J. Exp. Bot.* 58, 2307–2318.

Canton, F.R., Suarez, M.F. and Canovas, F.M. (2005) Molecular aspects of nitrogen mobilisation and recycling in trees. *Photosynth. Res.* 83, 265–278.

Chancerel, E., Lepoittevin, C., Le Provost, G., Lin, Y.C., Jaramillo-Correa, J.P., Eckert, A.J., Wegrzyn, J.L., Zelenika, D., Boland, A., Frigerio, J.-M., Chaumeil, P., Garnier-Gere, P., Boury, C., Grivet, D., Gonzalez-Martinez, S.C., Rouze, P., Van de Peer, Y., Neale, D.B., Cervera, M.T., Kremer, A. and Plomion, C. (2011) Development and implementation of a highly-multiplexed SNP array for genetic mapping in maritime pine and comparative mapping with loblolly pine. *BMC Genomics,* 12, 368. doi:10.1186/1471-2164-12-368.

Chancerel, E., Lamy, J.-B., Lesur, I., Noirot, C., Klopp, C., Ehrenmann, F., Boury, C., Le Provost, G., Label, P., Lalanne, C., Leger, V., Salin, F., Gion, J.-M. and Plomion, C. (2013) High-density linkage mapping in a pine tree reveals a genomic region associated with inbreeding depression and provides clues to the extent and distribution of meiotic recombination. *BMC Biol.* 11, 50. doi:10.1186/1741-7007-11-50.

Chen, J., Uebbing, S., Gyllenstrand, N., Lagercrantz, U., Lascoux, M. and Kallman, T. (2012) Sequencing of the needle transcriptome from Norway spruce (*Picea abies* Karst L.) reveals lower substitution rates, but similar selective constraints in gymnosperms and angiosperms. *BMC Genomics,* 13, 589. doi:10.1186/1471-2164-13-589.

Côté, C.L., Boileau, F., Roy, V., Ouellet, M., Levasseur, C., Morency, M.-J., Cooke, J.E.K., Seguin, A. and MacKay, J. (2010) Gene family structure, expression and functional analysis of HD-Zip III genes in angiosperm and gymnosperm forest trees. *BMC Plant Biol.* 10, 273. doi:10.1186/1471-2229-10-273.

Couturier, J., Montanini, B., Martin, F., Brun, A., Blaudez, D. and Chalot, M. (2007) The expanded family of ammonium transporters in the perennial poplar plant. *New Phytol.* 174, 137–150.

Durban, J., Juarez, P., Angulo, Y., Lomonte, B., Flores-Diaz, M., Alape-Giron, A., Sasa, M., Sanz, L., Gutierrez, J.M., Dopazo, J., Conesa, A. and Calvete, J.J. (2011) Profiling the venom gland transcriptomes of Costa Rican snakes by 454 pyrosequencing. *BMC Genomics,* 12, 259. doi:10.1186/1471-2164-12-259.

Falgueras, J., Lara, A.J., Fernandez-Pozo, N., Canton, F.R., Perez-Trabado, G. and Claros, M.G. (2010) SeqTrim: a high-throughput pipeline for pre-processing any type of sequence read. *BMC Bioinformatics*, 11, 38. doi:10.1186/1471-2105-11-38.

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

Fernandez-Pozo, N., Canales, J., Guerrero-Fernandez, D., Villalobos, D.P., Diaz-Moreno, S.M., Bautista, R., Flores-Monterroso, A., Guevara, M.A., Perdiguero, P., Collada, C., Cervera, M.T., Soto, A., Ordas, R., Canton, F.R., Avila, C., Canovas, M.G. and Claros, F.M. (2011) EuroPineDB: a high-coverage web database for maritime pine transcriptome. *BMC Genomics,* 12, 366. doi:10.1186/1471-2164-12-366.

Fickett, J.W. (1982) Recognition of protein coding regions in DNA-Sequences. *Nucleic Acids Res.* 17, 5303–5318.

Field, C.B., Behrenfeld, M.J., Randerson, J.T. and Falkowski, P. (1998) Primary production of the Biosphere: integrating terrestrial and oceanic components. *Science*, 281, 237–240.

Futamura, N., Totoki, Y., Toyoda, A., Igasaki, T., Nanjo, T., Seki, M., Sakaki, Y., Adriano, Mari.A., Shinozaki, K. and Shinohara, K. (2008) Characterization of expressed sequence tags from a full-length enriched cDNA library of *Cryptomeria japonica* male strobili. *BMC Genomics*, 9, 383. doi:10.1186/1471-2164-9-383.

Guillet-Claude, C., Isabel, N., Pelgas, B. and Bousquet, J. (2004) The Evolutionary implications of *knox-I* gene duplications in conifers: correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. *Mol. Biol. Evol.* 21, 2232–2245.

Harfouche, A., Meilan, R., Kirst, M., Morgante, M., Boerjan, W., Sabatti, M. and Scarascia Mugnozza, G. (2012) Accelerating the domestication of forest trees in a changing world. *Trends Plant Sci.* 17, 64–72.

Kinlaw, C.S. and Neale, D.B. (1997) Complex gene families in pine genomes. *Trends Plant Sci.* 2, 356–359.

Kolpakov, R., Bana, G. and Kucherov, G. (2003) mreps: efficient and flexible detection of tandem repeats in DNA. *Nucleic Acids Res.* 31, 3672–3678.

Koonin, E.V., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Krylov, D.M., Makarova, K.S., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Rogozin, I.B., Smirnov, S., Sorokin, A.V., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J. and Natale, D.A. (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5, R7.

Koski, L.B., Gray, M.W., Lang, B.F. and Burger, G. (2005) AutoFACT: an automatic functional annotation and classification tool. *BMC Bioinformatics*, 6, 151. doi:10.1186/1471-2105-6-151.

Leitch, A.R. and Leitch, I.J. (2012) Ecological and genetic factors linked to contrasting genome dynamics in seed plants. *New Phytol.* 194, 629–646.

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

Lepoittevin, C., Frigerio, J.M., Garnier-Gère, P., Salin, F., Cervera, M.-T., Vornam, B., Harvengt, L. and Plomion, C. (2010) In vitro vs. in silico detected SNPs for the development of a genotyping array: what can we learn from a non-model species? *PLoS ONE*, 5, 11034. doi:10.1371/journal.pone.0011034.

Lepoittevin, C., Harvengt, L., Plomion, C. and Garnier-Gère, P. (2012) Association mapping for growth, straightness and wood chemistry-traits in the *Pinus pinaster* Aquitaine breeding population. *Trees Genet. Genomes*, 8, 113–126.

Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics,* 25, 1754–1760.

Li, X., Wu, H., Dillon, S. and Southerton, S. (2009) Generation and analysis of expressed sequence tags from six developing xylem libraries in *Pinus radiata* D. Don. *BMC Genomics,* 10, 41. doi: 10.1186/1471-2164-10-41.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J. and Wang, J. (2010) De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20, 265–272.

Liang, F., Holt, I., Pertea, G., Karamycheva, S., Salzberg, S.L. and Quackenbush, J. (2000) An optimized protocol for analysis of EST sequences. *Nucleic Acids Res.* 28, 3657–3665.

Lijavetzky, D., Carbonero, P. and Vicente-Carbajosa, J. (2003) Genome-wide comparative phylogenetic analysis of the rice and *Arabidopsis* Dof gene families. *BMC Evol. Biol.* 3, 17. doi:10.1186/1471-2148-3-17.

Lin, Y., Li, J., Shen, H., Zhang, L., Papasian, C.J. and Deng, H.W. (2011) Comparative studies of de novo assembly tools for next-generation sequencing technologies. *Bioinformatics,* 27, 2031–2037.

Lorenz, W.W., Ayyampalayam, S., Bordeaux, J., Howe, G.T., Jermstad, K.D., Neale, D.B., Rogers, D.L. and Dean, J.D. (2012) Conifer DBMagic: a database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. *Trees Genet. Genom.* 8, 1477–1485.

Mackay, J., O'Malley, D.M., Presnell, T., Booker, F.-L., Campbell, M.M., Whethen, R.W. and Sederoff, R. (1997) Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase. *Proc. Natl Acad. Sci. USA,* 94, 8255–8260.

Mackay, J., Dean, J., Plomion, C., Peterson, D.G., Cánovas, F.M., Pavy, N., Ingvarsson, P.K., Savolainen, O., Guevara, M.Á., Fluch, S., Vinceti, B., Abarca, D., Díaz-Sala, C. and Cervera, M.T. (2012) Towards decoding conifer mega-genomes. *Plant Mol. Biol.* 80, 555–569.

Marth, G.T., Korf, I., Yandeli, M.D., Yeh, R.T., Gu, Z., Zakeri, H., Stitziel, N.O., Hiller, L., Kwok, P.-Y. and Gish, W.R. (1999) A general approach to single-nucleotide polymorphism discovery. *Nat. Genet.* 23, 452–456.

Martin, D.M., Fäldt, J. and Bohlmann, J. (2004) Functional characterization of nine norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol.* 135, 1908–1927.

Morse, A.M., Peterson, D.G., Islam-Faridi, M.N., Smith, K.E., Magbanua, Z., Garcia, S.A., Kubisiak, T.L., Amerson, H.V., Carlson, J.E., Nelson, C.D. and Davis, J.M. (2009) Evolution of Genome Size and Complexity in *Pinus*. *PLoS ONE*, 4, e4332.

Muller, T., Ensminger, I. and Schmid, K.J. (2012) A catalogue of putative unique transcripts from Douglas fir (*Pseudotsuga menziesii*) based on 454 transcriptome sequencing of genetically diverse, drought stressed seedlings. *BMC Genomics*, 13, 673. doi:10.1186/1471-2164-13-673.

Mullin, T.J., Andersson, B., Bastien, J.-C., Beaulieu, J., Burdon, R.D., Dvorak, W.S., King, J.N., Kondo, T., Krakowski, J., Lee, S.D., McKeand, S.E., Pâques, L., Raffin, A., Russell, J., Skrøppa, T., Stoehr, M. and Yanchuk, A. (2011) Economic importance, breeding objectives and achievements. In *Genetics, Genomics and Breeding of Conifers Trees* (Plomion, C., Bousquet, J. and Kole, C., eds), pp. 40–127, New York, NY: Edenbridge Science Publishers and CRC Press.

Neale, D.B. and Kremer, A. (2011) Forest tree genomics: growing resources and applications. *Nat. Rev. Genet.* 12, 111–122.

Nystedt, B., Street, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D.G., Vezzi, F., Delhomme, A., Vicedomini, R., Sahlin, K., Sherwood, E., Elfstrand, M., Gramzow, L., Holmberg, K., Hällman, J., Keech, O., Klasson, L., Koriabine, M., Kucukoglu, M., Käller, M., Luthman, J., Lysholm, F., Niittylä, T., Olson, A., Rilakovic, N., Ritland, C., Rossello, J.A., Sena, J., Svensson, T., Talavera-López, C., Theißen, G., Tuominen, H., Vanneste, K., Wu, Z-Q., Zhang, B., Zerbe, P., Arvestad, L., Bhalerao, R., Bohlmann, J., Bousquet, J., Gil, R.G., Hvidsten, T.R., de Jong, P., MacKay, J., Morgante, M., Ritland, K., Sundberg, B., Thompson, S.L., Van de Peer, Y., Andersson, B., Nilsson, B., Ingvarsson, P., Lundeberg, J., and Jansson, S. (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, doi:10.1038/nature12211.

Olson, J.S., Watts, J.A. and Allison, L.J. (1983) *Carbon in live vegetation of major world ecosystems*(Oak Ridge National Laboratory, Oak Ridge, TN), Report ORNL-. 5862.

Parchman, T.L., Geist, K.S., Grahnen, J.A., Benkman, C.W. and Buerkle, C.A. (2010) Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics*, 11, 180.

Park, Y., Lelu-Walter, M., Harvengt, L., Trontin, J., MacEacheron, I., Klimaszewska, K. and Bonga, J. (2006) Initiation of somatic embryogenesis in *Pinus banksiana*, *P. strobus*, *P. pinaster*, and *P. sylvestris* at three laboratories in Canada and France. *Plant Cell, Tissue Organ Cult.* 86, 87–101.

Pavy, N., Paule, C., Parsons, L., Crow, J., Morency, M.-J., Cooke, J., Johnson, J.E., Noumen, E., Guillet-Claude, C., Butterfield, Y., Barber, S., Yang, G., Liu, J., Stott, J., Kirkpatrick, R., Siddiqui, A., Holt, R., Marra, M., Seguin, A., Retzel, E., Bousquet, J. and MacKay, J. (2005) Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC Genomics*, 6, 144.

Pérez-Rodríguez, P., Riaño-Pachón, D.M., Guedes-Correa, L.G., Rensing, S.A., Kersten, B. and Mueller-Roeber, B. (2010) PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* 38 (Database issue), D822–D827.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J. and Pandolfi, P.P. (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumor biology. *Nature*, 465, 1033–1040.

Ralph, S., Chun, H., Kolosova, N., Cooper, D., Oddy, C., Ritland, C.E., Kirkpatrick, R., Moore, R., Barber, S., Holt, R.A., Jones, S.J., Marra, M.A., Douglas, C.J., Ritland, K. and Bohlmann, J. (2008) A conifer genomics resource of 200,000 spruce (*Picea* spp.) ESTs and 6,464 high-quality, sequence-finished full-length cDNAs for Sitka spruce (*Picea sitchensis*). *BMC Genomics*, 9, 484.

Resende, M.F.R., Muñoz, P., Resende, M.D.V., Garrick, D.J., Fernando, R.L., Davis, J.M., Jokela, E.J., Martin, T.A., Peter, G. and Kirst, M. (2012) Accuracy

of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). *Genetics*, 190, 1503–1510.

Rigault, P., Boyle, B., Lepage, P., Cooke, J.E.K., Bousquet, J. and Mackay, J. (2011) A white spruce catalogue for conifer genome analyses. *Plant Physiol.* 157, 14–28.

Ritland, K. (2012) Genomics of a phylum distant from flowering plants: conifers. *Tree Genet. Genomes,* 8, 573–582.

Rueda-López, M., García-Gutiérrez, A., Cánovas, F.M. and Avila, C. (2013) The family of Dof transcription factors in pine. *Trees Struct. Funct*. doi:10.1007/s00468-013-0903-z.

Sanz, F., Latour, S., Neves, M., Bastet, E., Pischedda, D., Pineiro, G., Gauthier, T., Lesbats, J., Plantier, C., Marques, A., Lanvin, J.-D., Santos, J.A., Touza, M., Pedras, F., Parrot, J., Reuliong, D. and Faria, C. (2006) *Industrial Applications of Pinus Pinaster*. 256 pp. Madeira: CIS Madeira, FIBA, AIMMP, CTBA.

Scaglione, D., Lanteri, S., Acquadro, A., Lai, Z., Knapp, S.J., Rieseberg, L. and Portis, E. (2012) Large-scale transcriptome characterization and mass discovery of SNPs in globe artichoke and its related taxa. *Plant Biotechnol. J.* 10, 956–969.

Soltis, D.E. and Soltis, P.S. (2003) The role of phylogenetics in comparative genetics. *Plant Physiol.* 132, 1790–1800.

Suzuki, S., Li, L., Sun, Y.-H. and Chiang, V.L. (2006) The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. *Plant Physiol.* 142, 1233–1245.

Trontin, J.F., Debille, S., Canlet, F., Harvengt, L. and Lelu-Walter, M.-A., Label, P., Teyssier, C., Miguel, C., De Vega-Bartol, J., Tonelli, M., Santos, R., Rupps, A., Hassani, S.B., Zoglauer, K., Carneros, E., Díaz-Sala, C., Abarca, D., Arrillaga, I., Mendoza-Poudereux, I., Segura, J., Avila, C., Rueda, M., Canales, J. and Cánovas, F.M. (2013) Somatic embryogenesis as an effective regeneration support for reverse genetics in maritime pine: the Sustainpine collaborative project as a case study. In: *Proceeding of the IUFRO Working Party 2.09.02 conference on "Integrating vegetative propagation, biotechnology and genetic improvement for tree production and sustainable forest management", 25-28/06/2012 (Brno, Czech Republic)*, (Park, Y.S. and Bonga, J.M., eds), pp 184–187. Published online (http://www.iufro20902.org/).

Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Déjardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjärvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leplé, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouzé, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y. and Rokhsar, D. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 313, 1596–1604.

UNECE (2013). *Forest product statistics 2007-2011*. Timber Bulletin ECE/TIM/BULL/65/2, Issue 2012, United Nations Economic Commission for Europe (UNECE), Trade and Sustainable Land management Division, http://www.unece.org/fileadmin/DAM/timber/statsdata/fps07-11.pdf.

Villalobos, D.P., Díaz-Moreno, S.M., El-Sayed, S.S., Cañas, R.A., Osuna, D., Van Kerckhoven, S.H., Bautista, R., Claros, M.G., Cánovas, F.M. and Cantón, F.R. (2012) Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. *BMC Plant Biol.* 12, 100. doi:10.1186/1471-2229-12-100.

Xu, Z., Zhang, D., Hu, J., Zhou, X., Ye, X., Reichel, K.L., Stewart, N.R., Syrenne, R.D., Yang, X., Gao, P., Shi, W., Doeppke, C., Sykes, R.W., Burris, J.N., Bozell, J.J., Cheng, M.Z., Hayes, D.G., Labbe, N., Davis, M., Stewart, C.N. Jr. and Yuan, J.S. (2009) Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. *BMC Bioinformatics,* 10, S3.

Supporting information

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

Figure S1 Distribution of GO terms among the annotated unigenes of *Pinus pinaster* transcriptome (Figure S1.pdf).

Figure S2 Hierarchical clustering analysis of TF families in plants (Figure S2.pdf).

Methods S1-Supplemental section on transcriptome assembly (Methods S1.pdf).

Table S1 A comparison of data set collection of transcription factor families present into the genomes of *Pinus pinaster, Picea glauca, Arabidopsis thaliana*, *Oryza sativa*, *Populus trichocarpa and Vitis vinifera*. (Table S1.docx).

Table S2 Maritime pine SNPs. (Table S2. Maritime pine SNP.txt).

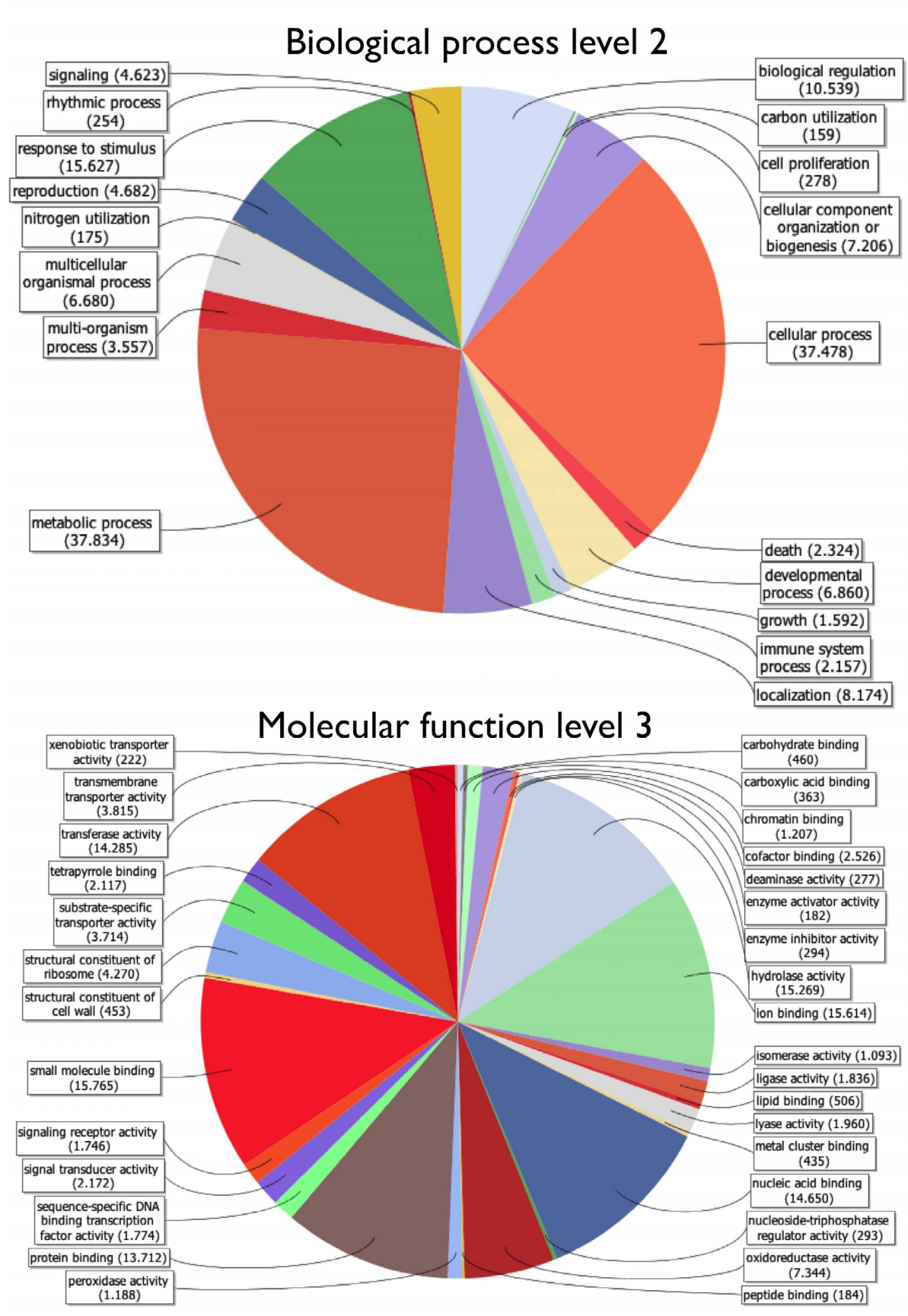


Figure S1 Distribution of GO terms among the annotated unigenes of *Pinus pinaster* transcriptome

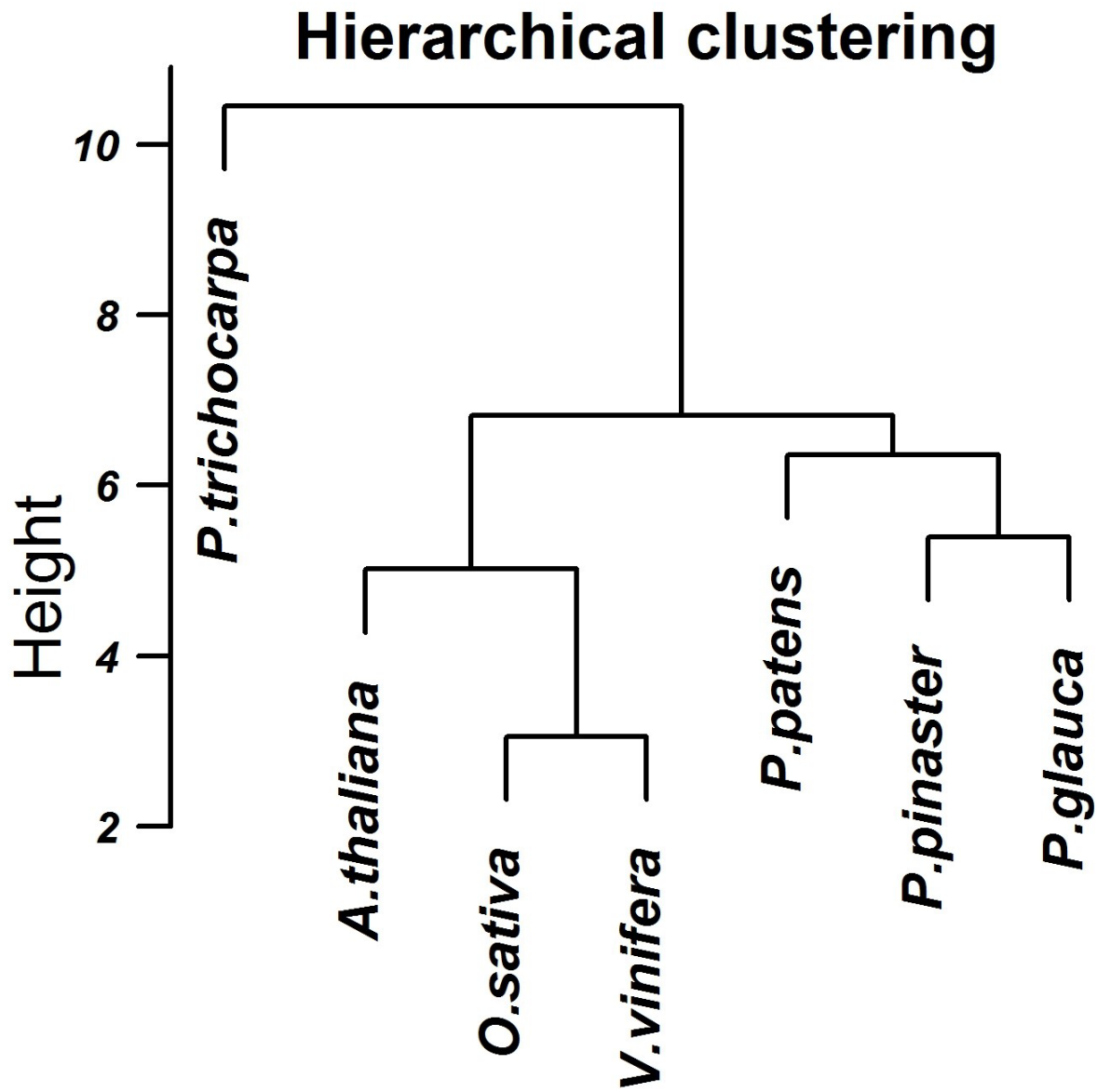


Figure S2 Hierarchical clustering analysis of TF families in plants

Methods S1
Supplemental section on transcriptome assembly

Short-reads and long-reads were following a different treatment. Long-read sequences were assembled with two different assemblers, MIRA3 (based on overlap-layout-consensus algorithm) and Euler-SR (based on a strict de Bruijn graph analyzed by an Eulerian path) using the default values. Due to memory and execution time limitations, each library was independently assembled with Euler-SR, while two big groups (A and B, see Figure 2) were provided for MIRA3 assembly. Original long-reads were mapped to Euler-SR contigs using Bowtie2 to discard artefactual contigs. «Debris» ones from MIRA3 and unmapped contigs from Euler-SR were analyzed with FLN (http://www.scbi.uma.es/fulllengthernext) to recover any coding read.

General guidelines for short-reads processing were well-established (Strickler *et al.,* 2012). Briefly, sequence cleaning from sequencing bias and overabundant sequences was followed by assembly using de Bruijn graph algorithm (i.e., ABySS) and overlap-based reconciliation (i.e., TGICL). Therefore, useful short-read sequences were pre-assembled separately as single- and paired-end in batches of four million sequences with ABySS (Birol *et al.,* 2009) and increasing k-mers from 43 to 73 bp by 6 bp steps. Default ABySS parameters were used, except for the minimum coverage of pairs that was set to 10. Useful pre-contigs should have a minimum of 100 bp in length. The pre-assembly procedure led to 444 pre-assemblies which were complexity-reduced using CD-HIT (Li and Godzik, 2006) with default parameters except for a word length of 10, a throw-away sequence length of 11, a both strands comparison, a clustering into the most similar cluster, a sequence identity threshold of 98% and an alignment coverage for the shorter sequence of 50% minimum. Reduced pre-assemblies were then meta-assembled using TGICL (Pertea *et al.,* 2003) with default parameters, with the exception of a 60 bp minimum overlap length, a 96% minimum identity for overlap and a 100,000 sequences cluster size limit. This meta-assembly from short-reads produced a set of 24,754 contigs ranging from 79 bp to 7,824 bp with an average length of 913 bp, an average coverage of 389X, and a median coverage of 182X. Only 0.04% nucleotides were covered three times or less.

Contigs separately obtained from short- and long-reads (355,483 in total) were reconciled using CAP3, resulting in 210,513 contigs to describe a *P. pinaster* transcriptome.

Notably, 1,241 unigenes were longer than 3000 nucleotides. Unigene length ranged from 40 to 7,876 bp, with an average of 495 bp and a median of 361 bp.

A BWA mapping was performed with unigenes to estimate the amount of original short-reads sequences finally placed into the assembly. Seventy-four percent of 77,262,636 filtered paired-end short-reads and 65% of the 215,333,910 filtered single-end short-reads matched unigenes, with only 2.4 % nucleotides having a coverage of three or less. Using BWA-SW for long-reads, 99.3% (4,069,232) of filtered sequences matched 202,303 unigenes (96.1%). Of 8,210 unmapped unigenes, 7,284 were less than 500 bp in length, 321 were only formed using Euler-SR assembler, 623 were only formed using the MIRA3 assembler and 6,319 were only formed from short-reads. Most unmapped unigenes with long-reads came from short-reads, which may have been due to low representation of embryonic tissues in long-read libraries, or the fact that short-reads and long-reads can produce different assemblies.

References

Birol, I., Jackman, S.D., Nielsen, C.B., Qian, J.Q., Varhol R. *et al.* (2009) De novo transcriptome assembly with ABySS. *Bioinformatics* **25**, 2872-2877.

Li, W. and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamicheva, S., Lee, Y., Cheung, F., Parvizi, B., Tsai, J. and Quakenbush, J. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* **19**, 651-652.

Strickler, S.R., Bombarely, A. and Mueller L.A. (2012) Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *Am. J. Bot.* **99**, 257-266.

Table S1 A comparison of data set collection of transcription factor families present into the genomes of *Pinus pinaster, Picea glauca, Arabidopsis thaliana, Oryza sativa, Populus trichocarpa and Vitis vinifera*.

Pfam ID	Pfam Symbol	*P. patens*	*P. pinaster*	*P. glauca*	*A. thaliana*	*O. sativa*	*P. trichocarpa*	*V. vinifera*
PF00010	HLH	101	67	51	138	90	158	102
PF00046	Homeobox	40	76	86	91	70	119	71
PF00072	Response_reg	5	6	19	42	30	95	41
PF00097	zf-C3HC4	45	137	188	311	236	421	218
PF00170	bZIP_1	37	39	31	73	76	73	42
PF00249	Myb-DNA-binding	62	122	122	263	189	385	250
PF00319	SRF-TF	22	4	45	108	42	109	54
PF00320	GATA	12	17	12	30	23	36	19
PF00412	LIM	3	9	5	13	10	20	10
PF00447	HSF_DNA-bind	8	21	8	24	23	31	19
PF00643	zf-B_box	8	3	6	28	18	38	18
PF00808	CBFD_NFYB_HMF	23	26	52	39	38	45	25
PF00847	AP2	149	106	63	145	113	209	93
PF02045	CBFB_NFYA	1	9	5	10	10	11	7
PF02362	B3	30	22	13	94	61	117	46
PF02365	NAC	32	39	43	113	94	162	70
PF02701	zf-Dof	20	10	5	36	24	41	22
PF03106	WRKY	37	38	25	72	79	104	58
PF03110	SBP	14	17	11	17	17	29	17
PF03514	GRAS	49	21	17	33	37	96	41
PF03634	TCP	6	10	7	24	18	34	15
PF03791	KNOX2	3	14	7	8	10	12	12
PF03859	CG-1	1	1	5	6	5	4	4
PF04640	PLATZ	13	6	10	12	10	20	12
PF04690	YABBY	0	4	3	6	7	13	7
PF04770	ZF-HD_dimer	8	17	19	17	15	23	10
PF05687	DUF822	6	2	5	8	6	11	6
PF06203	CCT	11	1	13	40	32	45	27
PF06507	Auxin_resp	13	31	14	23	22	35	19
PF08536	Whirly	0	2	2	3	2	3	2
PF08744	NOZZLE	0	0	0	1	0	0	0
TOTAL		802	877	892	1828	1407	2499	1337

Table S2 Maritime pine SNPs.

online access :

http://onlinelibrary.wiley.com/store/10.1111/pbi.12136/asset/supinfo/pbi12136-sup-0005-TableS2.txt?v=1&s=3e57b1ae3c0115a277667461c4279bd71c6db2c7

L'embryogenèse somatique : une méthode de multiplication végétative du pin maritime pour demain ?

FCBA et l'INRA ont engagé des recherches il y a maintenant près de 20 ans pour développer une méthode de multiplication végétative performante du pin maritime. Ses applications dans le programme d'amélioration génétique seraient multiples, depuis la gestion durable des ressources génétiques jusqu'à la sélection plus efficace des meilleures variétés et leur déploiement facilité dans les plantations. Suite aux développements pionniers chez l'épicéa auxquels FCBA a contribué (années 80, Afocel) et comme chez la plupart des autres conifères, c'est le processus d'embryogenèse somatique à partir de graines immatures couplé à la cryoconservation des embryons somatiques obtenus qui offre actuellement les meilleures perspectives pratiques. Des progrès importants ont été obtenus pour la maîtrise de cette technologie, particulièrement depuis la mise en place en 2004 (et continue depuis) d'une collaboration spécifique sur ce thème entre les équipes « Biotechnologie & Sylviculture Avancée » de FCBA et «Amélioration, Génétique et Physiologie Forestières » de l'INRA. Nous faisons ici un bilan des avancées majeures à différentes étapes de l'embryogenèse somatique jusqu'à la mise en place d'essais au champ en cours d'évaluation. Nous portons également un regard sur les verrous techniques et les contraintes socio-économiques qu'il faudrait lever pour amener cette technologie prometteuse à franchir les portes de l'application pratique.

La multiplication végétative : un outil précieux en amélioration génétique

Lorsqu'elles sont maîtrisées, les techniques de multiplication végétative (clonage) sont un atout pour l'amélioration génétique car elles permettent d'accéder à la sélection individuelle (tests clonaux) plutôt qu'uniquement familiale (tests de descendances). La plus grande efficacité de la sélection qui en résulte (utilisation de toute la variance génétique disponible, héritable et non héritable) permet d'optimiser les gains génétiques à chaque cycle d'amélioration ce qui en fait un avantage énorme pour les espèces à maturité sexuelle tardive comme les arbres forestiers. Par ailleurs cette capacité à réaliser des copies végétatives d'arbres sélectionnés facilite la mise en place des vergers de conservation des ressources génétiques (parcs à clones) ou de production des nouvelles variétés (vergers de clones). Il est également possible d'envisager une sylviculture « multivariétale » via le déploiement de variétés polyclonales composées d'un assemblage de génotypes testés répondant le mieux aux objectifs de production (optimisation du gain génétique) pour un niveau de diversité génétique déterminé (choisi) dans les plantations. Cette plus grande souplesse et précision dans la création variétale est particulièrement souhaitable pour adapter les nouvelles variétés à la demande et augmenter leur « résilience » aux aléas environnementaux (changement climatique, maladies, insectes, *etc*.).

La culture *in vitro* : un outil pour restaurer des capacités de multiplication végétative

Une telle stratégie est déjà à l'œuvre chez le peuplier pour lequel il n'y a pas de problème de multiplication végétative des arbres sélectionnés. Les variétés actuellement mises sur le marché sont effectivement produites chez les pépiniéristes par bouturage horticole classique (plançons). Cependant chez beaucoup d'espèces forestières et la plupart des conifères la récalcitrance des arbres matures au bouturage est un phénomène récurrent en lien avec le vieillissement ontogénétique (ou changement de phase depuis le stade juvénile jusqu'aux stades adultes végétatif et reproducteur). Il est souvent possible de restaurer des capacités organogénétiques perdues (vigueur juvénile, capacité d'enracinement des boutures, caractères morphologiques juvéniles, *etc*.) par des techniques de taille associées ou non avec des traitements hormonaux, de bouturage ou de greffage en cascade. Mais ces conditionnements *ex vitro* sont généralement longs et leur succès aléatoire (faible reproductibilité). L'apport des biotechnologies et en particulier de la culture *in vitro* a été déterminant dans beaucoup de situations. Cette culture d'explants (apex, bourgeons, méristèmes, fragments de tiges, feuilles, etc.) en conditions artificielle, axénique et miniature (*in vitro*) autorise un meilleur contrôle des paramètres environnementaux. Il est alors souvent

Copyrigh FCBA INFO, Septembre 2013

possible de revigorer voire rajeunir des arbres adultes qui récupèrent alors, au moins partiellement ou transitoirement, des capacités de multiplication végétative. C'est le cas par exemple pour les variétés d'eucalyptus résistantes au froid sélectionnées par FCBA (*E. gunnii*, *E. gundal*) qui bénéficient de l'appui du microbouturage pour produire des pieds-mères à forte valeur ajoutée pour la production de plants par bouturage chez les pépiniéristes (têtes de clones).

L'embryogenèse somatique (ES) : vers un support performant pour la multiplication végétative du pin maritime

Chez le pin maritime, la multiplication végétative d'arbres sélectionnés est actuellement limitée au greffage horticole. Le bouturage reste très difficile. Si le greffage est utilisé en routine dans le programme d'amélioration pour installer des parcs à clones (mise en collection) ou certains vergers de production (vergers de clones), son utilisation en sélection (tests clonaux) doit prendre en compte l'effet intrinsèque du porte-greffe sur l'évaluation des caractères du greffon. La recherche d'une technique de production de copies végétatives « entières » (sur leurs propres racines) mobilise de gros efforts tant à FCBA qu'à l'INRA depuis plus de 20 ans. Microgreffage de méristèmes et/ou microbouturage *in vitro* sont des techniques explorées à FCBA avec succès pour rajeunir et obtenir des copies d'arbres matures (Dumas et Monteuuis 1995, Trontin *et al.* 2004) mais leur efficacité est trop faible (taux de multiplication d'environ 1,2 par mois) pour envisager des applications autres qu'en recherche. C'est l'avènement dans les années 80 d'une technique de bouturage à « haut débit » de l'embryon zygotique contenu dans chaque graine qui a permis d'envisager la mise au point d'une stratégie de multiplication végétative performante de génotypes sélectionnés de pin maritime. Cette technique de clonage nommée « embryogenèse somatique » (ES, **Fig. 1**) mime la production d'embryons telle qu'elle se déroule dans les graines mais sans recours à la reproduction sexuée (du grec *sôma* = corps, ensemble des cellules non reproductrices par opposition au *germen*). L'ES consiste à entretenir un processus naturellement à l'œuvre dans chaque graine. A un stade immature du développement de l'embryon zygotique (cône vert) celui-ci est capable de se multiplier dans la graine pour produire 4 embryons génétiquement identiques (des jumeaux, **Fig. 1A**). Cette capacité disparait rapidement au cours de la maturation des graines et l'un de ces embryons devient dominant (le futur semis) tandis que les 3 autres dégénèrent. En excisant ces embryons immatures de la graine à un stade de développement déterminé (autour de la mise en place de la dominance) il est possible d'induire *in vitro* leur multiplication par un processus de formation continu de nouveaux embryons (polyembryonie de clivage = formation *de novo* + division de ces embryons) jusqu'à former une masse embryogène (étape d'initiation, **Fig. 1B**) contenant une multitude de copies conformes (un clone d'embryons immatures). Concrètement 1 g de masse embryogène peut contenir plus de 1000 embryons immatures ! Ces masses sont très faciles à propager avec des taux de multiplication pouvant atteindre 5 à 10 par semaine (étape de multiplication, **Fig. 1C**). Dans des conditions particulières de culture il est possible de provoquer le développement des embryons immatures en embryons somatiques cotylédonaires (étape de maturation, **Fig. 1D**) aptes à la germination (**Fig. 1E**) puis à la conversion en plants somatiques homologues à des semis (**Fig. 1F**).

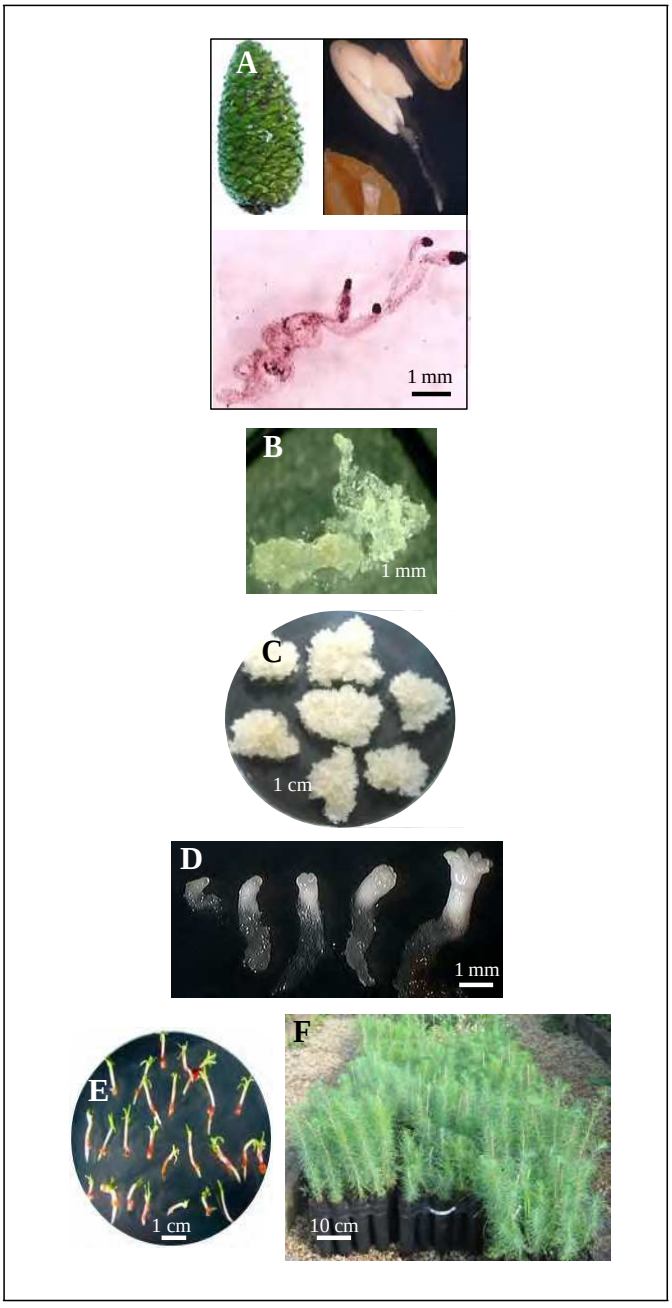


Figure 1. Embryogenèse somatique du pin maritime.

Ouverture des graines de cônes verts et excision de l'embryon zygotique (**A**). Initiation (**B**) et multiplication (**C**) des masses embryogènes. Développement des embryons (**D**), germination (**E**) et conversion en plants somatiques similaires au semis (**F**).

Copyrigh FCBA INFO, Septembre 2013

L'ES couplée à la cryoconservation : la clef pour multiplier rétroactivement des arbres matures

La multiplication d'un embryon de semis dont nous ne connaissons pas *a priori* les performances au champ (contrairement aux parents) n'aurait qu'un intérêt limité si les masses embryogènes n'étaient pas conservables à long terme par congélation à très basse température sous azote (-196°C). Dans un programme d'amélioration bien engagé comme celui du pin maritime (depuis les années 60, 4^{ième} cycle de sélection en cours), la possibilité de sécuriser les ressources génétiques en cryocollection dans un état très juvénile pendant toute la phase de test clonal sur le terrain (10 ans environ) permet de bénéficier rétroactivement d'un stock de tissu embryogène pour chacun des génotypes finalement sélectionnés pour leurs performances (**Fig. 2**). Il est alors possible d'envisager la multiplication de ces arbres élites par ES pour contribuer aux sorties variétales. C'est en ce sens que la multiplication végétative est actuellement un objectif crédible chez le pin maritime comme chez d'autres arbres forestiers en Europe (synthèse dans Lelu-Walter *et al.* 2013). Elle intègre au programme d'amélioration génétique en cours l'appui de deux biotechnologies, l'ES et la cryoconservation.

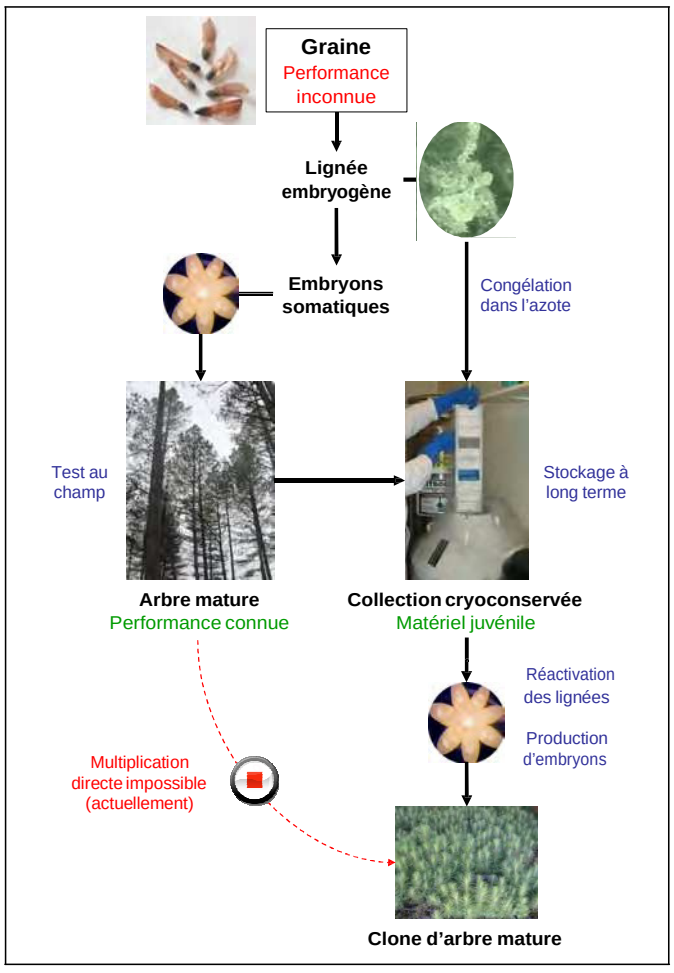


Figure 2. L'appui de l'ES et de la cryoconservation pour le clonage rétroactif d'arbres matures.

Que savons-nous faire exactement en ES du pin maritime ?

Depuis les premières données ayant démontré la possibilité d'initier l'ES chez le pin maritime (thèse de Jarlet-Hughes 1989) et de produire des plants somatiques (Lelu *et al.* 1999), les développements sont continus tant à FCBA qu'à l'INRA (**Tab. 1,** synthèse dans Klimaszewska *et al.* 2007) avec un partenariat formel sur ce thème entre les 2 instituts depuis 2004 doublé de plusieurs collaborations nationales (Univ. Paris VI, Paris VII et Nancy I) ou internationales (Canada : CFS, Suède : SLU, Espagne : Univ. Málaga et Oviedo). Ce programme est largement soutenu par la Région Aquitaine, la Région Centre, l'Etat, ainsi que par des fonds Européens. L'avancement des recherches autorise la mise en place d'essais au champ dès 1999 par FCBA. Les travaux ont longtemps porté sur l'initiation des masses embryogènes qui est maintenant considérée comme suffisamment performante pour ne pas induire de biais de sélection comme d'ailleurs les étapes de multiplication et de cryoconservation qui sont maîtrisées. La plupart des efforts concernent maintenant l'étape de maturation (maîtrise du développement embryonnaire). Il s'agit surtout d'augmenter la fréquence des lignées capables de produire des embryons somatiques avec des rendements élevés et avec une qualité suffisante pour germer correctement, c'est-à-dire avec une vigueur initiale similaire à celle de l'embryon zygotique. Etat des lieux aux étapes clefs …

Tableau 1. Développement de l'ES du pin maritime à FCBA et à l'INRA

Référence[a]		Collaboration	Thème[b]
1989	Jarlet-Hughes	FCBA/U. Paris VI	I
1995	Bercetche/Pâques	FCBA	I/M
1999	Lelu *et al.*	INRA/CFS	I/M/D/G/P
2001	Ramarosandra-tana *et al.*	FCBA/U. Paris VII	M/D/G
2003	Jordy/Favre	FCBA/U. Nancy I	D/G (qualité)
2005	Breton *et al.*	FCBA/U. Nancy I	M/D
2005	Harvengt	FCBA	I/M/D/G/P/E
2006	Perez-Rodriguez *et al.*	FCBA/U. Málaga/SLU	D (qualité)
2006	Lelu-Walter *et al.*	INRA/CFS	I/M/D/G/P
2006	Park *et al.*	FCBA/INRA/CFS	I
2006	Breton *et al.*	FCBA/U.Nancy I	M/D (qualité)
2007	Klimaszewska *et al.*	FCBA/INRA/CFS	I/M/D/G/P/E (synthèse)
2009	Klimaszewska *et al.*	INRA/CFS/U. Oviedo	M/D (qualité)
2011	Trontin *et al.*	FCBA/INRA	I/M/D/G/P/E

[a]Référence précise des articles dans Klimaszewska *et al.* (2007).
[b]**I** : initiation ; **M** : multiplication ; **D** : développement embryonnaire ; **G** : germination ; **P** : production de plants ; **E** : essais au champ.

Copyrigh FCBA INFO, Septembre 2013

Le taux de succès est élevé à l'étape d'initiation (77 %)

Suite à une collaboration entre FCBA, INRA et CFS (Canadian Forest Service) initiée en 2003 portant simultanément sur plusieurs espèces de pin (Park *et al.* 2006) nous avons pu mettre en évidence la nette supériorité d'une base de milieu de culture *in vitro* (mLV, milieu Litvay modifié) à cette étape chez le pin maritime (**Fig. 3**). De 2003 à 2009 le taux d'initiation moyen de l'ES obtenu à FCBA avec le milieu mLV pour une large gamme de lots de graines issus de croisements contrôlés (27) ou open (17) s'élève à 63,0 ± 11,2 % C'est 3 fois plus qu'avec la base de milieu utilisée de 2000 à 2005 (mDCR, 21,7 ± 7,6 %). Toujours dans le cadre de cette collaboration inter-laboratoires nous avons également révélé l'intérêt d'un régulateur de croissance (CPPU, une hormone de type cytokinine) en combinaison avec mLV (effets synergiques). Après détermination de la concentration optimale en CPPU chez le pin maritime nous avons pu améliorer le taux d'initiation moyen jusqu'à 76,8 ± 6,4 % (Trontin *et al.* 2011). Un tel résultat place le pin maritime parmi les conifères les plus performants à cette étape. L'effet famille reste significatif ainsi que la variation interannuelle mais avec un niveau aussi élevé de génotypes réceptifs au sein des familles le risque de biais (perte de diversité génétique) introduit par une éventuelle sélection sur le critère d'aptitude à l'ES apparaît à présent très limité. Cette étape assez longue (2-3 mois pour initier et stabiliser une lignée) est donc considérée comme maitrisée pour des applications pratiques. Elle nécessite toutefois de bien cibler la courte fenêtre de prélèvement (7-10 jours) des embryons zygotiques immatures au stade de développement optimal pour l'initiation (le stade précotylédonaire) qui se situe généralement courant juillet dans le massif aquitain.

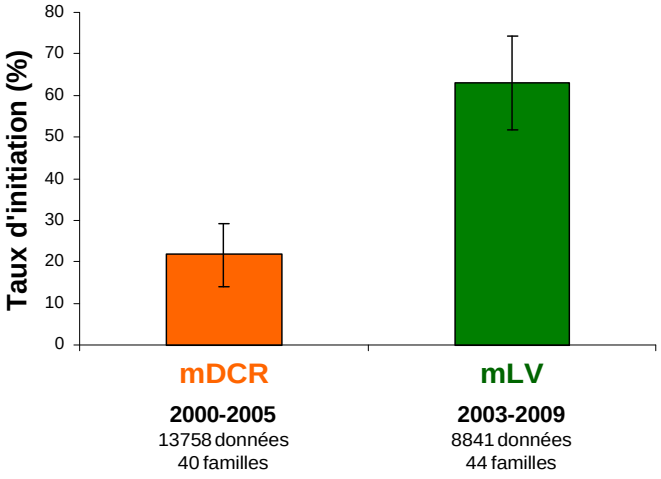


Figure 3. Le taux de succès à l'étape d'initiation (%) a considérablement augmenté en développant à partir de 2003 une base de milieu mLV plutôt que mDCR (barres = intervalle de confiance à 5 %).

Les masses embryogènes initiées sont aisément multipliées et cryoconservées

La multiplication des masses embryogènes est très rapide tant sur milieu mLV que mDCR contenant une auxine (2,4-D) et une cytokinine (BAP). La culture est possible au contact direct du milieu gélifié ou sur filtre de papier qui permet de simplifier le transfert des masses sur milieu neuf. Les accroissements relatifs moyens en matière fraîche estimés pour 6 lignées embryogènes de 3 familles cultivées sur filtre de papier avec un repiquage toutes les 2 semaines sont supérieurs à 1000 % avec un avantage pour mLV (1582 ± 152 %) comparé à mDCR (1268 ± 113 %). Sur la base d'une telle capacité de multiplication, un rapide calcul montre qu'il est possible d'obtenir en 6 semaines (3 repiquages) pour une lignée environ 1 kg de masse embryogène à partir d'un g initial. A raison de 1000 embryons immatures par g de masse embryogène (cf. point 5.3) cela représente un potentiel d'un million d'embryons somatiques ! Même si nous ne réalisons actuellement en pratique que 10 % de ce potentiel, la dimension industrielle du process par rapport aux techniques classiques est évidente. Cependant une chute progressive du potentiel embryogène est observée au cours du vieillissement des lignées en multiplication (généralement en moins de 6 mois). Ce phénomène semble bien spécifique aux pins car il apparaît beaucoup plus limité chez d'autres conifères (épicéa, mélèze). Des facteurs épigénétiques (qui ne modifient pas la structure de base de l'ADN) seraient à l'œuvre comme la méthylation de l'ADN, la modification de la structure de la chromatine ou l'expression de petits ARNs régulateurs. Ce problème est techniquement contourné via la bonne gestion à long terme des stocks cryoconservés de masses embryogènes dans un état juvénile (**Fig. 2**). Les masses embryogènes peuvent être facilement cryoconservées dans l'azote (-196°C) après un prétraitement d'un peu plus de 2 jours à l'aide de cryoprotectants et une descente en température contrôlée. Les très basses températures permettent de figer non seulement l'eau libre dans les cellules mais également l'eau liée aux molécules. Toute réaction métabolique est ainsi stoppée, en particulier les phénomènes de vieillissement. La décongélation de ces tissus ne pose pas de problème de survie (près de 100 % de réactivation). Il est donc possible d'établir des collections statiques de milliers d'individus dans un état juvénile au sein de cryo-conservateurs qui occupent quelques m² seulement de surface utile. Nous entretenons ainsi à FCBA une collection de plus de 2000 génotypes de pin maritime qui représentent une bonne quinzaine des meilleures familles issues de croisements contrôlés. Une telle collection peut être maintenue à très long terme sous vapeurs d'azote (plus de 15 ans de recul à FCBA). La bonne gestion des échantillons permet de garantir le niveau du stock cryoconservé pour chaque lignée embryogène. Le coût de maintien des collections est uniquement lié à l'alimentation en azote, ce qui représente environ 0,30 €/an/échantillon.

Copyrigh FCBA INFO, Septembre 2013

• Le rendement en embryons somatiques et le taux de génotypes réceptifs restent insuffisants à l'étape de maturation

L'intérêt du milieu mLV aux premières étapes de l'ES (initiation, multiplication) a été confirmé à l'étape de maturation. Si l'on considère les résultats obtenus entre 2000 et 2009 pour de nombreuses lignées, nous constatons (**Fig. 4**) que le rendement en embryons somatiques cotylédonaires s'est très nettement amélioré autour de 2004, date à laquelle nous avons progressivement opté pour mLV au lieu de mDCR. En moyenne nous obtenons 50 embryons/g de tissu embryogène avec mLV (547 lignées testées sur cette période). C'est près de 15 fois plus que sur mDCR (moins de 4 embryons/g, 1830 lignées testées).

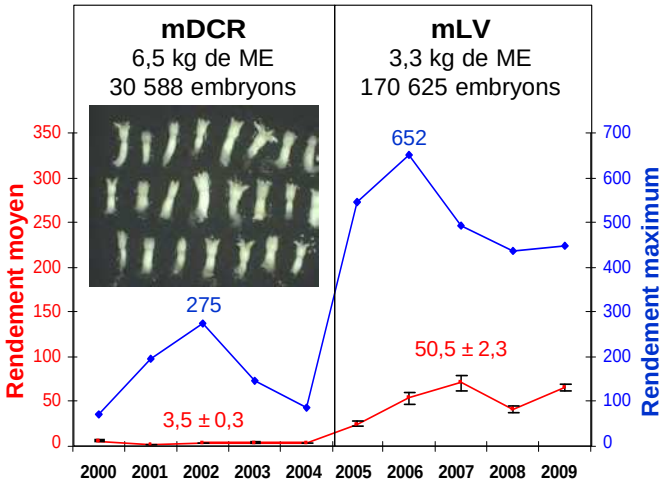


Figure 4. Le rendement en embryons somatique cotylédonaires (photo) par g de masse embryogène (ME) est beaucoup plus élevé sur milieu de maturation mLV comparé à mDCR. Moyenne (courbe rouge ; barres : intervalle de confiance à 5 %) et maximum annuel observé (courbe bleue).

Le taux de génotypes réceptifs à cette étape, estimé comme la fréquence des lignées capables de produire au moins 50 embryons/g est de 44 %. En considérant l'efficacité des étapes précédentes d'initiation (77 %) et de multiplication (100 %), la performance globale de l'ES se situe actuellement autour de 34 %. Ce chiffre reste trop faible mais suffisant pour envisager des applications pratiques. Nos meilleures lignées expriment des potentiels qui peuvent dépasser les 600 embryons/g (**Fig. 4**). C'est encore loin des potentiels estimés sur la base d'observation microscopiques (1000-5000 embryons immatures/g) mais il semble possible de s'en approcher. Les études portant sur notre génotype modèle (PN519) montrent en effet que la récalcitrance des lignées à la maturation serait surtout d'origine physiologique et/ou épigénétiques (interactions génotype x conditions de culture).

Les conditions de réalisation des étapes de l'embryogenèse au cours de la maturation sont maintenant relativement bien établies et reproduisent là encore ce qui est observé dans les graines. La transition des embryons immatures en multiplication dans les masses embryogènes vers le stade cotylédonaire apte à la germination est en effet provoquée par une réduction de la disponibilité en eau (milieu fortement gélifié), l'augmentation de la pression osmotique (milieu riche en saccharose) et l'ajout d'un signal hormonal (l'acide abscissique) qui est classiquement impliqué dans la régulation du développement embryonnaire. La durée du processus est de 12 semaines *in vitro.* C'est très comparable à ce qui est observé *in vivo* dans chaque graine (embryogenèse d'avril à juillet en Aquitaine).

Les embryons somatiques sont-ils capables de se développer normalement ?

• Ce que nous disent les essais au champ

Les équipes de FCBA et de l'INRA ont été capables de produire des embryons somatiques cotylédonaires dès 1999 (**Tab 1**, Lelu *et al.* 1999). Depuis, un objectif constant est d'estimer au travers d'essais au champ si les embryons somatiques se développent normalement, s'ils peuvent fructifier à l'âge mature et produire des graines. Huit essais ont été installés par FCBA (7) et l'INRA (1) avec des plants produits par FCBA ou en collaboration avec l'INRA depuis 2004 (**Tab. 2**). Cela représente plus de 3200 plants somatiques installés au champ. Ces essais testent pour la plupart des clones produits dans différentes conditions de culture *in vitro* comme les milieux mLV et mDCR.

Tableau 2. Tests au champ d'embryons somatiques produits par FCBA et l'INRA depuis 1999.

Plantation		Site	Nf*	Nc*	Np*
1999	FCBA (pépinière)	Sivaillan	12	12	59
2003	FCBA (pépinière)	Sivaillan	9	32	144
2004	FCBA (pépinière)	Sivaillan	7	35	99
2004	FCBA (forêt)	Landriole	9	25	448
2005	FCBA (forêt)	Landriole	10	25	667
2005	INRA (forêt)	Laperge	14	78	540
2008	FCBA (forêt)	Landriole	6	11	770
2012	FCBA (forêt)	Domaine de France	17	36	480

**Nombres de familles (Nf), clones (Nc) et plants (Np) à l'installation des essais.*

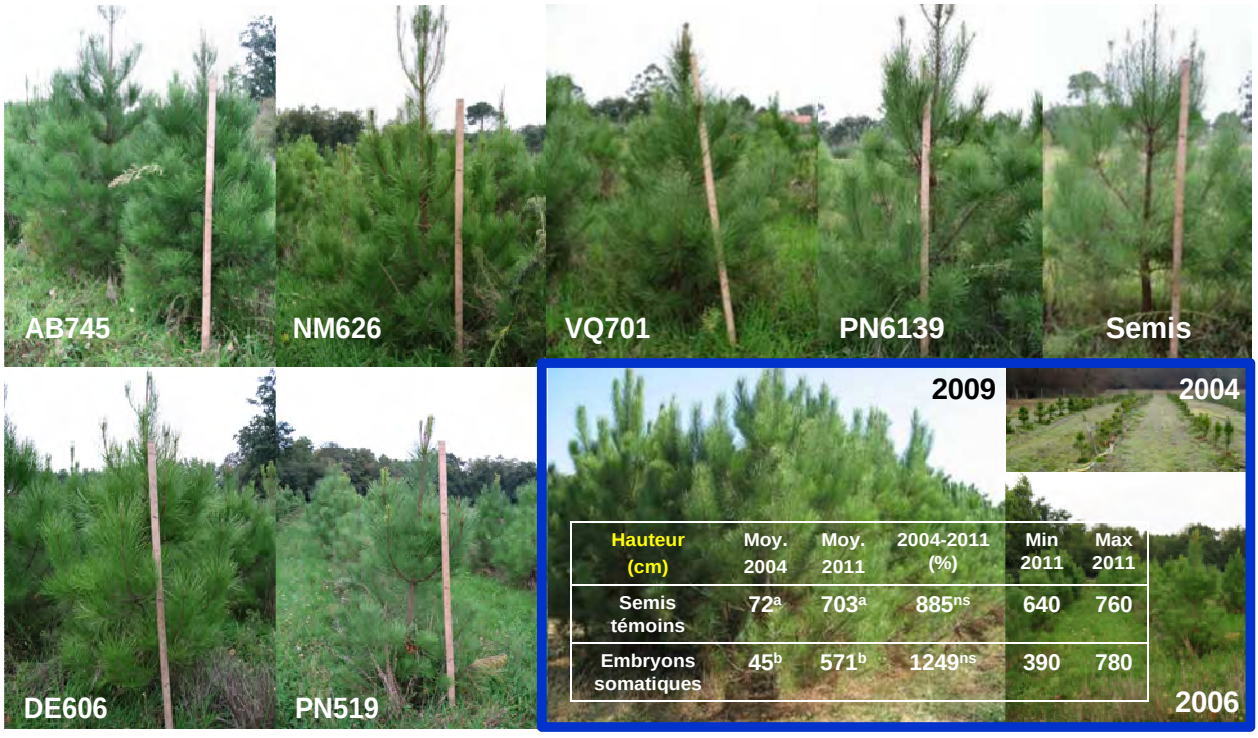
Copyrigh FCBA INFO, Septembre 2013

Le plus ancien essai, d'ampleur limité (12 clones FCBA, 59 plants en pépinière) date de 1999. Lors de la dernière mesure à 12 ans (2011) les arbres somatiques mesuraient de 6,50 à 12,10 m pour un diamètre de 9 à 29 cm. Les mensurations de certains arbres somatiques sont donc comparables aux semis témoins de la parcelle (hauteur : 9,60 à 11,20 ; diamètre : 19 à 33 cm). Les premières fructifications ont été notées dès l'âge de 5 ans pour les embryons et les semis témoin et elles sont à présent régulières. Cet essai démontre donc que les embryons somatiques sont capables d'exécuter les phases juvénile et adulte végétative tout à fait normalement et d'entrer dans la phase adulte reproductive (**Fig. 5**).



Figure 5. Arbres somatiques de 12 ans (essai 1999, Sivaillan) dans leur phase adulte reproductive.

Des essais plus conséquents ont été mis en place à partir de 2004-2005 (25 à 79 clones FCBA ou INRA, quelques centaines de plants selon les essais, **Tab. 2**). Ils intègrent comme référence des semis issus des familles correspondant aux clones (témoins génétiques) ou aux variétés VF1 (témoins améliorés). Ces essais montrent que divers clones peuvent présenter des morphologies comparables aux semis (**Fig. 6**). Les plants somatiques présentent néanmoins une plus faible vigueur initiale qui affecte les paramètres de productivité. Dans l'essai de Sivaillan (2004) planté en pépinière (**Fig. 6,** encart bleu), la hauteur moyenne des embryons à 7 ans (5,71 m) est ainsi significativement plus faible que celle des semis témoins (7,03). Cependant la croissance relative des embryons mesurée sur la même période (1249 %) est similaire à celle des semis (885 %). Si cette plus faible croissance se retrouve en situation forestière (essais de Landriole, **Fig. 7**), certains clones n'en montrent pas moins un très bon comportement (ex : 25C, 29C, DE737, CM815, ET826). Après 6 saisons sur le terrain, les accroissements relatifs des clones sont pour la plupart dans la moyenne des semis et même parfois supérieurs (ex : 29C, PN6128, NM626, NM18c). Les meilleurs clones pourraient donc être capables de rattraper leur retard. Ce résultat démontre la faisabilité et l'intérêt de la sélection clonale chez le pin maritime. Un projet FCBA récent soutenu par la Région Aquitaine a en particulier abouti à la plantation du premier test de comparaison de clones (essai 2012, **Tab. 2**). Bien sûr la moindre vigueur initiale ne permet pas actuellement de réaliser les gains génétiques espérés pour la voie clonale dans



Hauteur (cm)	Moy. 2004	Moy. 2011	2004-2011 (%)	Min 2011	Max 2011
Semis témoins	72[a]	703[a]	885[ns]	640	760
Embryons somatiques	45[b]	571[b]	1249[ns]	390	780

Figure 6. Aspect d'arbres somatiques (6 clones) et semis témoin 2,5 ans après plantation (essai 2004, Sivaillan). Encart bleu : vues générales de l'essai. Le tableau donne les hauteurs moyennes, minimales et maximales observées ainsi que l'accroissement relatif entre 2004 et 2011. [a,b]Différences significatives. [ns]non significatif.

Copyrigh FCBA INFO, Septembre 2013

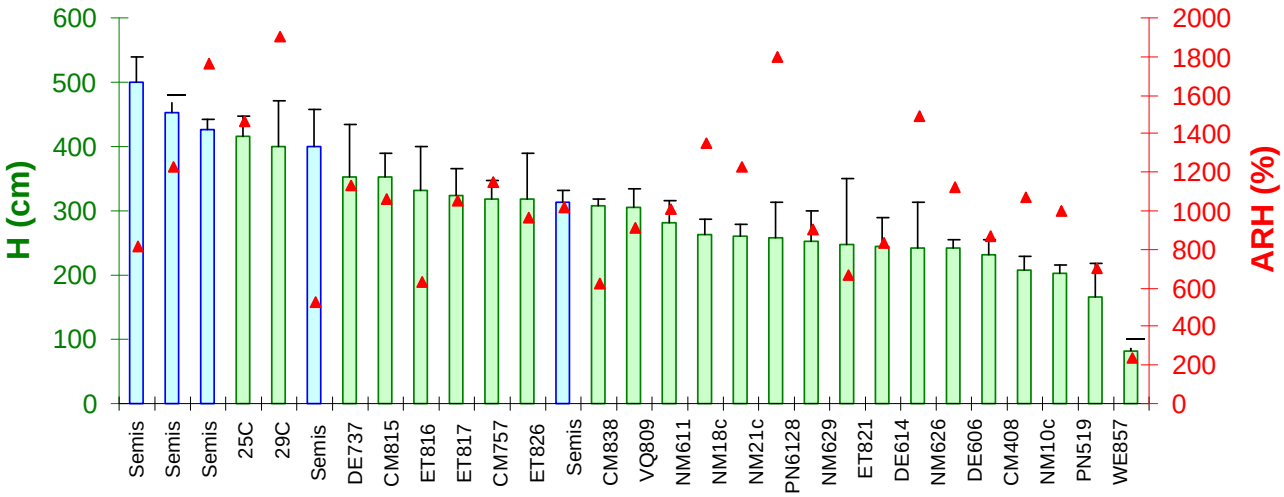


Figure 7. Moyenne des hauteurs (H) et accroissements relatifs en hauteur (ARH) à 6 ans de 24 clones (barres vertes) en situation forestière (essais de Landriole, 2004, 2005) en comparaison avec 5 lots de semis témoins (barres bleues). L'intervalle de confiance à 5 % est donné pour les hauteurs.

des délais standards. Mais dans l'hypothèse de la maîtrise du process de l'ES (recherches en cours à FCBA et à l'INRA) cette stratégie apparaît très prometteuse. Ces essais conduisent aussi actuellement à la conclusion contre-intuitive que l'homogénéité au sein des clones est plus faible que celle des lots de semis témoins. Là encore, la moindre vigueur initiale des clones dans l'état actuel de la technologie est une cause plausible. Alors que la forte vigueur initiale des semis tend à lisser les différences entre génotypes au démarrage d'un essai, l'effet inverse est obtenu pour les clones. Il faut probablement attendre plusieurs années pour que les effets génétiques soient visibles et vérifier que l'homogénéité intraclonale augmente avec le temps tandis qu'elle diminue chez les semis. De telles observations ont été faites en Nouvelle-Zélande (M. Menzies, Scion, com. personnelle) dans le cas du pin radiata pour lequel le processus d'ES entre à présent dans le champ commercial.

• Quelle origine au retard de développement des plants somatiques comparé au semis ?

Il faut chercher l'origine de la faible vigueur initiale des plants somatiques dans la qualité intrinsèque des embryons cotylédonaires produits. Leur comparaison avec les embryons zygotiques excisés de la graine et cultivés dans les mêmes conditions *in vitro* démontre leur plus faible capacité germinative (**Fig. 8**). Si les embryons somatiques ont bien terminé leur programme d'embryogenèse à l'issue de l'étape dite de « maturation » (morphologie similaire à celle des embryons zygotiques), ils ne sont probablement pas complètement matures. La recherche du déterminisme physiologique et moléculaire de la maturité des embryons est actuellement un axe de développement majeur à l'INRA et à FCBA, en particulier dans le cadre d'une thèse en cours (A. Morel) dirigée par M.-A. Lelu-Walter (INRA Orléans) et soutenue par la Région Centre. Porter la qualité des embryons somatiques au niveau de leur contrepartie zygotique, augmenter le taux de génotypes réceptifs à l'étape de maturation et exprimer au mieux l'énorme potentiel des lignées doivent à termes concourir à la mise en œuvre pratique de cette technologie chez le pin maritime.

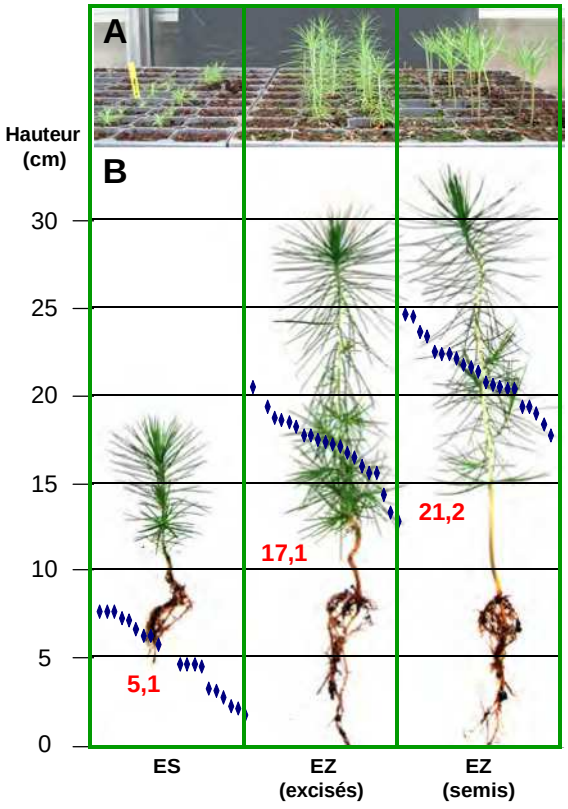


*Figure 8. Aspect des embryons somatiques (ES) et zygotiques (EZ) d'une même famille (NL) excisés ou non de la graine (semis) après 6 (**A**) et 22 semaines (**B**) de développement ex vitro.*

ES et EZ excisés ont germé 10 semaines in vitro puis ont été transférés ex vitro au moment du semis témoin. Le graphique B présente la variabilité (points bleus) et la hauteur moyenne (en rouge) de chaque catégorie de plants. Les différences observées sont significatives au seuil de 5%.

Copyrigh FCBA INFO, Septembre 2013

Quelles possibilités d'automatisation de l'ES pour réduire les coûts ?

Le potentiel indéniable de l'ES pour le déploiement industriel de nouvelles variétés offrant un gain génétique maximisé pour une diversité génétique choisie dans les plantations est bien compris à l'étranger (Canada, USA, Nouvelle-Zélande) où les échelles commerciales ont été atteintes par divers groupes forestiers privés (ArBorGen, CellFor, Forest Genetics) chez plusieurs conifères (ex : douglas, *Picea glauca*, *Pinus radiata*, *Pinus taeda*). En revanche l'intérêt reste faible en Europe même si des tentatives de développement commercial sont en cours pour l'épicéa commun (Suède), l'épicéa de Sitka (Irlande) et le Sapin de Nordmann (Danemark, production de sapins de Noël). Au-delà des raisons structurelles (forêt morcelée car largement privée, retour sur investissement long) et sociétales (acceptabilité des variétés clonales, réglementations) qui apparaissent comme majeures en Europe, c'est le coût actuellement élevé du plant somatique comparé au semis traditionnel (5 à 10 fois plus) qui freine son emploi chez les pins (Lelu-Walter *et al.* 2013). La manipulation de chaque embryon cotylédonaire depuis la récolte et le tri sous loupe en fin de maturation jusqu'à l'acclimatation en serre devient en effet rapidement fastidieuse. Des dispositifs industriels de maturation et de germination des embryons en milieu semi liquide (immersion temporaire) sont déjà employés avec succès pour la production de millions de plants somatiques chez des ligneux comme le caféier (CIRAD, Nestlé). Les solutions techniques retenues sont très loin des coûteux bioréacteurs industriels et s'apparentent à de simples poches plastiques jetables dans lesquelles les embryons évoluent en vrac. Chez les conifères et en particulier les pins, le développement des embryons n'est pas complètement synchrone et ils restent souvent liés entre eux et/ou solidaires des tissus embryogènes environnants. Il est donc nécessaire de séparer les embryons les uns des autres puis de les trier sur des critères morphologiques. Des robots « conventionnels » (à l'image des chaînes de production d'automobiles) capables de reproduire les gestes humains (reconnaître, trier, orienter, transférer des objets de quelques millimètres un par un) sont en développement depuis plusieurs années (ex : Find et Krogstrup 2009) mais peinent à démontrer leur efficacité. Une approche radicalement différente a été présentée récemment (Aidun et Egertsdotter 2012) dans laquelle l'automatisation est réalisée par homogénéisation en milieu liquide des tissus embryogènes contenant les embryons à récolter puis injection dans de fins canaux au diamètre approprié (« microfluidic system »). Des jeux de pression appliqués dans les canaux permettent de faire circuler rapidement les embryons (50 cm/s), de les séparer ou de les espacer régulièrement. Il est alors aisé de les positionner devant une caméra ultra-rapide qui inspecte taille et morphologie (critères prédéfinis) par analyse automatique d'images. Les embryons sélectionnés sont aiguillés vers un circuit de reconnaissance et d'orientation des pôles racinaires et caulinaires qui aboutit à leur mise en germination sur milieu solide. Deux lignes de production pilotes ont été installées en Suède récemment (SweeTree Technologies) pour des essais concluants chez l'épicéa commun (1 embryon planté par seconde). Un tel système hydraulique est beaucoup plus simple et abordable que l'approche robotisée. Breveté en 2011 (C. Aidun, USA), il pourrait être disponible sous licence prochainement (dès 2014-2015) pour adaptation à d'autres espèces.

Et la multiplication directe d'arbres matures par ES ?

Une telle perspective permettrait l'accélération de la création variétale mais les conifères sont largement récalcitrants à ce type de traitement. C'est donc une éventualité encore très prospective chez le pin maritime. Son intérêt est néanmoins capital si on considère la nécessité dans un futur proche d'augmenter la fréquence et la diversité des sorties variétales pour mieux faire face à la demande en bois dans un contexte d'augmentation des aléas biotiques et abiotiques. Elle permettrait de gagner 10-12 ans (soit la durée d'un cycle de sélection) comparé à la stratégie actuelle d'ES à partir de graines immatures couplée à la cryoconservation (**Fig. 2**). Si celle-ci s'intègre parfaitement dans les programmes d'amélioration bien rodés, c'est moins le cas lorsqu'il s'agit de démarrer des programmes ou d'intégrer de nouveaux critères de sélection. Des progrès récents chez *Picea glauca* (Klimaszewska *et al.* 2011) montrent qu'il est possible de réaliser l'ensemble du processus de l'ES à partir de bourgeons d'arbres âgés d'une dizaine d'année. Nous avons obtenu à FCBA des données comparables chez *Picea abies* à partir d'arbres de 3 ans (Harvengt *et al.* 2001). A ceci près, dans les 2 cas, qu'il s'agit d'arbres somatiques (= obtenus par ES, propagation végétative). Etrangement, il n'y a toujours pas de démonstration tangible de cette possibilité à partir d'arbres zygotiques (= obtenus par semis, reproduction sexuée). Ces résultats ont le mérite de suggérer que le processus d'ES pourrait réactiver des capacités organogénétiques caractéristiques de l'état juvénile chez les conifères comme cela vient d'être démontré pour des chênes centenaires. FCBA et l'INRA prospectent cette possibilité chez le pin maritime dans le cadre de collaborations internationales avec les meilleures équipes du domaine intéressées par les pins au Canada, en Nouvelle-Zélande et en Finlande. Des tissus embryogènes ont pu être obtenus à partir de coupes de bourgeons chez les pins sylvestre et contorta (arbres de 10 à 20 ans) sans toutefois parvenir à régénérer des plants somatiques. Chez le pin maritime nous avons pu conclure à FCBA (données non publiées, Trontin JF, Quoniou S) qu'il est possible d'obtenir des structures

Copyrigh FCBA INFO, Septembre 2013

embryoïdes similaires à des embryons à partir de coupes de bourgeons d'arbres de 35 ans. Même si leur multiplication n'est pas encore possible, ces résultats laissent entrevoir la possibilité d'ES à partir de tissus de pins maritimes matures.

Conclusion

Grâce aux efforts continus des équipes de recherche à FCBA et l'INRA depuis plus de 20 ans, une méthode de multiplication végétative performante par ES est envisageable à terme pour le pin maritime. Les collaborations soutenues entre les deux instituts et à l'international depuis 2004 ont permis de réaliser des progrès considérables en ce sens. Les étapes initiales (initiation, multiplication) de l'ES ainsi que la cryoconservation des tissus embryogènes sont maintenant bien maîtrisées. Les essais au champ conduits depuis 1999 démontrent que les plants somatiques régénérés à partir de ces tissus peuvent franchir les différentes étapes du développement normal d'un semis jusqu'à la fructification. Mais ils montrent également la moindre vigueur initiale de ce type de plant qui révèle un défaut de maturité des embryons somatiques à la récolte. Les recherches en cours visent donc surtout à mieux maîtriser l'étape de maturation des embryons, tant quantitativement (taux de génotypes réceptifs, expression de l'énorme potentiel embryogène des lignées) que qualitativement (amélioration de la capacité germinative et de la croissance initiale). C'est à cette condition que l'ES du pin maritime sera techniquement prête à affronter les questions socio-économiques (coût, marché, acceptabilité, gestion du gain et de la diversité dans les plantations).
Lorsqu'elles sont parfaitement maîtrisées et bon marché, les techniques de multiplication végétative comme l'ES sont un atout considérable pour l'amélioration génétique, la création et le déploiement rapide de nouvelles variétés chez les arbres forestiers. C'est pour favoriser leur développement et leur application qu'un groupe IUFRO mondial (http://www.iufro.org/science/divisions/division-2/20000/20900/20902/) a été créé par YS Park (CFS, Canada) en 2009 et auquel participent activement FCBA (co-coordination) et l'INRA. Les partenariats qui se nouent à l'occasion des manifestations de ce groupe (prochain meeting :Vitoria-Gasteiz, 9-14/09/2014) contribuent à accélérer la mise au point de l'ES.

Jean-François Trontin, Francis Canlet, Isabelle Reymond, Sandrine Debille, Karine Durandeau, Luc Harvengt, Jean-Pierre Rousseau, Jean-Mathieu de Boisseson, Jean-Yves Fraysse, Pierre Alazard, Alain Bailly (FCBA)

Caroline Teyssier, Claire Le Metté, Alexandre Morel, Philippe Label, Marie-Anne Lelu-Walter (INRA)

Bibliographie

1. Aidun CK, Egertsdotter EMU (2012). Fluidics-Based automation of clonal propagation via somatic embryogenesis: SE-fluidics system. 2nd International conference of the IUFRO working party 2.09.02, June 25-28, Brno, Czech Republic.
2. Dumas E, Monteuuis O (1995). *In vitro* rooting of micropropagated shoots from juvenile and mature *Pinus pinaster* explants: influence of activated charcoal. Plant Cell Tissue & Organ Culture 40: 231–235.
3. Find J, Krogstrup P (2009). Integration of biotechnology, robot technology and visualisation technology for development of methods for automated mass production of elite trees. Working Papers of the Finnish Forest Research Institute 114: 72-77
4. Harvengt L, Trontin JF, Reymond I, Canlet F, Pâques M (2001). Molecular evidence of true-to-type propagation of a 3-year-old Norway spruce through somatic embryogenesis. Planta 213: 828-832
5. Klimaszewska K, Trontin JF, Becwar M, Devillard C, Park YS, Lelu-Walter MA (2007). Recent progress on somatic embryogenesis on four *Pinus* spp. Tree Forest Science & Biotechnology 1: 11-25.
6. Klimaszewska K, Overton C, Stewart D, Rutledge RG (2011). Initiation of somatic embryos and regeneration of plants from primordial shoots of 10-year-old somatic white spruce and expression profiles of 11 genes followed during the tissue culture process. Planta 233: 635-647.
7. Lelu MA, Bastien C, Drugeault A, Gouez ML, Klimaszewska K (1999). Somatic embryogenesis and plantlet development in *Pinus sylvestris* and *Pinus pinaster* on medium with and without growth regulators. Physiol. Plant. 105: 719-728.
8. Lelu-Walter MA, Thompson D, Harvengt L, Sanchez L, Toribio M, Pâques L (2013). Somatic embryogenesis in forestry with a focus on Europe: state-of-the-art, benefits, challenges and future direction. Tree Genetics & Genomes 9: 883-899.
9. Park YS, Lelu MA, Harvengt L, Trontin JF, Mac Eacheron I, Klimaszewska K, Bonga JM (2006). Initiation of somatic embryogenesis in *P. banksiana*, *P. strobus*, *P. pinaster* and *P. sylvestris* at three laboratories in Canada and France. Plant Cell Tissue and Organ Culture 86: 87-101.
10. Trontin JF, Reymond I, Quoniou S, Canlet F, Debille S, Bruneau G, Harvengt L, Le Metté C, Vallance M, Teyssier C, Label P, Lelu-Walter MA (2011). An overview of current achievements and shortcomings in developing maritime pine somatic embryogenesis and enabling technologies in France. IUFRO 2.09.02 proceedings on Advances in Somatic Embryogenesis of Trees and Its Application for the Future Forests and Plantations (Eds : YS Park, JM Bonga, SY Park, HK Moon), 19-21/08/2010 (Suwon, South Korea), pp 100-102.
11. Trontin JF, Alazard P, Dumas E, Quoniou S, Canlet F, Chantre G, Harvengt L (2004). Prospects for clonal propagation of selected maritime pine (*Pinus pinaster* Ait.) using micropropagation techniques. 9th Conference on Biotechnology in the Pulp & Paper Industry (Durban, South Africa), pp. 9.5.

Ces développements ont reçu depuis 1999 le soutien de la Région Aquitaine, de la Région Centre, du Ministère de l'Agriculture et de la Forêt ainsi que de l'Union Européenne

Contact :

Jean-François TRONTIN
Tél. 05 56 79 95 03
jean-francois.trontin@fcba.fr

FCBA – Pôle Biotechnologie & Sylviculture Avancée
Equipe Génétique & Biotechnologie
71, Route d'Arcachon – Pierroton, 33610 Cestas

Copyrigh FCBA INFO, Septembre 2013

Annexe 4 : Composition des milieux de cultures

	Milieu de Base* (1/2 Litvay's medium)	Milieu de Multiplication*	Milieu de Maturation*
	(Litvay et al., 1985)	(Lelu-Walter et al., 1999)	(Lelu-Walter et al., 1999)
KNO3	0,95 g	0,95 g	0,95 g
NH4NO3	0,825 g	0,825 g	0,825 g
CaCl2, 2H2O	0,011 g	0,011 g	0,011 g
MgSO4, 7H2O	0,925 g	0,925 g	0,925 g
KH2PO4	0,170 g	0,170 g	0,170 g
Fer séquestrène (EDTA)	280 mg	280 mg	280 mg
KI	4,15 mg	4,15 mg	4,15 mg
H3BO3	31 mg	31 mg	31 mg
MnSO4, H2O	21 mg	21 mg	21 mg
ZnSO4, 7H2O	43 mg	43 mg	43 mg
Na2MoO4, 2H2O	1,25 mg	1,25 mg	1,25 mg
CuSO4, 5H2O	0,5 mg	0,5 mg	0,5 mg
CoCl2, 6H2O	0,13 mg	0,13 mg	0,13 mg
Acide Nicotinique	0,5 mg	0,5 mg	0,5 mg
Pyridoxine HCl	0,1 mg	0,1 mg	0,1 mg
Thiamine HCl	0,1 mg	0,1 mg	0,1 mg
Myo-inositol	100 mg	100 mg	100 mg
Hydrolysat de caséine		1 g	1 g
L-glutamine		0,5 g	0,5 g
Saccharose		20 g (0,06 M)	68,4 g (0,2 M)
2,4-D		0,5 mg (2,3 µM)	–
BA		0,5 mg (2,2 µM)	–
ABA		–	21,12 mg (80 µM)
pH		5,8	5,8
Phytagel		4 g	9 g

*Concentration par litre de milieu

Alexandre MOREL

Physiologie moléculaire du développement des embryons somatiques de pin maritime (Pinus pinaster Ait.) : approches transcriptomique et protéomique

Chez le pin maritime l'embryogenèse somatique, méthode de multiplication végétative performante, n'est pas optimisée la limite étant le contrôle du développement des embryons somatiques (ES). Nos objectifs ont été **1)** d'étudier les mécanismes physiologiques, cellulaires et moléculaires précoces contrôlant la différenciation des ES en réponse à une disponibilité en eau réduite **2)** d'évaluer pour l'ES cotylédonaire de 12 semaines son état de maturité, son protéome afin de les comparer à l'embryon zygotique (EZ).
1) Pour le 1er objectif, le rapprochement de données de transcriptomique et de protéomique a été entrepris. Nous avons observé une réponse physiologique et moléculaire ABA-dépendante entrainant une transition précoce de la prolifération vers la différenciation des ES (surexpression de protéines impliquées dans la division cellulaire, l'embryogenèse et la synthèse de l'amidon). Une protéine de type germine et une ubiquitine ligase apparaissent comme marqueurs potentiels de l'embryogenèse somatique précoce du pin maritime, alors que la protéine phosphatase 2C marque la réponse adaptative à l'environnement de culture.
2) La maturité de l'ES cotylédonaire a été étudiée aux niveaux physiologiques (masse sèche, teneur en eau) et biochimiques (teneur en protéines totales, en sucres solubles). Des ES de 10, 12 ou 14 semaines se révèlent semblables. Une méthode de classification hiérarchique ascendante basée sur 9 variables explicatives, montre que l'ES est similaire à l'EZ cotylédonaire frais (protéomes présentant 94% d'homologie). Parmi les protéines communes, 3 familles ont été identifiées (protéines de réserve, protéines de réponse au stress HSP, protéines LEA) ainsi que l'aldose réductase et l'adénosine kinase. Nous les proposons comme marqueurs génériques du stade cotylédonaire de l'embryogenèse tardive du pin maritime. L'ensemble des résultats contribue à une meilleure compréhension de l'embryogenèse somatique du pin maritime.
Mots clés : Embryogenèse somatique, maturation, marqueurs du développement, protéomique, transcriptomique, pin maritime, embryon zygotique

Molecular physiology of somatic embryo development of maritime pine (Pinus pinaster Ait.): proteomic and transcriptomic approaches

In maritime pine somatic embryogenesis, a powerful method of vegetative propagation, is not optimized the limitation being the control of the somatic embryo (SE) development. Our objectives were **1)** To study the early physiological, cellular and molecular mechanisms controlling SE differentiation in response to a reduced water availability **2)** to estimate for cotyledonary SE 12 weeks old, its maturity, its proteome to compare them to the zygotique embryo (ZE).
1) For the 1st objective, transcriptomic and proteomic data were combined. We observed a physiological and molecular answer ABA-dependent, inducing an early transition from proliferation towards SE differentiation (surexpression of proteins involved in the cellular division, the embryogenesis and in starch synthesis). A protein of germin type and an ubiquitine ligase appear as potential markers of the early somatic embryogenesis of maritime pine, while the phosphatase protein 2C stands out the adaptive answer to the environment of culture.
2) Cotyledonary SE maturity was studied at the physiological (dry weight, water content) and biochemical levels (total protein content, soluble sugars). SE of 10, 12 or 14 weeks old appeared very similar. A hierarchical ascendant cluster analysis based on 9 explanatory variables, shows that SE is similar to the fresh cotyledonary EZ; proteome profiling further confirmed high similarity (94.5%) between them. Protein profiling revealed biomarkers belonging to 3 large families of proteins (HSP, LEA and storage proteins) or the aldose reductase and the adenosine kinase. We propose them as generic markers of the cotyledonary stage of the late embryogenesis of the maritime pine.
All the results contribute to a better understanding of the somatic embryogenesis of maritime pine.
Keywords : somatic embryogenesis, maturation, development molecular markers, proteomic, transcriptomic, maritime pine, zygotic embryo

LABORATOIRE INRA Centre Val de Loire
Unité Amélioration, Génétique et Physiologie Forestières
2163, avenue de la pomme de pin -CS 40001 Ardon