



HAL
open science

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

Alexandre Morel, Caroline Teyssier, Jean-François Trontin, Kateřina Eliášová, Bedřich Pešek, Martine Beaufour, Domenico Morabito, Nathalie Boizot, Claire Le Metté, Leila Belal-Bessai, et al.

► **To cite this version:**

Alexandre Morel, Caroline Teyssier, Jean-François Trontin, Kateřina Eliášová, Bedřich Pešek, et al.. Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water availability: transcriptomic and proteomic analyses. *Physiologia Plantarum*, 2014, 152 (1), pp.184-201. 10.1111/ppl.12158 . hal-01179486

HAL Id: hal-01179486

<https://hal.science/hal-01179486v1>

Submitted on 27 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



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

Journal:	<i>Physiologia Plantarum</i>
Manuscript ID:	PPL-2013-00496.R1
Manuscript Type:	Regular manuscript - Development, growth and differentiation
Date Submitted by the Author:	n/a
Complete List of Authors:	Lelu-Walter, M.-A.; INRA, MOREL, Alexandre; INRA, UR 0588 AGPF Teyssier, Caroline; INRA, Unit for Breeding, Genetics and Physiology of Forest trees Trontin, Jean-Francois; FCBA, Eliášová, Kateřina; Institute of Experimental Botany, ASCR Pešek, Bedřich; Institute of Experimental Botany, ASCR Beaufour, Martine; CNRS, Morabito, Domenico; Univ Orleans, UFR Fac Sci BOIZOT, Nathalie; INRA, UR 0588 AGPF REYMOND, Isabelle; FCBA, HARVENGT, Luc; FCBA, CADENE, Martine; CNRS, Corbineau, Françoise; UPMC, Vagner, Martin; Institute of Experimental Botany, ASCR Label, Philippe; INRA, UMR_A 547 PIAF
Key Words:	somatic embryo development, germin-like protein, glycolysis, protein phosphatase 2C, ubiquitin-protein ligase

SCHOLARONE™
Manuscripts

Comment citer ce document :

Morel, A., Teyssier, C., Trontin, J.-F., Eliášová, K., Pešek, B., Beaufour, M., Morabito, D., Boizot, N., Le Mette, C., Belal-Bessai, L., Reymond, I., Harvengt, L., Cadene, M., Corbineau, F., Vágner, M., Label, P., Lelu-Walter, M.-A. (Auteur de correspondance) (2014). Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

3
4 Alexandre Morel¹, Caroline Teyssier¹, Jean-François Trontin², Kateřina Eliášová³, Bedřich Pešek³,
5 Martine Beaufour⁴, Domenico Morabito⁵, Nathalie Boizot¹, Claire Le Metté¹, Leila Belal-Bessai¹,
6 Isabelle Reymond², Luc Harvengt², Martine Cadene⁴, Françoise Corbineau⁶, Martin Vágner³, Philippe
7 Label^{1†**}, Marie-Anne Lelu-Walter^{1†*}

8
9 1: INRA, UR 0588 Unité Amélioration, Génétique et Physiologie Forestières, 2163 Avenue de la
10 Pomme de Pin, CS 4001, Ardon, F-45075 Orléans Cedex 2, France

11 2 : FCBA, Pôle Biotechnologie et Sylviculture Avancée, Equipe Génétique et Biotechnologie, Campus
12 Forêt-Bois de Pierroton, 71 route d'Arcachon, F-33610 Cestas, France.

13 3 : Institute of Experimental Botany ASCR, Rozvojová 263, Praha 6-Lysolaje, 165 02, Czech
14 Republic.

15 4 : CNRS, UPR4301, Centre de Biophysique Moléculaire, Rue Charles Sadron, F-45100 Orléans,
16 France.

17 5: Univ. Orléans, INRA, LBLGC , EA1207, F-45067, Orléans, France

18 6: Sorbonne Universités, UPMC Univ Paris 06, UMR 7622, Biologie du Développement, 4 place
19 Jussieu, F-75005, Paris, France.

20
21 † These authors contributed equally to this work

22
23 * Corresponding author: Marie-Anne Lelu-Walter

24 Phone : +33 (0) 238417800

25 Fax : +33 (0) 238417879

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

27
28 ** Present address : INRA-Université Blaise Pascal, UMR_A 547 PIAF, Les Cézeaux, 24 Avenue des
29 Landais, 63177 Aubière cedex, France

1 Summary

2 Maritime pine somatic embryos require a reduction in water availability (high gellan gum
3 concentration in the maturation medium) to reach the cotyledonary stage. This key switch, reported
4 specifically for pine species, is not yet well understood. To facilitate the use of somatic embryogenesis
5 for mass propagation of conifers, we need a better understanding of embryo development. Comparison
6 of both transcriptome (Illumina RNA sequencing) and proteome (2D-SDS-PAGE with MS
7 identification) of immature somatic embryos, cultured on either high (9G) or low (4G) gellan gum
8 concentration, was performed, together with analysis of water content, fresh and dry mass,
9 endogenous ABA (GC-MS), soluble sugars (HPLC), starch, and confocal laser microscope
10 observations. This multi-scale, integrated analysis was used to unravel early molecular and
11 physiological events involved in somatic embryo development. Under unfavourable conditions (4G),
12 the glycolytic pathway was enhanced, possibly in relation to cell proliferation which may be
13 antagonistic to somatic embryo development. Under favourable conditions (9G), somatic embryos
14 adapted to culture constraint by activating specific protective pathways, and ABA-mediated molecular
15 and physiological responses promoting embryo development. Our results suggest that on 9G, germin-
16 like protein and ubiquitin-protein ligase could be used as predictive markers of somatic embryo
17 development whereas protein phosphatase 2C could be a biomarker for culture adaptive responses.
18 This is the first characterization of early molecular mechanisms involved in development of pine
19 somatic embryos following an increase in gellan gum concentration in the maturation medium, and it
20 is also the first report on somatic embryogenesis in conifers combining transcriptomic and proteomic
21 datasets.

22

23

24

25 Abbreviations

26 ABA, abscisic acid ; EMs, embryonal masses; SE, somatic embryo.

27

28

1 Introduction

Maritime pine, *Pinus pinaster* Ait., is industrially important in South-Western 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 Lelu-Walter et al. 2013). However, somatic embryo (SE) maturation a challenging step (Stasolla et al. 2002) towards 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 (Lara-Chavez et al. 2012; Pérez Rodríguez et al. 2006). 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 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, embryonal masses 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

1 proteins (Gutmann et al. 1996). ABA can regulate the switch from embryonal mass proliferation to
2 embryo development (Stasolla et al. 2002).

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

9 In addition to water availability and ABA, carbohydrate supply is of great importance for pine somatic
10 embryogenesis. In *Pinus pinaster*, maturation of SE is enhanced by raising the sucrose concentration
11 to 0.2M (Lelu-Walter et al. 2006). Clearly, endogenous carbohydrate status varies throughout somatic
12 embryogenesis (see review by Lipavská and Konrádová, 2004). In general, during the proliferation
13 phase embryonal masses have a high hexose content and almost no sucrose, resulting in a high
14 hexose/sucrose ratio. On the other hand, maturation results in a decrease in hexose content whereas a
15 gradual increase in sucrose content is observed in the SE (Lipavská et al. 2000). Starch, accumulates
16 mainly during the first weeks of the maturation process.

17 In addition to these physiological studies, efforts have recently been made to describe SE development
18 in conifers at the molecular level, but studies so far remain limited. Proteomic studies, combining 2D-
19 gel electrophoresis with a high degree of accuracy in protein identification by mass spectrometry, are
20 available only for *Cupressus sempervirens* (Sallandrouze et al. 1999), *Picea glauca* (Lippert et al.
21 2005), and *Larix x eurolepis* (Teyssier et al. 2013). These studies have contributed to the identification
22 of proteins involved in the embryogenesis process. For gene expression, only a few reports are
23 available, mainly for *Picea abies*. In this species, transcript levels have been compared between
24 different lines (embryogenic vs developmentally arrested) at different SE developmental stages
25 (Stasolla et al. 2004). More recently, changes in expression of genes with products, involved in auxin
26 biosynthesis and gibberellin-mediated signaling were reported, in a study that made the assumption
27 that conifer genes are homologous to those in Arabidopsis (Vestman et al. 2011). In support of this
28 assumption, genes related to embryogenesis in model plants have sequence similarity to expressed
29 sequence tags (ESTs) from *Pinus taeda* and *Larix leptolepis* (Cairney and Pullman, 2007; Zhang et al.
30 2012).

31 This study provides new insights into early molecular events regulating the development of *Pinus*
32 *pinaster* SE under reduced water availability. Our original approach consists in integrated multi-scale
33 analyses, including comparisons of the transcriptome and proteome with histological, biological and
34 biochemical observations. This is the first report combining transcriptomic and proteomic data for
35 conifers somatic embryogenesis. The molecular mechanisms underlying SE response to reduced water
36 availability are discussed in detail. We have found novel candidate predictive markers for conifer SE
37 development as well as for adaptive responses, which will facilitate practical application of the

1 knowledge gained in this study to monitor early responses of embryogenic tissue to maturation
2 conditions.

3 4 **Material and Methods**

5 ***Pinus pinaster* somatic embryogenesis**

6 Experiments were conducted with an embryogenic line (AAY06006) of *Pinus pinaster*. Embryonal
7 Masses (EMs) were initiated from immature zygotic embryos (Park et al. 2006) originating from a
8 Landes x Morocco polycross. EMs were subcultured and matured in darkness according to Lelu-
9 Walter et al. (2006). Briefly, filter paper discs with dissociated EMs were transferred onto maturation
10 medium which consisted of modified Litvay medium (Litvay et al. 1985) supplemented with 0.2 M
11 sucrose, 80 μ M cis-trans (\pm) abscisic acid (ABA, Sigma) and either 4 g or 9 g L⁻¹ gellan gum
12 (Phytigel™ Sigma), indicated here as 4G and 9G respectively. After 12 weeks of maturation,
13 cotyledonary SE were counted and embryogenic potential (number of SE per g FW) was estimated.

14 15 **Fresh Weight (FW), Dry Weight (DW), and water content of EMs**

16 Measurements were performed after 0, 1, 2, 3 and 4 weeks of maturation on both media. Each week,
17 each filter paper disc with dispersed EMs was weighed and replaced on the maturation medium. FW
18 of EMs was obtained by subtracting the mass of the filter paper disc from the total FW. DW was
19 determined after oven-drying at 70°C for 24 h and calculated by multiplying the DW to FW ratio by
20 100 (Teyssier et al. 2011). Water content was calculated as (FW-DW)/DW and expressed as g H₂O g⁻¹
21 DW (Dronne et al. 1997). Finally, water availability from the maturation media was determined by
22 measuring the amount of water absorbed by the filter paper disc (Teyssier et al. 2011). Each assay
23 (FW, DW, water content, water availability) was carried out on ten biological replicates for each
24 medium at each time point.

25 26 **Histology and microscopy**

27 EMs were collected after 0, 1, 2 and 3 weeks of maturation on both media. EM viability was
28 determined by double staining with fluorescein diacetate (FDA) (Sigma-Aldrich) and propidium
29 iodide (PI) (Sigma-Aldrich) according to modified protocols (Jones and Senft, 1985; Petřek et al.
30 2005; Vondráková et al. 2010). A LSM5 Duo confocal laser scanning microscope (Zeiss, Jena,
31 Germany) equipped with an Argon/2 laser (FDA excitation at 488 nm, emission filter-set BP 505–550)
32 and a DPSS laser (PI excitation at 561 nm, emission filter-set LP 650) was used to visualize stained
33 EMs. Viable cells exhibit bright green fluorescence, non-viable or dead cells show bright red
34 fluorescence. Starch grains in squashed fresh EMs were stained with iodine/potassium iodide and
35 observed with a Jenaval transmission light microscope (Zeiss, Jena, Germany) equipped with a DS-
36 5M digital camera (Nikon, Tokyo, Japan). All images were processed using the computer image
37 analysis system NIS Elements, version 3.2 (Laboratory Imaging, Prague, Czech Republic).

ABA and abscisic acid 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 Bielecki solution; a further 2 x 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, U.S.A.) to eliminate non-polar compounds. The eluate was partly evaporated by the rotary vacuum evaporator (RVE) to approx. 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, U.S.A.), 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, USA) filters. The filtrates were divided into two parts, one for determination of ABA, the second for determination of ABA-GE.

ABA determination by GC-MS. ABA was analyzed by GC-MS/MS (Vandenbusche 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 an LC-MS system consisting of an HTS PAL autosampler (CTC Analytics, Switzerland), Rheos 2200 gradient HPLC pump (Flux, Switzerland) and Quantum Ultra AM triple-quadrupole mass spectrometer (Thermo Instruments, USA) equipped with an ESI interface. Separation was performed on an HPLC column, 50 x 2.1 mm Kinetex C18 (Phenomenex, USA) using gradient elution with a mobile phase consisting of acetonitrile, water and 1% acetic acid. The mass spectrometer was set to record SRM transitions 425→153 for ABA-GE and, as internal standard, 430→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 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, New York, USA), and then analyzed by HPLC. Samples were eluted with

1 80:20 (v/v) acetonitrile/H₂O at a flow rate of 1 mL min⁻¹ on a 200 x 4.6 mm Spherisorb-NH₂ column
2 (Thermo Fisher Scientific, France). The eluates were analyzed with a differential refractometer (RI
3 Plus Detector, Thermo Fisher Scientific, France) and the peak areas were electronically integrated
4 using ChromQuest 5.0 (Thermo Fisher Scientific, France). The various sugars were identified by co-
5 elution with standards (Sigma, St Louis, MO, USA). For starch, undissolved sample pellets resulting
6 from soluble carbohydrate extraction were dried and incubated in 0.02N NaOH for 2 h at 95°C. After
7 cooling to room temperature, starch was reduced to glucose: amyloglucosidase (35 U) was added and
8 samples were incubated at 50°C for 1 h with stirring. After centrifugation, aliquots of the clear
9 supernatants were assayed for glucose (Bergmeyer et al. 1974).

10 *Total protein assay.* Total protein extracts were prepared from five replicates for each type of sample
11 (25 to 50 mg FW of frozen EMs) as previously described (Teyssier et al. 2013). Protein concentrations
12 were determined using the Bradford assay with bovine serum albumin as a standard.

13 Results, which were expressed as µg g⁻¹ DW, corresponded to the mean of 3 (carbohydrates), 5 (total
14 protein) or 6 (starch) replicates ± confidence interval (CI) at the 5% error level.

16 **Transcriptomic analysis**

17 Maturation was carried out as described above. EMs were collected after 1 week of maturation on 4G
18 and 9G, frozen in liquid nitrogen, and ground for 5 min with a mortar and pestle to a fine powder.
19 Frozen powder samples were divided into two equal portions (each about 400 mg FW) for
20 transcriptomic and proteomic analyses. For transcriptomic analysis, three biological replicates were
21 sampled for each maturation conditions, producing a total of 6 samples.

22 *Sequencing.* Library construction and sequencing was done by a commercial service provider (GATC
23 Biotech, Germany). Total RNA was extracted, mRNAs purified and cDNA synthesized according to
24 the Smart Kit (Clontech, USA) manufacturer's recommendations. The six samples were paired-end
25 sequenced on an Illumina Genome Analyser II producing 100 bp paired-end short reads. Primer-
26 adapter removal, quality filtering and ambiguous base trimming produced sequences 82 ± 6 bp long.
27 The six samples each produced 5,586,402 ± 1,514,289 paired-end sequence pairs, giving a total of
28 33,518,411 paired-end sequence pairs. Original data were deposited in the NCBI' Sequence Read
29 Archive under accession number SRR609713.

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

35 *Functional Characterization and Gene Ontology (GO) Annotation.* Gene Ontology (GO) annotations
36 to retrieve molecular function, biological process and cellular component terms were performed using
37 Blast2GO (<http://www.blast2go.org/>) (Conesa et al. 2005). The sequences were loaded into the

1 Blast2GO program, and BLASTx with a minimum E-value of 1.0 was performed by the program prior
2 to mapping. The mapping step was used to annotate the sequences with GO terms from the GO
3 database. Fisher's exact test was performed within Blast2GO to determine whether GO identifiers
4 occurred more often in a group than would appear by chance. For GO analysis, biological process GO
5 identifiers were considered only if they occurred more than 5 times.

6 *Functional Classification by KEGG.* Annotation against the Kyoto Encyclopedia of Genes and
7 Genomes (KEGG) database (<http://www.genome.jp/kegg/>) was performed in order to exploit enzyme
8 data to integrate genomic, chemical and network information (Ogata et al. 1999) using the Blast2GO
9 program. All mapped sequences were annotated to KEGG database accessions to obtain their enzyme
10 commission (EC) numbers. The EC numbers were then mapped onto KEGG to obtain the
11 corresponding KEGG Pathway-Maps.

12 **2-D gel proteomic analysis**

13 *2-D PAGE.* Total protein extracts were prepared from 5 replicates for each type of sample
14 (approximately 400 mg FW) as previously described in Teyssier et al. (2013). Briefly, extraction was
15 realised with phenol and precipitation with ammonium acetate in ethanol then the proteins assayed.
16 First dimension separation was performed with 300 µg protein loaded onto 24-cm IPG strips, pH 4-7
17 (Protean IEF Cell system, BioRad, France), whereas the 2-D PAGE was performed with 11%
18 polyacrylamide gels. Gels were stained with colloidal CBB-G. The numerical analyse of the gels was
19 performed with Progenesis software (Nonlinear Dynamics, United Kingdom) as described in Teyssier
20 et al. (2013).

21 *Gel spot processing.* The destain, wash and proteolysis steps and sample preparation for mass
22 spectrometry are described elsewhere (Teyssier et al. 2013).

23 *NanoLC-MS/MS analysis.* Liquid chromatographic separations of peptides were performed using an
24 UltiMate 3000 nano-ultraLC system (Dionex, USA) after preconcentration on a 15 cm reverse-phase
25 column (Acclaim Pepmap RSLC C18). Online analysis via a distal-coated PicoTip emitter (New
26 Objective, USA) and a nanospray source was performed on an HCT Ultra ion trap (Bruker, Germany),
27 using HyStar 3.2 software (Bruker) for integration of liquid chromatography and mass spectrometry.
28 The spray voltage was set between 1800 and 2200 V through EsquireControl 6.2 (Bruker) and the ion
29 trap was scanned in MS mode from m/z 100 to 1500. Peaks with the four highest intensities were
30 selected for CID MS/MS fragmentation within a 4 Da isolation window, using an exclusion m/z list
31 and dynamic exclusion as previously described (Beaufour et al. 2012). Fragment ions were scanned
32 from m/z 100 to 3000.

33 *Protein identification by database sequence assignment.* The procedure was as described elsewhere
34 (Teyssier et al. 2013), except for the database used. MS/MS spectra were searched against the *Pinus*
35 *pinaster* contig database (see following paragraph). Only proteins with a minimum of four peptides
36 reaching individual scores greater than 60% were taken into account. Protein lists for each spot were
37

1 curated manually to remove redundancies. The false discovery rate (FDR) at the peptide level was
2 calculated by combining all data and applying the decoy feature of the identification engine. A FDR
3 value of 0.57% was found for peptide scores equal or greater than 60%.

5 **Statistical analysis**

6 Statistical analysis was carried out with R software (version 2.8.0; R Development Core Team, 2011).
7 Effects of the treatments on the FW, DW, water content, sugar analysis, total protein analysis and
8 ABA measurements were evaluated using one-way ANOVA. Variations in these parameters during
9 maturation in relation to the gellan gum concentration in the medium were analysed with multiple
10 comparisons of means with Tukey contrasts ($P < 0.05$). For analysis of transcriptomic data, each
11 paired-end sequence was mapped onto the *Pinus pinaster* transcriptome assembly ([http://genotoul-
12 contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster3/index.html](http://genotoul-contigbrowser.toulouse.inra.fr:9092/Pinus_pinaster3/index.html)) with BWA (Li and Durbin, 2009).
13 Paired-end matches per transcript were counted, and corrected for transcript length and sample
14 sequencing depth to produce transcript expression levels in FPKM (Roberts et al. 2011). Differential
15 expression was assessed by a *t*-test using biological replicates of each treatment. *P*-values were
16 adjusted for multiple hypothesis testing by the Benjamini-Hochberg correction (Benjamini and
17 Hochberg, 1995) to control for the false discovery rate. Transcripts with an adjusted *P*-value below
18 0.05 were considered to be differentially expressed. Finally, for 2-D PAGE analysis, the intensity
19 change for each spot was analysed with Student's *t*-test on the basis of the normalized spot volume ($P <$
20 0.05).

22 **Results**

23 **Histological and biological characteristics of maritime pine EMs**

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

1 Compared with EMs matured on 4G, those transferred onto 9G showed no FW increase (Fig. 3A), but
2 there was a significant increase in their DW ($P < 0.05$, Fig. 3B) and decrease in their water content ($P <$
3 0.05 , Fig. 3C). After 4 weeks, EMs matured on 9G had a significantly higher DW content (10.1% FW)
4 and lower water content ($9.0 \text{ g H}_2\text{O g}^{-1} \text{ DW}$) than those on 4G (5.1% FW, and $19.0 \text{ g H}_2\text{O g}^{-1} \text{ DW}$;
5 Fig. 3B, 3C respectively). Availability of water from the maturation medium was significantly lower
6 on 9G ($0.580 \text{ g H}_2\text{O} \pm 0.009$, mean \pm CI at the 5% error level.) than on 4G ($0.635 \text{ g H}_2\text{O} \pm 0.008$).

8 **ABA content**

9 At the beginning of the maturation period, the ABA content of EMs was low ($2.1 \text{ nmol g}^{-1} \text{ DW}$). EMs
10 subsequently showed a significant increase in endogenous ABA (Fig. 4) whereas ABA glucose ester
11 conjugate (ABA-GE) remained undetectable whatever the conditions tested (data not shown). On 9G,
12 ABA content increased slowly during the 3 first weeks of maturation but peaked at $21.1 \text{ nmol g}^{-1} \text{ DW}$
13 after 4 weeks. A similar initial increase in ABA content was observed in EMs matured on 4G, but after
14 4 weeks of maturation, the ABA content was less than half ($9.0 \text{ nmol g}^{-1} \text{ DW}$) that on 9G.

16 **Carbohydrate and total protein contents**

17 At 1 week of maturation, similar amounts of total protein were observed for EMs developed on 4G or
18 9G (Table 1). In contrast, significant differences in carbohydrates contents were observed. EMs
19 matured on 4G had a greater amount of fructose (Fru, $27.4 \text{ } \mu\text{g g}^{-1} \text{ DW}$) and glucose (Glu, $26.6 \text{ } \mu\text{g g}^{-1}$
20 DW) but a similar sucrose content (Su, $6.88 \text{ } \mu\text{g g}^{-1} \text{ DW}$) to that of EMs matured on 9G. In
21 consequence, the ratio (Fru+Glu)/Su was significantly higher for EMs matured on 4G than for those
22 matured on 9G (7.9 and 4.4, respectively). Other carbohydrates such as stachyose, raffinose, maltose
23 and mannose were not detected whatever the conditions. Finally, EMs matured on 9G contained a
24 significantly larger amount of starch ($46.30 \text{ } \mu\text{g g}^{-1} \text{ DW}$) than those matured on 4G ($32.70 \text{ } \mu\text{g g}^{-1} \text{ DW}$).

26 **Transcriptomic analysis**

27 1103 transcripts were considered to be differentially expressed, 632 for EMs matured on 9G and 471
28 for EMs matured on 4G (Appendix S1 in Supporting Information).

29 *Data annotation by enrichment analysis and functional classification.* Using the Blast2GO program,
30 significant transcripts were annotated against the Plants/*Arabidopsis thaliana* section of the protein
31 sequences database. Next, the annotated significant transcripts were mapped against the GO database
32 to retrieve the relevant GO terms. In total 4865 GO terms were assigned to the 529 mapped transcripts
33 with an average of five GO terms assigned to each transcript. Of these GO terms, the majority were
34 assigned to the Biological Process category (3118, 64%), followed by Molecular Function (1217,
35 25%), and Cellular Component (530, 11%).

36 GO analysis at level 2, illustrating the Biological Process categories, is presented in the “Hit
37 description” column of Appendix S2 in Supporting Information. Within the Biological Process

1 ontology, proteins involved in cellular process, metabolic process and response to stimulus
2 predominated.

3 Figure 5 shows the results of GO enrichment analysis, illustrating the distribution of Biological
4 Process GO terms that were differentially enriched for EMs matured on 4G versus 9G. Twenty-one
5 functional categories were represented by 298 sequences originating from EMs matured on 4G and
6 239 sequences from 9G. The majority of the significant categories for EMs matured on 4G were
7 involved in carbohydrate metabolism, energy metabolism and cellular respiration, while for EMs
8 matured on 9G the significant categories were cell growth and embryo-related development, regulation
9 response to stimulus and secondary metabolic processes.

10 In order to identify the biological pathways active in early EM maturation, the assembled ESTs were
11 used to obtain the corresponding Enzyme Commission numbers (ECs) by annotation against the
12 KEGG database. A total of 207 transcripts were assigned to 40 ECs. The ECs were then grouped into
13 56 biochemical pathways, of which 33 pathways were found to be involved in metabolism. Most
14 enzymes active in EMs matured on 4G participate in the glycolysis/gluconeogenesis pathway (23) and
15 methane metabolism pathway (16).

16 17 *Transcript expression levels.*

18 Among the 471 transcripts overexpressed in EMs grown on 4G, the most prominent were 9 encoding
19 "pyruvate decarboxylase" with a mean of 12.0 FPKM, 7 encoding "alcohol dehydrogenase" (8.3
20 FPKM), 7 annotated as "bifunctional enolase 2" and related to a glucose catabolic process (15.2
21 FPKM), 1 annotated as "6-phosphofructokinase 2" involved in glycolysis (7.5 FPKM) and 1 annotated
22 as "sucrose synthase 3" responsive to oxygen levels (9.8 FPKM).

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

34 35 **2-D gel proteomic analysis**

36 Changes in EM protein profiles were investigated using 2D gel electrophoresis. Separation with high
37 resolution was achieved, allowing detection of a total of 1,428 spots (representative gel shown in

Appendix S3 in the Supporting Information). 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 mass spectrometry. 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 Appendix S4 in the Supporting Information. All identified proteins were grouped into functional categories within the Biological Process ontology (level 2) using Blast2GO (Appendix S5 in Supporting Information). 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.

Since the *Pinus 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 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 somatic embryos?

When compared to the lower concentration, 4G, it appears that on 9G, maritime pine SE differentiated better, based on macroscopical and histological observations (Fig. 1, 2), and on measurements of DW and water content (Fig. 3). Our results reinforce findings previously reported for somatic

1 embryogenesis in white pine (Klimaszewska et al. 2000), and Scots pine (Lelu-Walter et al. 2008),
2 where high gellan gum concentrations were linked to lower water availability and enhanced SE
3 development. Low water and carbon source availability also had a major impact on SE maturation and
4 germination in carrot (Lee et al. 2001), suggesting that the induction of embryo development by
5 temporary growth starvation observed in natural conditions can be mimicked *in vitro* in a broad range
6 of species. Our results indicate that gellan gum concentration and associated modulation of water
7 availability in the medium is critical in controlling SE development in maritime pine.

9 **What happens in EMs when somatic embryo maturation is hindered?**

10 Microscopic observations (Fig. 1) showed that EMs grown on 4G produced less structured immature
11 embryos and more undifferentiated tissues through cell proliferation (Fig. 2), an observation
12 corroborated by the increase in FW due mainly to water uptake throughout the culture period (Fig. 3).

13 In general, for pine species maturation in the presence of low concentrations of gelling agent sustains
14 proliferation of immature SE but is detrimental to SE development (Klimaszewska et al. 2000).

15 In 4G-grown EMs, endogenous synthesis of ABA is limited and there is no detectable ABA
16 conjugation, at least via ABAGE synthesis. We avoided exogenous application of ABA at levels high
17 enough to obscure endogenous ABA synthesis by several orders of magnitude because of the
18 occurrence of high levels of ABA conjugation (Daie et al. 1984) as reported for larch (Label and Lelu,
19 2000).

20 Carbohydrate contents in EMs cultured on 4G is characterized by a high hexose:sucrose ratio, and
21 fairly low starch accumulation which is in accordance with microscopic observations (Fig. 1e). These
22 results suggest stimulation of the glycolytic pathway, possibly related to cell proliferation (Fig.2a). A
23 high hexose:sucrose ratio has been observed in proliferating conifer EMs (Lipavská and Konrádová,
24 2004).

25 Transcriptionally, our results show an increased level of expression of alcohol dehydrogenase (ADH)
26 and pyruvate decarboxylase (PDC) genes in 4G-grown EMs, along with over-expression of sucrose
27 synthase. At the proteomic level, ADH, phosphoglycerate kinase and PDC isozyme 2 are also
28 significantly up-regulated. The two datasets are consistent, and support possible stimulation of the
29 glycolytic pathway, as also suggested by the results of carbohydrate analysis. However a recent paper
30 on *Arabidopsis* reported that ADH and PDC enzymes and their corresponding transcripts could be
31 modulated to overcome unfavorable conditions such as anoxia brought about by submergence in
32 culture medium (Mithran et al. 2013). In addition, utilization of sucrose, provided in our study by the
33 culture medium, is important in anoxia tolerance mediated by alcoholic fermentation (Loreti et al.
34 2005). These findings, taken together with our results, suggest that for EMs, 4G is *i*) a condition
35 unfavorable for SE differentiation and maturation, possibly due to an excess of water at this low gellan
36 gum concentration and *ii*) 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 over-expression of several bifunctional enolase-2 transcripts which 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 which 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, since 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 somatic embryos?

ABA is synthesized at a low level during the three first weeks of culture (Fig. 4), but peaks at week 4 when SE 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 ubiquitinated proteins are recognized and catabolized by the 26S proteasome

1 complex. This pathway supplies amino acids for the biosynthesis of new proteins (Vierstra, 2003). A
2 recent paper (Rode et al. 2012) links high levels of enzymes involved in the proteasome-dependent
3 proteolysis pathway and the switch from callus to globular embryos in cyclamen.
4 During somatic embryogenesis in gymnosperms, chitinases are reportedly excreted into the culture
5 medium. These enzymes are involved in the regulation of the EM-to-SE transition and massive
6 programmed cell death (Wiweger et al. 2003; Vestman et al. 2011). Our results showed over-
7 expression of transcripts for chitinase in EMs grown on 9G. In Norway spruce, programmed cell death
8 is responsible for the degradation of EMs when early SEs differentiate (Vestman et al. 2011). The
9 timing of programmed cell death can also be modulated by disulfide isomerase, which was up-
10 regulated on 9G. This enzyme seems to be required for proper embryo development in *Arabidopsis*
11 seeds (Andème-Ondzighi et al. 2008). Expansin S2 and putative expansin-B14 are known to play roles
12 in embryogenesis. In *Arabidopsis*, they are involved in the cell cycle, as key regulators of wall and cell
13 expansion (Thibaud-Nissen et al. 2003; Spencer et al. 2007). The proteins tubulin beta-2 chain and
14 GTP-binding nuclear protein Ran-A1 participate in cytoskeletal process such as nucleocytoplasmic
15 transport (Pan et al. 2009; Vernoud et al. 2003). All these proteins, which are up-regulated in 9G, may
16 direct cell division towards cell remodeling for embryogenesis, a hypothesis consistent with our
17 histological observations.
18 Several serine/threonine kinase (SRK2) transcripts are highly over-expressed in EMs growing on.
19 SRK2s are plant homologs of AMPK/SNF1 genes, which in other eukaryotes are considered to be
20 crucial in linking stress and metabolic responses (Kulik et al. 2011). They are reported to be key
21 players in osmotic stress responses and ABA-dependent plant development (Fujii et al. 2011).
22 Endogenous ABA synthesis was significantly enhanced on 9G, i.e when water availability was
23 reduced. In addition, the detection of highly over-expressed protein phosphatase 2C (PP2C) transcripts
24 in 9G grown EMs reinforces the likelihood of such a - possibly adaptive - molecular response (Sheen,
25 1998). PP2Cs are reported to be key protein-protein interactors mediating ABA responses in plants
26 (Yoshida et al. 2006). Our results suggest that immature SE respond to reduced availability of water in
27 the maturation medium through ABA-mediated molecular mechanism.
28 Two superoxide dismutases are overexpressed at the proteomic level, as is one germin-like protein.
29 This latter is also detected at the transcriptomic level, suggesting an ongoing process of antioxidant
30 protein production through mRNA synthesis EMs growing on 9G. The role of germin-like proteins in
31 plant development and defense is well documented (Bernier and Berna, 2001). Their involvement as
32 markers of early stages of somatic embryogenesis in both conifers (Neutelings et al. 1998; Mathieu et
33 al. 2006) and angiosperms (Caliskan et al. 2004; Kim and Triplett, 2004) makes them especially
34 valuable in this context. Superoxide dismutase is involved in the regulation of oxidative stress. It
35 destroys free radicals which are normally produced within the cells and are toxic to biological systems.
36 In a proteomic study on the early stages of seed development in *Pinus massoniana*, a correlation was
37 established between the level of this protein and the transition from the pro-embryo to the columnar

1 embryo stage (Zhen et al. 2012). It has been suggested that the ability of embryos to develop an
2 efficient redox homeostasis system may be key for successful regulation and maintenance of cell
3 differentiation.

4 Our results also show that many transcripts related to F3H, as well as several other transcripts related
5 to flavonol metabolism, are over-expressed in 9G-grown EMs. F3H catalyses hydroxylation of
6 flavanones to dihydroflavonols. Flavonoids are ubiquitous in plants, and some of these compounds
7 play specific roles in seed development (Lepiniec et al. 2006). Among these roles, stress resistance is a
8 key biological function which leads to the synthesis of several secondary compounds all related to
9 flavanones hydroxylation and subsequent metabolism down to condensed tannins (Winkel-Shirley,
10 2002). Our results suggest that EMs cultivated on 9G may overcome constraints imposed by culture
11 conditions by synthesizing protective secondary metabolites. This does not seem to occur in EMs
12 grown on 4G, where many stress response-related transcripts were identified as discussed above.

13 At the proteomic level, two enzymes (glucose-1-phosphate adenylyltransferase AGPP; 4-alpha-
14 glucanotransferase DPE2, disproportionating enzyme) reportedly involved in starch biosynthesis
15 (Keeling and Myers, 2010) were found to be overexpressed on 9G. This is consistent with the high
16 starch content of immature SE cultivated in these favorable conditions (Fig. 1, Table 1). It is
17 interesting to note that glucose and fructose contents were correlated with the difference in starch grain
18 accumulation on 4G and 9G (Fig. 1e, f; Table 1). In maritime pine, a low abundance of starch granules
19 was reported in proliferating EMs (Breton et al. 2005) whereas starch was found to accumulate
20 throughout developing embryos during early maturation (Tereso et al. 2007). A similar picture was
21 reported during larch SE maturation (Gutmann et al. 1998).

22 **Conclusions**

23 This is the first study on the early molecular mechanisms involved in somatic embryogenesis of pine
24 following an increase in gellan gum concentration in the medium, and it is also the first report
25 combining transcriptomic and proteomic data on somatic embryogenesis in conifers. Our
26 transcriptomic and proteomic results indicate that glycolysis is enhanced under conditions
27 unfavourable for SE maturation (4G) which lead to EM proliferation. Under favourable conditions
28 (9G), we observed adaptive, ABA-mediated molecular and physiological responses to reduced water
29 availability resulting in early transition of EMs from proliferation to the SE developmental pathway
30 (indicated by active protein synthesis, and overexpression of proteins involved in cell division,
31 embryogenesis and starch synthesis). Specific pathways (synthesis of protective secondary
32 metabolites, regulation of oxidative stress) are concomitantly activated, apparently to overcome
33 constraints due to culture conditions.

34 Based on both our transcriptomic and proteomic data and related information available in the
35 literature, we proposed the germin-like protein and ubiquitin-protein ligase as potential marker genes
36 and proteins for early evidence of effective SE development on maturation medium. Protein-

1 phosphatase 2C (PP2C) may be also a suitable marker for monitoring adaptive molecular responses to
2 favourable maturing conditions that may have a beneficial impact on regeneration capacity of cultures.
3 The identification of relevant biomarkers at the molecular level needs validation experiments but has
4 clear practical implications for implementation of somatic embryogenesis in both conifer fundamental
5 research and commercial applications. Functional genomics based on reverse genetics in maritime pine
6 is a long process requiring transgenic plant regeneration through somatic embryogenesis and
7 cryopreservation of selected transgenic lines (Trontin et al. 2002). Labor intensiveness would be
8 significantly decreased if the regeneration capacity of transgenic lines could be predicted from early
9 molecular monitoring of rapid maturation tests (1 week). Furthermore untargeted (genome-wide) or
10 targeted transcriptomic and proteomic profiling (candidate genes and proteins markers) may be
11 invaluable for qualitative and quantitative selection of elite embryogenic lines that are receptive to
12 maturing conditions and productive enough for breeding application (clonal and varietal tests) and
13 new variety deployment in the frame of multivarietal forestry.

14 15 16 **Author contributions**

17 AM participated in the acquisition of the data, carried out all types of analyses and drafted the
18 manuscript. CT conceived in the design of the study, participated in 2-D gel proteomic analysis and
19 drafted the manuscript. JFT conceived in the design of the study, participated in somatic
20 embryogenesis and drafted the manuscript. KE carried out histological and microscopical analysis and
21 drafted the manuscript. BP carried out ABA and ABA GE analyses. MB carried out Gel spot
22 processing and protein identification by database sequence assignment. DM participated in the design
23 of the study and helped to draft the manuscript. NB carried out carbohydrate analyses and helped to
24 draft the manuscript. CLM carried out somatic embryogenesis and collected the material. LB-B
25 participated in 2-D gel proteomic analyses. IR carried out somatic embryogenesis and collected the
26 material. LH participated in the design of the study and helped to draft the manuscript. MC
27 participated in gel spot processing and protein identification by database sequence assignment and
28 helped to draft the manuscript. FC participated in study design, participated in carbohydrate analysis
29 and drafted the manuscript. MV participated in the design of the study and drafted the manuscript. PL
30 conceived in the design of the study, carried out transcriptomic analysis, performed the statistical
31 analysis and drafted the manuscript. MALW conceived in the design of the study and its coordination,
32 and drafted the manuscript. All authors read and approved the final manuscript.

33 34 35 **Acknowledgments**

36 This work was supported by grants from the French "Conseil Régional de la Région Centre"
37 (EMBRYOME project, contract 33639), the "Conseil Régional de la Région Aquitaine" (Embryo2011,

Version définitive du manuscrit publié dans / Final version of the manuscript published in :
Physiologia Plantarum, 2014, 152(1), 184-201 <http://dx.doi.org/10.1111/ppl.12158>

1 contract 09012579-045), and the French Ministry of Foreign Affairs and the French Ministry of
2 Higher Education and Research through the France/Czech Republic Science Cooperation
3 BARRANDE Programme. Data analysis was made possible through the involvement of INRA's
4 GenoToul bioinformatics platform in Toulouse (France).
5
6

For Peer Review

1 **Bibliography**

- 2 Andème-Ondzighi C, Christopher DA, Cho EJ, Chang S-C, Staehelin LA (2008) Arabidopsis protein
3 disulfide isomerase-5 inhibits cysteine proteases during trafficking to vacuoles before programmed
4 cell death of the endothelium in developing seeds. *Plant Cell* 20: 2205-2220
- 5 Attree SM, Fowke LC (1993) Embryogeny of gymnosperms: advances in synthetic seed technology of
6 conifers. *Plant Cell Tiss Org* 35: 1-35
- 7 Beaufour M, Godin F, Vallée B, Cadene M, Bénédicti H (2012) Interaction proteomics suggests a new
8 role for the Tfs1 protein in yeast. *J Proteome Res* 11: 3211-3218
- 9 Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: a practical and powerful
10 approach to multiple testing. *J Royal Stat Society, Series B* 57: 289-300
- 11 Bernier F, Berna A (2001) Germins and germin-like proteins: Plant do-all proteins. But what do they
12 do exactly? *Plant Physiol Bioch* 39: 545-554
- 13 Bergmeyer HU, Bernt E, Schmidt F, Stork H (1974) D-glucose. In: Bergmeyer HU (ed) *Methods of*
14 *enzymatic analysis*, vol III. Verlag Chemie, Weinheim, pp 1196-1201
- 15 Black M (1991) Involvement of ABA in the physiology of developing and mature seeds. In: Davies
16 WJ, Jones HG (eds) *Abscisic acid: physiology and biochemistry*. Oxford: BIOS Scientific Publisher
17 pp 99-124
- 18 Black M, Corbineau F, Grzesik M, Guy P, Côme D (1996) Carbohydrate metabolism in the
19 developing and maturing wheat embryo in relation to its desiccation tolerance. *J Exp Bot* 47: 161-
20 169
- 21 Breton D, Harvengt L, Trontin J-F, Bouvet A, Favre J-M (2005) High subculture frequency, maltose-
22 based and hormone-free medium sustained early development of somatic embryos in maritime pine.
23 *In Vitro Cell Dev Biol Plant* 41: 494-504
- 24 Caliskan M, Turet M, Cuming A (2004) Formation of wheat (*Triticum aestivum* L.) embryogenic
25 callus involves peroxide-generating germin-like oxalate oxidase. *Planta* 219: 132-140
- 26 Cairney J, Pullman GS (2007) The cellular and molecular biology of conifer embryogenesis. *New*
27 *Phytologist* 176: 511-536
- 28 Chinnusamy V, Zhu J, Zhu J-K (2007) Cold stress regulation of gene expression in plants. *Trends*
29 *Plant Sci* 12: 444-451
- 30 Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool
31 for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21: 3674-
32 3676
- 33 Cubenas-Potts C, Matunis MJ (2013) SUMO: A multifaceted modifier of chromatin structure and
34 function. *Developmental Cell* 24: 1-12
- 35 Daie J, Wyse R, Hein M, Brenner ML (1984) Abscisic acid metabolism by source and sink tissues of
36 sugar beet. *Plant Physiol* 74: 810-814

- 1 Dembinsky D, Woll K, Saleem M, Liu Y, Fu Y, Borsuk LA, Lamkemeyer T, Fladerer C, Madlung J,
2 Barbazuk B, Nordheim A, Nettleton D, Schnable PS, Hochholdinger F (2007) Transcriptomic and
3 proteomic analyses of pericycle cells of the maize primary root. *Plant Physiol* 145: 575-588
4 Dogra PD (1978) Morphology, development and nomenclature of conifer embryo. *Phytomorphology*
5 28: 307-322
6 Dronne S, Label P, Lelu M-A (1997) Desiccation decreases abscisic acid content in hybrid larch
7 (*Larix × leptoeuropaea*) somatic embryos. *Physiol Plant* 99: 433-438
8 Fujii H, Verslues PE, Zhu J-K (2011) Arabidopsis decuple mutant reveals the importance of SnRK2
9 kinases in osmotic stress responses *in vivo*. *Proceedings National Acad Sci* 108: 1717-1722
10 Gutmann M, Von Aderkas P, Label P, Lelu M-A (1996) Effects of abscisic acid on somatic embryo
11 maturation of hybrid larch. *J Exp Bot* 47: 1905-1917
12 Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, Scarascia Mugnozza G (2012)
13 Accelerating the domestication of forest trees in a changing world. *Trends Plant Sci* 17: 64-72
14 Jennissen H (1995) Ubiquitin and the Enigma of Intracellular Protein Degradation. *European J Bioch*
15 231: 1-30
16 Jones KH, Senft JA (1985) An improved method to determine cell viability by simultaneous staining
17 with fluorescein diacetate-propidium iodide. *J Histochem Cytochem* 33: 77-79
18 Kapik RH, Dinus RJ, Dean, JFD (1995) Abscisic acid and zygotic embryogenesis in *Pinus taeda*. *Tree*
19 *Physiol* 15: 485-490
20 Keeling PL, Myers AM (2010) Biochemistry and genetics of starch synthesis. *Annual Review Food*
21 *Sci Technol* 1: 271-303
22 Kim H, Triplett B (2004) Cotton fiber germin-like protein. I. Molecular cloning and gene expression.
23 *Planta* 218: 516-524
24 King RW (1982) Abscisic acid in seed development. In: Khan AA (ed) *Physiology and biochemistry*
25 *of seed development, dormancy and germination*, Elsevier Biomedical press, Amsterdam, pp 157-
26 181
27 Klimaszewska K, Smith D (1997) Maturation of somatic embryos of *Pinus strobus* is promoted by a
28 high concentration of gellan gum. *Physiol Plant* 100: 949-957
29 Klimaszewska K, Bernier-Cardou M, Cyr DR, Sutton BCS (2000) Influence of gelling agents on
30 culture medium gel strength, water availability, tissue water potential, and maturation response in
31 embryogenic cultures of *Pinus strobus* L. *In Vitro Cell Dev Biol Plant* 36: 279-286
32 Kong L, Attree SM, Fowke LC (1997) Changes of endogenous hormone levels in developing seeds,
33 zygotic embryos and megagametophytes in *Picea glauca*. *Physiol Plant*. 101: 23-30
34 Kulik A, Wawer I, Krzywinska E, Bucholc M, Dobrowolska G (2011) SnRK2 protein kinases—key
35 regulators of plant response to abiotic stresses. *Omic: a journal of integrative biology* 15: 859-872
36 Label P, Lelu MA (2000) Exogenous abscisic acid fate during maturation of hybrid larch (*Larix x*
37 *leptoeuropaea*) somatic embryos. *Physiol Plant* 109: 456-462

- 1 Lara-Chavez A, Egertsdotter U, Flinn BS (2012) Comparison of gene expression markers during
2 zygotic and somatic embryogenesis in pine. *In Vitro Cell Dev Biol Plant* 48: 341-354
- 3 Lee EK, Cho DY, Soh WY (2001) Enhanced production and germination of somatic embryos by
4 temporary starvation in tissue cultures of *Daucus carota*. *Plant Cell Rep* 20: 408-415
- 5 Lepiniec L, Debeaujon I, Routaboul J-M, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics
6 and biochemistry of seed flavonoids. *Annual Review Plant Biol* 57: 405-430
- 7 Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2006) Simplified and improved somatic
8 embryogenesis for clonal propagation of *Pinus pinaster* Ait. *Plant Cell Rep* 25: 767-776
- 9 Lelu-Walter M-A, Bernier-Cardou M, Klimaszewska K (2008) Clonal plant production from self- and
10 cross-pollinated seed families of *Pinus sylvestris* (L.) through somatic embryogenesis. *Plant Cell*
11 *Tiss Org Cult* 92: 31-45
- 12 Lelu-Walter M-A, Thompson D, Harvengt L, Sanchez L, Toribio M, Pâques LE (2013) Somatic
13 embryogenesis in forestry with a focus on Europe: state-of-the-art, benefits, challenges and future
14 direction. *Tree Genet Gen* 9: 883-899
- 15 Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform.
16 *Bioinformatics* 25: 1754-1760
- 17 Lipavská H, Konrádová (2004) Somatic embryogenesis in conifers: the role of carbohydrate
18 metabolism. *In Vitro Cellular and Developmental Biology - Plant* 40: 23-30
- 19 Lipavská H, Svobodová H, Albrechtová J, Kumstýřová L, Vágner M, Vondráková Z (2000) Somatic
20 embryogenesis in Norway spruce: carbohydrate status during embryo maturation and the effect of
21 polyethylene glycol treatment. *In Vitro Cell Dev Biol Plant* 36: 260-267
- 22 Lippert D, Jun Z, Ralph S, Ellis DE, Gilbert M, Olafson R, Ritland K, Ellis B, Douglas CJ, Bohlmann
23 J (2005) Proteome analysis of early somatic embryogenesis in *Picea glauca*. *PROTEOMICS* 5: 461-
24 473
- 25 Litvay JD, Verma DC, Johnson MA (1985) Influence of loblolly pine (*Pinus taeda* L.) culture medium
26 and its components on growth and somatic embryogenesis of the wild carrot (*Daucus carota* L.).
27 *Plant Cell Rep* 4: 325-328
- 28 Loreti E, Poggi A, Novi G, Alpi A, Perata P (2005) A genome-wide analysis of the effects of sucrose
29 on gene expression in *Arabidopsis* seedlings under anoxia. *Plant Physiol* 137: 1130-1138
- 30 Mathieu M, Lelu-Walter M-A, Blervacq AS, David H, Hawkins S, Neutelings G (2006) Germin-like
31 genes are expressed during somatic embryogenesis and early development of conifers. *Plant Mol*
32 *Biol* 61: 615-627
- 33 Mithran M, Paparelli E, Novi G, Perata P, Loreti E (2013) Analysis of the role of the pyruvate
34 decarboxylase gene family in *Arabidopsis thaliana* under low-oxygen conditions. *Plant Biology*, (in
35 press) doi: 10.1111/plb.12005

- 1 Neutelings G, Domon JM, Membre N, Bernier F, Meyer Y, David A, David H (1998) Characterization
2 of a germin-like protein gene expressed in somatic and zygotic embryos of pine (*Pinus caribaea*
3 Morelet). *Plant Mol Biol* 38: 1179–1190
- 4 Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto Encyclopedia of
5 Genes and Genomes. *Nucleic Acids Research* 27: 29-34
- 6 Pan Z, Guan R, Zhu S, Deng X (2009) Proteomic analysis of somatic embryogenesis in Valencia
7 sweet orange (*Citrus sinensis* Osbeck). *Plant Cell Rep* 28: 281-289
- 8 Pan Z, Zeng Y, An J, Ye JXu Q, Deng X (2012) An integrative analysis of transcriptome and
9 proteome provides new insights into carotenoid biosynthesis and regulation in sweet orange fruits. *J*
10 *Proteomics* 75: 2670-2684
- 11 Park YS (2002) Implementation of conifer somatic embryogenesis in clonal forestry: technical
12 requirements and deployment considerations. *Annals For Sci* 59: 651-656
- 13 Park YS, Lelu-Walter M-A, Harvengt L, Trontin JF, McEacheron I, Klimaszewska K, Bonga JM
14 (2006) Initiation of somatic embryogenesis in *Pinus banksiana*, *P. strobus*, *P. pinaster*, and *P.*
15 *sylvestris* at three laboratories in Canada and France. *Plant Cell Tiss Org Cult* 86: 87-101
- 16 Pérez Rodríguez MJ, Suárez MF, Heredia R, Ávila C, Breton D, Trontin J-F, Filonova L, Bozhkov P,
17 Von Arnold S, Harvengt L, Cánovas FM (2006) Expression patterns of two glutamine synthetase
18 genes in zygotic and somatic pine embryos support specific roles in nitrogen metabolism during
19 embryogenesis. *New Phytologist* 169: 35-44
- 20 Petřek J, Vítěček J, Vlášínová H, Kizek R, Kramer KJ, Adam V, Klejdus B, Havel L (2005)
21 Application of computer imaging, stripping voltammetry and mass spectrometry to study the effect
22 of lead (Pb-EDTA) on the growth and viability of early somatic embryos of Norway spruce (*Picea*
23 *abies* L. Karst.). *Analytical Bioanalytical Chemistry* 383: 576-586
- 24 R Development Core Team (2011) R: A language and environment for statistical computing. R
25 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL [http://www.R-](http://www.R-project.org/)
26 [project.org/](http://www.R-project.org/)
- 27 Ramarosandratana A, Harvengt L, Bouvet A, Calvayrac R, Pâques M (2001) Effects of carbohydrate
28 source, polyethylene glycol and gellan gum concentration on embryonal-suspensor mass (ESM)
29 proliferation and maturation of maritime pine somatic embryos. *In Vitro Cell Dev Biol Plant* 37: 29-
30 34
- 31 Roberts A, Trapnell C, Donaghey J, Rinn J, Pachter L (2011) Improving RNA-Seq expression
32 estimates by correcting for fragment bias. *Genome Biol* 12: R22
- 33 Rode C, Lindhorst K, Braun H-P, Winkelmann T (2012) From callus to embryo: a proteomic view on
34 the development and maturation of somatic embryos in *Cyclamen persicum*. *Planta* 235: 995-1011
- 35 Sallandrouze A, Faurobert M, El Maataoui M, Espagnac H (1999) Two-dimensional electrophoretic
36 analysis of proteins associated with somatic embryogenesis development in *Cupressus sempervirens*
37 L. *Electrophoresis* 20: 1109-1119

- 1 Sheen J (1998) Mutational analysis of protein phosphatase 2C involved in abscisic acid signal
2 transduction in higher plants. *Proceedings National Acad Sci* 95: 975-980
- 3 Stasolla C, Kong L, Yeung EC, Thorpe T (2002) Maturation of somatic embryos in conifers:
4 morphogenesis, physiology, biochemistry, and molecular biology. *In Vitro Cell Dev Biol Plant* 38:
5 93-105
- 6 Stasolla C, Bozhkov PV, Chu T-M, Van Zyl L, Egertsdotter U, Suarez MF, Craig D, Wolfinger RD,
7 Von Arnold S, Sederoff RR (2004) Variation in transcript abundance during somatic embryogenesis
8 in gymnosperms. *Tree Physiol* 24: 1073-1085
- 9 Spencer MWB, Casson SA, Lindsey K (2007) Transcriptional profiling of the arabidopsis embryo.
10 *Plant Physiol* 143: 924-940
- 11 Tereso S, Zoglauer K, Milhinhos A, Miguel C, Oliveira M (2007) Zygotic and somatic embryo
12 morphogenesis in *Pinus pinaster*: comparative histological and histochemical study. *Tree Physiol* 27:
13 661-669
- 14 Teyssier C, Grondin C, Bonhomme L, Lomenech A-M, Vallance M, Morabito D, Label P, Lelu-
15 Walter M-A (2011) Increased gelling agent concentration promotes somatic embryo maturation in
16 hybrid larch (*Larix x eurolepis*): a 2-DE proteomic analysis. *Physiol Plant* 141: 152-165
- 17 Teyssier C, Maury S, Beaufour M, Grondin C, Delaunay A, Le Metté C, Ader K, Cadene M, Label P,
18 Lelu-Walter M-A (2013) In search of markers for somatic embryo maturation in hybrid larch (*Larix*
19 *x eurolepis*): global DNA methylation and proteomic analyses. *Physiol Plant* (in press) doi:
20 10.1111/ppl.12081
- 21 Thibaud-Nissen F, Shealy RT, Khanna A, Vodkin LO (2003) Clustering of microarray data reveals
22 transcript patterns associated with somatic embryogenesis in soybean. *Plant Physiol* 132: 118-136
- 23 Trontin J-F, Harvengt L, Garin E, Vernaza ML, Arancio L, Hoebeke J, Canlet F, Paques M (2002)
24 Towards genetic engineering of maritime pine (*Pinus pinaster* Ait.). *Annals of Forest Science* 59:
25 687-697
- 26 Umezawa T, Sugiyama N, Mizoguchi M, Hayashi S, Myouga F, Yamaguchi-Shinozaki K, Ishihama
27 Y, Hirayama T, Shinozaki K (2009) Type 2C protein phosphatases directly regulate abscisic acid-
28 activated protein kinases in Arabidopsis. *Proceedings National Acad Sci* 106: 17588-17593
- 29 Vandebussche F, Petraské J, Zadniková P, Hoyerová K, Pesek B, Raz V, Swarup R, Bennett M,
30 Zazimalova E, Benkova E, Van Der Straeten D (2010) The auxin influx carriers AUX1 and LAX3
31 are involved in auxin-ethylene interactions during apical hook development in *Arabidopsis thaliana*
32 seedlings. *Development* 137: 597-606
- 33 Vierstra RD (2003) The ubiquitin/26S proteasome pathway, the complex last chapter in the life of
34 many plant proteins. *Trends Plant Sci* 8: 135-142
- 35 Vernoud V, Horton AC, Yang Z, Nielsen E (2003) Analysis of the small GTPase gene superfamily of
36 arabidopsis. *Plant Physiol* 131:1191-1208

Version définitive du manuscrit publié dans / Final version of the manuscript published in :
Physiologia Plantarum, 2014, 152(1), 184-201 <http://dx.doi.org/10.1111/ppl.12158>

- 1 Vestman D, Larsson E, Uddenberg D, Cairney J, Clapham D, Sundberg E, von Arnold S (2011)
2 Important processes during differentiation and early development of somatic embryos of Norway
3 spruce as revealed by changes in global gene expression. *Tree Genet Genomes* 7: 347–362
4 Vondráková Z, Cvikrová M, Eliášová K, Martinčová O, Vágner M (2010) Cryotolerance in Norway
5 spruce and its association with growth rates, anatomical features and polyamines of embryogenic
6 cultures. *Tree Physiol* 30: 1335-1348
7 Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. *Current Opinion Plant Biol*
8 5: 218-223
9 Wiweger M, Farbos I, Ingouff M, Lagercrantz U, Von Arnold S (2003) Expression of Chia4-Pa
10 chitinase genes during somatic and zygotic embryo development in Norway spruce (*Picea abies*):
11 similarities and differences between gymnosperm and angiosperm class IV chitinases. *J Exp Bot* 54:
12 2691-2699
13 Yoshida T, Nishimura N, Kitahata N, Kuromori T, Ito T, Asami T, Shinozaki K, Hirayama T (2006.)
14 ABA-hypersensitive germination3 encodes a protein phosphatase 2C (AtPPP2CA) that strongly
15 regulates abscisic acid signaling during germination among arabidopsis protein phosphatase 2Cs.
16 *Plant Physiol* 140: 115-126
17 Zhang Y, Zhang S, Han S, Li X, Qi L (2012) Transcriptome profiling and in silico analysis of somatic
18 embryos in Japanese larch (*Larix leptolepis*) *Plant Cell Rep.* 31: 1637-1657
19 Zhen Y, Zhao Z-Z, Zheng R-H, Shi J (2012) Proteomic analysis of early seed development in *Pinus*
20 *massoniana* L. *Plant Physiol Biochem* 54: 97-104

21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **1 Supporting information**

4 Additional Supporting Information may be found in the online version of this article:
5
6

7 **4 Appendix S1** Identification of differentially expressed transcripts in immature somatic embryos of
8 maritime pine matured for 1 week on either 4 or 9 g L⁻¹ gellan gum.
9

10 **6 Appendix S2** Functional annotation of assembled sequences (significant transcripts) based on gene
11 ontology (GO) categorization.
12

13 **10 Appendix S3** Representative 2-D PAGE map for proteins from immature somatic embryos of
14 maritime pine matured for 1 week on high gellan gum (9 g L⁻¹).
15

16 **12 Appendix S4** Identification of differentially expressed proteins in immature somatic embryos of
17 maritime pine matured for 1 week on either 4 or 9 g L⁻¹ gellan gum.
18

19 **15 Appendix S5** Assignment of the putative proteins identified in immature somatic embryos of
20 maritime pine matured for 1 week to functional categories, according to Blast2GO
21 (http://www.blast2go.com/b2ghome).
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Carbohydrate and protein content in maritime pine immature somatic embryos of maritime pine matured for one week in the presence of either 4 or 9 g L⁻¹ gellan gum.

Compound ($\mu\text{g g}^{-1}$ DW)*	Gellan gum concentration (g L ⁻¹)	
	4	9
<u>Carbohydrate</u>		
Fructose (Fru)	27.44 \pm 2.22 ^a	17.27 \pm 1.39 ^b
Glucose (Glu)	26.60 \pm 2.86 ^a	19.29 \pm 0.71 ^b
Sucrose (Su)	6.88 \pm 0.14 ^a	8.29 \pm 1.24 ^a
<i>Total soluble carbohydrates</i>	60.91 \pm 5.17 ^a	44.85 \pm 2.65 ^b
<i>Ratio (Fru+Glu/Su)</i>	7.9 ^a	4.4 ^b
Starch	32.70 \pm 0.5 ^a	46.30 \pm 0.92 ^b
<u>Total protein</u>	654.40 \pm 81.72 ^a	733.10 \pm 100.4 ^a

*Values are means of 3 (carbohydrates), 5 (total protein), or 6 (starch) measurements \pm confidence interval at the 5% error level. Within figure, different letters represent statistically significant differences

Table 2: Characteristics 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 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 over-expressed on 9G, 4>9 : protein over-expressed 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	<i>Camellia sinensis</i>	31 / 26.3	4.95 / 4.96
56	SC_STAD_RICCO.1.4	4>9	Acyl-[acyl-carrier-protein] desaturase	A9NVD4	<i>Picea sitchensis</i>	46.9 / 44.3	5.65 / 5.5
36	SC_00001024.1.1	4>9	Alanine aminotransferase	Q84UX4	<i>Oryza sativa</i>	53.6 / 52.3	8.01 / 5.9
30	SC_ADH3_SOLTU.1.15	4>9	Alcohol dehydrogenase	Q43025	<i>Pinus banksiana</i>	40.5 / 44.6	6.12 / 6.02
34	SC_ADH_FRAAN.16.30	4>9	Alcohol dehydrogenase	Q43025	<i>Pinus banksiana</i>	40.5 / 46.4	6.12 / 5.86
32	SC_00002407.1.1	4>9	At5g54100 (Similarity to stomatin like protein)	Q9LVW0	<i>A. thaliana</i>	43.6 / 47.2	9.11 / 5.96
59	SC_00001003.1.1	4>9	BAHD family acyltransferase, clade V	D8T9Y4	<i>S. moellendorffii</i>	49.4 / 55.4	5.91 / 5.5
10	SC_RRAA3_ARATH.2.3	4>9	Cell division protein FtsQ	A3TJS8	<i>Oryza sativa</i>	18 / 9.8	5.6 / 5.7
55	SC_CH62_CUCMA.1.9	4>9	Chaperonin CPN60-2, mito.	Q05046	<i>Cucurbita maxima</i>	57.6 / 52.7	5.35 / 5.67
62	SC_CH62_CUCMA.2.9	4>9	Chaperonin CPN60-2, mito.	Q05046	<i>Cucurbita maxima</i>	57.6 / 61.5	5.35 / 5.47

60	SC_CH62_CUCMA.2.9	4>9	Chaperonin CPN60-2, mito.	Q05046	<i>Cucurbita maxima</i>	57.6 / 60.9	5.35 / 5.52
38	SC_00001092.1.1	4>9	DEAD-box ATP-dependent RNA helicase 2	Q5VNM3	<i>Oryza sativa</i>	45.6 / 48.7	5.98 / 5.96
78	SC_MCM31_MAIZE.1.5	4>9	DNA replication licensing factor MCM3 homolog 3	Q9SX03	<i>Zea mays</i>	85.2 / 86	5.92 / 5.91
33	SC_EFTM_ARATH.1.6	4>9	Elongation factor Tu, mito.	Q9ZT91	<i>A. thaliana</i>	44.1 / 45.1	5.93 / 5.9
45	SC_ENO1_HEVBR.7.31	4>9	Enolase 1	Q9LEJ0	<i>Picea sitchensis</i>	48.1 / 52.5	5.79 / 5.75
58	SC_GLGS_ARATH.1.1	4>9	Glucose-1-phosphate adenylyltransferase	B8LNV7	<i>Picea sitchensis</i>	57.9 / 51.2	5.99 / 5.46
37	SC_MDARS_CUCSA.1.10	4>9	Monodehydroascorbate reductase	A5JPK7	<i>Vitis vinifera</i>	47.3 / 47.1	5.93 / 5.94
29	SC_LOC100274026_MAIZE.1.3	4>9	NAD-dependent isocitrate dehydrogenase c	Q6ZI55	<i>Oryza sativa</i>	40.6 / 42.3	8.14 / 6.02
28	SC_PGKH1_ARATH.1.6	4>9	Phosphoglycerate kinase	F4MKM1	<i>Pinus pinaster</i>	52.9 / 44.6	8.83 / 6.11
31	SC_PGKH1_ARATH.1.6	4>9	Phosphoglycerate kinase	F4MKM1	<i>Pinus pinaster</i>	52.9 / 45	8.83 / 5.99
75	SC_PREP1_ARATH.1.3	4>9	Presequence protease 1, chloro./mito.	Q9LJL3	<i>A. thaliana</i>	111.5 / 92	5.03 / 5.35
50	SC_PDC2_ORYSI.1.19	4>9	Pyruvate decarboxylase isozyme 2	A2XFI3	<i>Oryza sativa</i>	64.8 / 63.7	5.6 / 5.74
51	SC_PDC2_ORYSI.1.19	4>9	Pyruvate decarboxylase isozyme 2	A2XFI3	<i>Oryza sativa</i>	64.8 / 63.2	5.6 / 5.75
52	SC_PDC2_ORYSI.3.19	4>9	Pyruvate decarboxylase isozyme 2	Q10MW3	<i>Oryza sativa</i>	64.7 / 63	5.53 / 5.72
53	SC_PDC2_ORYSI.3.19	4>9	Pyruvate decarboxylase isozyme 2	Q10MW3	<i>Oryza sativa</i>	64.7 / 63.3	5.53 / 5.71
44	SC_METK1_PICSI.34.55	4>9	S-adenosylmethionine synthase 1	A9NUH8	<i>Picea sitchensis</i>	42.9 / 47.4	5.42 / 5.7
46	SC_00154823.1.1	4>9	UDP-glucose 6-dehydrogenase	Q6RK07	<i>C. osmophloeum</i>	52.9 / 55.8	5.99 / 5.72
76	SC_DPE2_ARATH.1.1	9>4	4-alpha-glucanotransferase	DPE2 Q69Q02	<i>Oryza sativa</i>	108.3 / 90	6 / 5.61

			(Amylomaltase)						
43	SC_00000649.1.1	9>4	6-phosphogluconate decarboxylating	dehydrogenase.	B9RVA7	<i>Ricinus communis</i>	54.2 / 55.9	6.09 / 5.94	
40	SC_00000578.1.1	9>4	Acetyl-CoA carboxylase)	carboxylase (Biotin	O23960	<i>Glycine max</i>	52.3 / 57.6	5.82 / 6.11	
42	SC_SAHH_NICSY.1.26	9>4	Adenosylhomocysteinase		B8LLL7	<i>Picea sitchensis</i>	53.2 / 55.8	5.75 / 5.96	
35	SC_ADH1_PETHY.12.16	9>4	Alcohol dehydrogenase		Q43300	<i>Pinus banksiana</i>	40.4 / 47.3	5.76 / 5.82	
6	SC_00003988.1.1	9>4	AT3g23600/MDB19_9		Q9LUG8	<i>A. thaliana</i>	25.9 / 23.4	5.17 / 5.12	
22	SC_00002431.1.1	9>4	Benzoquinone reductase		A3F7Q3	<i>Picea sitchensis</i>	21.7 / 21.4	5.94 / 6.17	
70	SC_CALR_RICCO.2.19	9>4	Calreticulin		Q9FYV2	<i>Pinus taeda</i>	49.6 / 50.1	4.6 / 4.84	
69	SC_CALR_RICCO.2.19	9>4	Calreticulin		Q9FYV2	<i>Pinus taeda</i>	49.6 / 0.4	4.6 / 4.86	
49	SC_CH62_CUCMA.1.9	9>4	Chaperonin CPN60-2. mito.		Q05046	<i>Cucurbita maxima</i>	57.6 / 60.6	5.35 / 5.75	
17	SC_00004509.1.1	9>4	Expansin S2		Q39626	<i>Cucumis sativus</i>	25.4 / 21.2	9.39 / 6.02	
64	SC_00000119.1.1	9>4	FAM10 family protein At4g22670		Q93YR3	<i>A. thaliana</i>	46.6 / 57.5	4.83 / 5.31	
4	SC_GLGL1_BETVU.1.2	9>4	Glucose-1-P subunit	adenylyltransferase large	P55233	<i>Beta vulgaris</i>	57.7 / 32.9	6.05 / 4.76	
24	SC_00007308.1.1	9>4	GTP-binding nuclear protein Ran-A1		P41918	<i>Nicotiana tabacum</i>	25 / 23.8	6.38 / 6.12	
47	SC_ILV5_SPIOL.1.21	9>4	Ketol-acid reductoisomerase. chloro.		O82043	<i>Pisum sativum</i>	57.8 / 61.7	5.85 / 5.84	
9	SC_UMC2374_MAIZE.1.3	9>4	Probable 6-phosphogluconolactonase chloro.	4.	Q69NG5	<i>Oryza sativa</i>	28.3 / 23.8	5.65 / 5.34	
1	SC_PROF1_PHAVU.6.6	9>4	Profilin		A9NMR7	<i>Picea sitchensis</i>	14.3 / 6.1	4.61 / 4.81	

26	SC_00151055.1.1	9>4	Proteasome subunit beta type 1	P42742	<i>A. thaliana</i>	40.3 / 42.3	9.46 / 6.19
27	SC_00151055.1.1	9>4	Proteasome subunit beta type 1	P42742	<i>A. thaliana</i>	40.3 / 42.8	9.46 / 6.14
71	SC_PDI14_ARATH.1.13	9>4	Protein disulfide isomerase	A7KY78	<i>Oldenlandia affinis</i>	54.6 / 70.9	5.1 / 4.98
72	SC_PDI14_ARATH.1.13	9>4	Protein disulfide isomerase	A7KY78	<i>Oldenlandia affinis</i>	54.6 / 73.5	5.1 / 4.98
57	SC_PDI23_ORYSJ.1.5	9>4	Protein disulfide isomerase family	A9CPA7	<i>Glycine max</i>	47.7 / 53.9	5.41 / 5.58
81	SC_XYL1_ARATH.1.6	9>4	Putative alpha-xylosidase	Q8VWV9	<i>Pinus pinaster</i>	100.6 / 91	6.1 / 6.09
16	SC_EXB14_ORYSJ.1.2	9>4	Putative expansin-B14	Q6H677	<i>Oryza sativa</i>	25.5 / 27.2	5.3 / 5.97
23	SC_00146588.1.1	9>4	Putative germin-like protein 9-1	Q652Q1	<i>Vitis vinifera</i>	23.5 / 22.4	6.4 / 6.21
15	SC_SODF_NICPL.4.4	9>4	Superoxide dismutase	Q6QJL0	<i>Pinus pinaster</i>	28.6 / 21.1	7.18 / 5.89
21	SC_SODM_NICPL.1.5	9>4	Superoxide dismutase	A9NNA1	<i>Picea sitchensis</i>	25.5 / 20.4	8.7 / 6.08
66	SC_TBB2_ELEIN.1.4	9>4	Tubulin beta-2 chain	Q9ZPN9	<i>Eleusine indica</i>	50.4 / 54.8	4.71 / 5.15
8	SC_00009054.1.1	9>4	UMP - CMP kinase	O04905	<i>A. thaliana</i>	22.5 / 25.4	5.79 / 5.19

Figure legends:

Fig.1. Viability and starch accumulation in immature somatic embryos of maritime pine matured on either 4 g L⁻¹ or 9 g L⁻¹ gellan gum. Confocal images (projections of optical sections) of embryos stained with fluorescein diacetate and propidium iodide (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 1 or 2 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), 200 μm (c, d, e, f); m – meristematic cells, s – suspensor cells.

Fig. 2. Morphological appearance of embryonal masses of maritime pine matured for 12 weeks on either 4 g L⁻¹ (a) or 9 g L⁻¹ (b) gellan gum. Note embryonal mass proliferation on 4G (a1, a2) and presence of cotyledonary somatic embryos on 9G (b1, b2). Scale bar = 1cm (a1, b1), 1mm (a2, b2).

Fig.3. Biological characteristics of immature somatic embryos of maritime pine during maturation in the presence of either low (4 g L⁻¹, ○) or high (9 g L⁻¹, ■) gellan gum. A, Fresh weight (FW); B, Dry weight (DW); C, Water content. Bars represent confidence interval at the 5% error level. Letters represent statistical groups defined by the Multiple Comparisons of Means method ($P < 0.05$, $N = 10$).

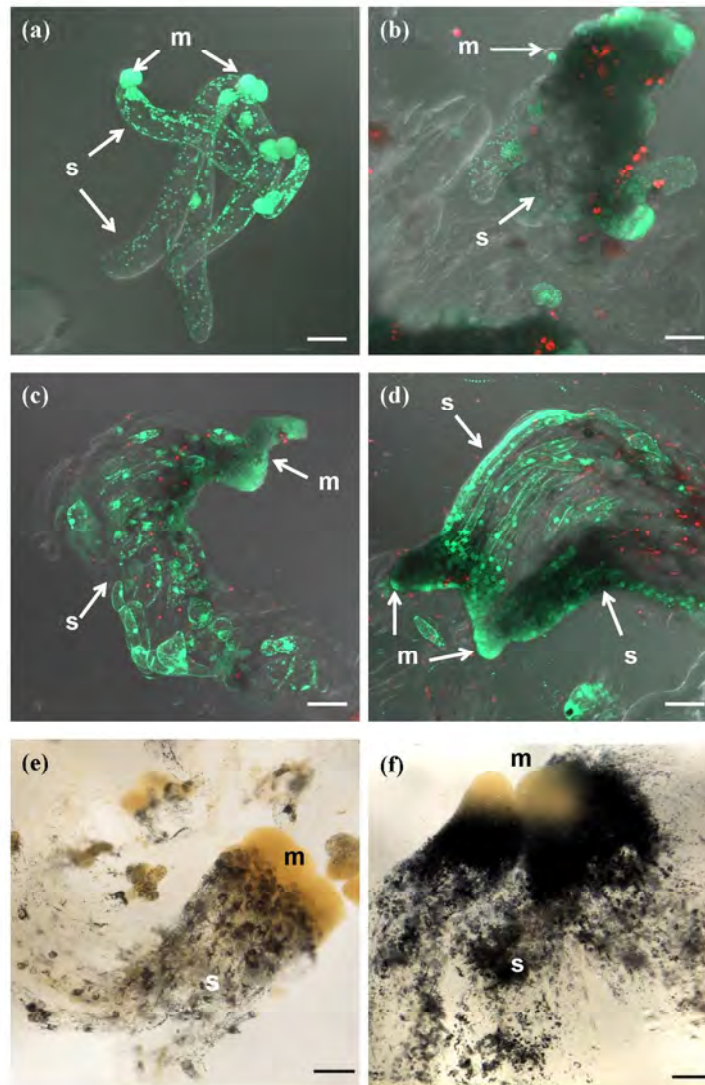
Fig.4. Endogenous abscisic acid (ABA) content of immature somatic embryos of maritime pine during maturation in the presence of either low (4 g L⁻¹, ○) or high (9 g L⁻¹, ■) gellan

1
2
3 gum. Bars represent confidence interval at the 5% error level. Letters represent statistical
4 groups defined by the Multiple Comparisons of Means method ($P < 0.05$, $N = 10$).
5
6
7

8 **Fig.5.** Gene ontology (Biological Process categories at level 2) terms found to be significantly
9 enriched after 1 week of maturation with either low (4 g L^{-1} , □) or high (9 g L^{-1} , ■) gellan
10 gum. Only GO categories for which >5 genes were represented are included. To test
11 whether Gene Ontology terms were enriched in a test group, when compared to a reference
12 group, Fisher's Exact Test with Multiple Testing Correction of FDR ($\text{FDR} < 0.05$, Benjamini
13 and Hochberg, 1995) was used.
14
15
16
17

18 Numbers of sequences: Numbers of sequences involved in each biological process.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Fig.1



Viability and starch accumulation in immature somatic embryos of maritime pine matured on either 4 g L-1 or 9 g L-1 gellan gum. Confocal images (projections of optical sections) of embryos stained with fluorescein diacetate and propidium iodide (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 1 or 2 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

Comment citer ce document :

Morel, A., Teyssier, C., Trontin, J.-F., Eliášová, K., Pešek, B., Beaufour, M., Morabito, D., Boizot, N., Le Mette, C., Belal-Bessai, L., Reymond, I., Harvengt, L., Cadene, M., Corbineau, F., Vágner, M., Label, P., Lelu-Walter, M.-A. (Auteur de correspondance) (2014). Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water

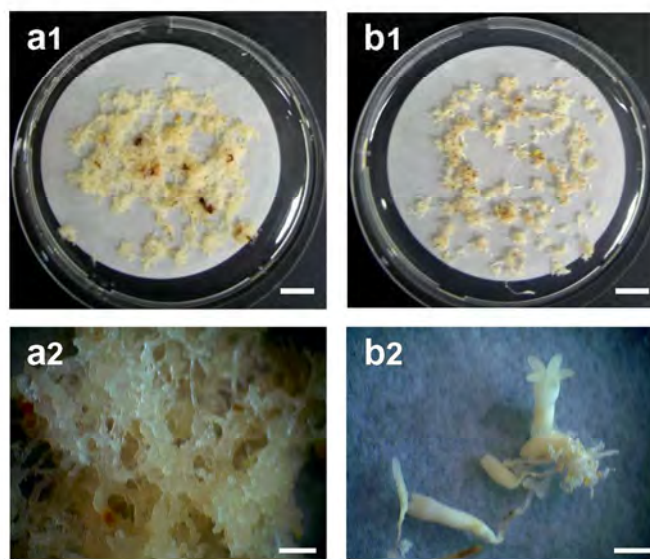
Version définitive du manuscrit publié dans / Final version of the manuscript published in :
Physiologia Plantarum, 2014, 152(1), 184-201 <http://dx.doi.org/10.1111/ppl.12158>

1
2
3 accumulation was evident. Scale bar =100 μm (a, b), 200 μm (c, d, e, f); m – meristematic cells, s –
4 suspensor cells.
5 190x275mm (220 x 220 DPI)
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Comment citer ce document :

Morel, A., Teyssier, C., Trontin, J.-F., Eliášová, K., Pešek, B., Beaufour, M., Morabito, D., Boizot, N., Le Mette, C., Belal-Bessai, L., Reymond, I., Harvengt, L., Cadene, M., Corbineau, F., Vágner, M., Label, P., Lelu-Walter, M.-A. (Auteur de correspondance) (2014). Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water



Morphological appearance of embryonic masses of maritime pine matured for 12 weeks on either 4 g L-1 (a) or 9 g L-1 (b) gellan gum. Note embryonic mass proliferation on 4G (a1, a2) and presence of cotyledonary somatic embryos on 9G (b1, b2). Scale bar = 1cm (a1, b1), 1mm (a2, b2).
190x275mm (223 x 224 DPI)

Comment citer ce document :

Morel, A., Teyssier, C., Trontin, J.-F., Eliášová, K., Pešek, B., Beaufour, M., Morabito, D., Boizot, N., Le Mette, C., Belal-Bessai, L., Reymond, I., Harvengt, L., Cadene, M., Corbineau, F., Vágner, M., Label, P., Lelu-Walter, M.-A. (Auteur de correspondance) (2014). Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water

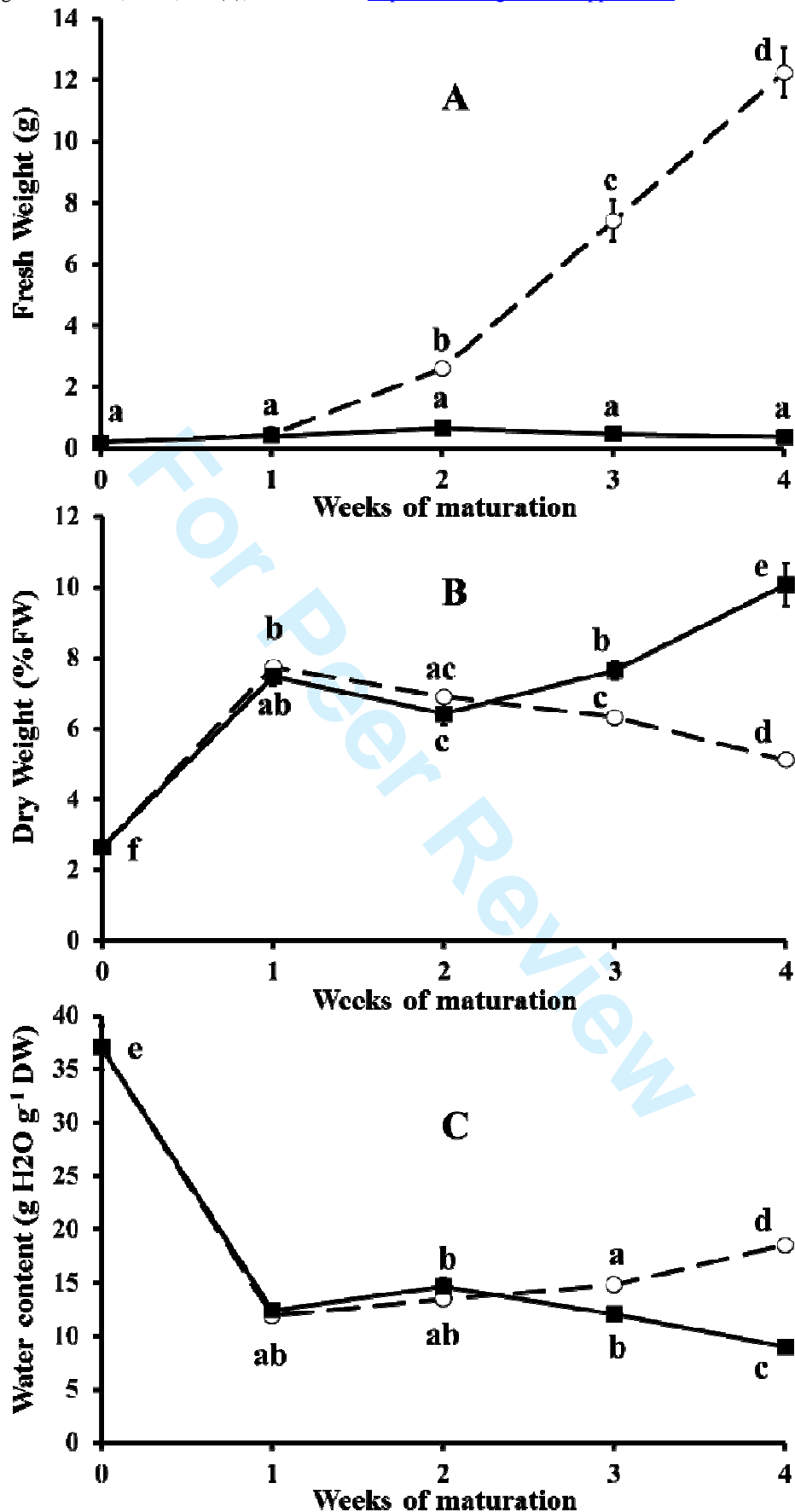


Figure 3.

Comment citer ce document :

Morel, A., Teyssier, C., Trontin, J.-F., Eliášová, K., Pešek, B., Beaufour, M., Morabito, D., Boizot, N., Le Mette, C., Belal-Bessai, L., Reymond, I., Harvengt, L., Cadene, M., Corbineau, F., Vágner, M., Label, P., Lelu-Walter, M.-A. (Auteur de correspondance) (2014). Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

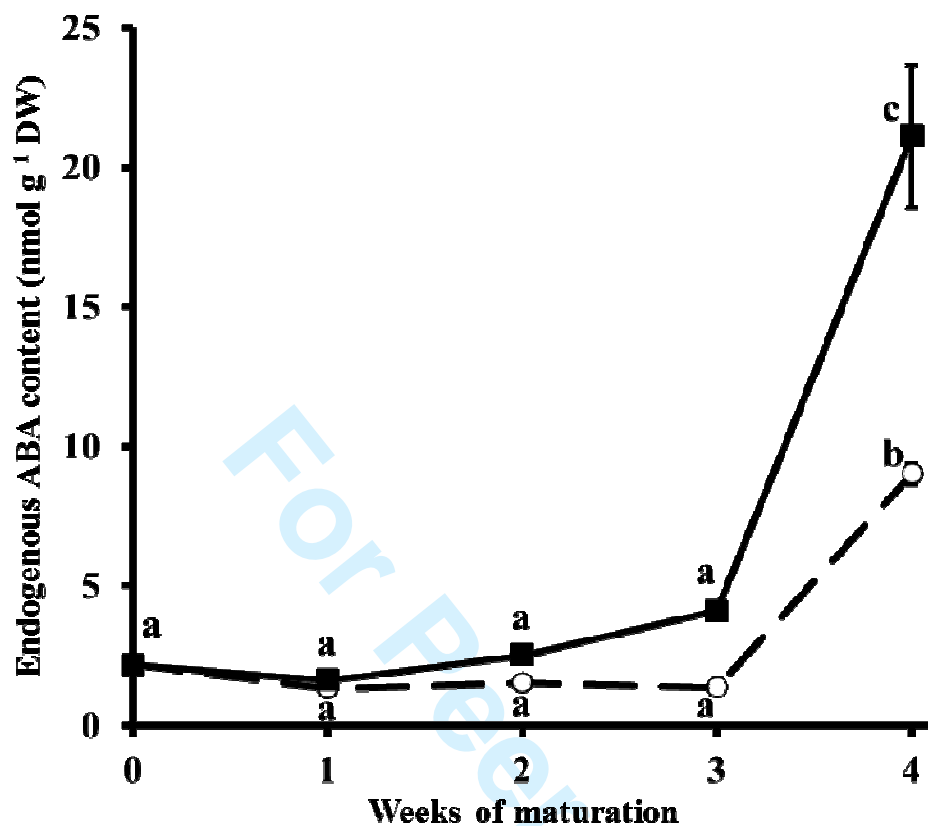


Figure 4

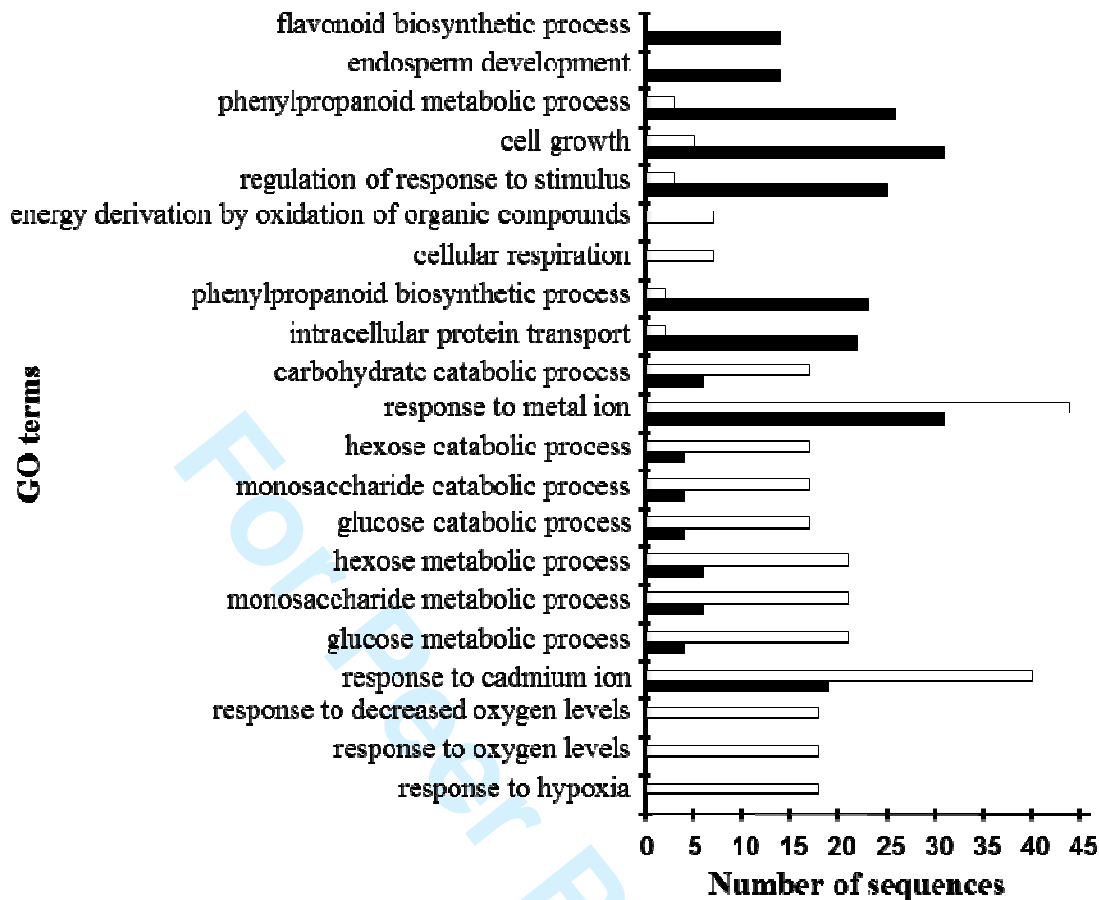


Figure 5.

FOR PEER REVIEW