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

Alexandre Morel, Jean-François Trontin, Françoise Corbineau, Anne-Marie Lomenech, Martine Beaufour, Isabelle Reymond, Claire Le Metté, Kevin Ader, Luc Harvengt, Martine Cadene, et al.

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

PLANTA --Manuscript Draft—
Corrected version R1
DOI 10.1007/s00425-014-2125-z

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; 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” full-sib 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) ²-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 (water content, fructose, galactose, glucose, maltose, 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. S1b).

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 proteins (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 (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 e 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 et al. 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 ²-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* (KumstýYová 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

Bonga JM, Klimaszewska K, von 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 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

KumstýÞová 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 M-A, 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 M-A, 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 M-A, 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 M-A (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*. Madeira: CIS Madeira, FIBA, AIMMP, CTBA. 256 pp.

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, 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)

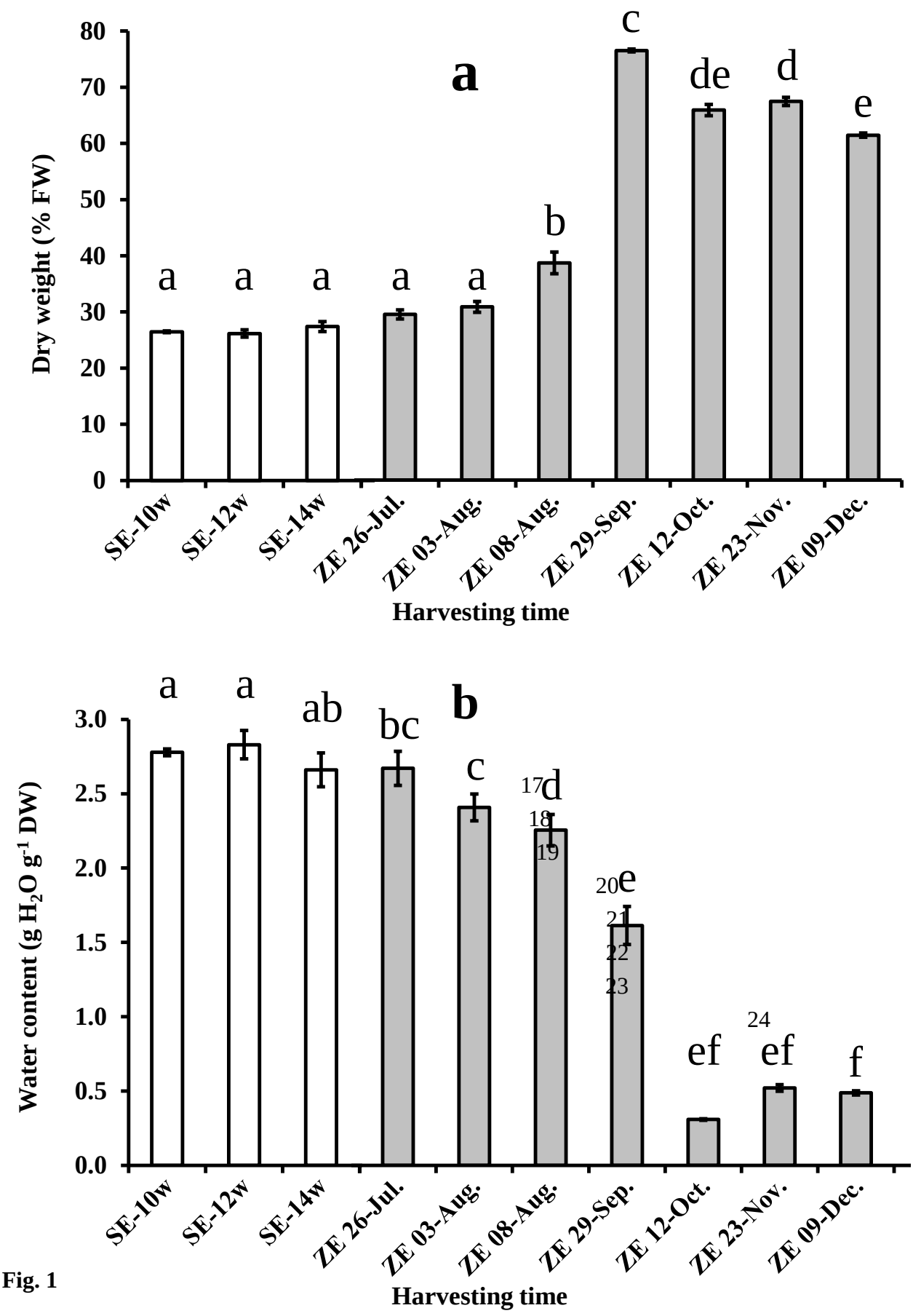

Fig. 1

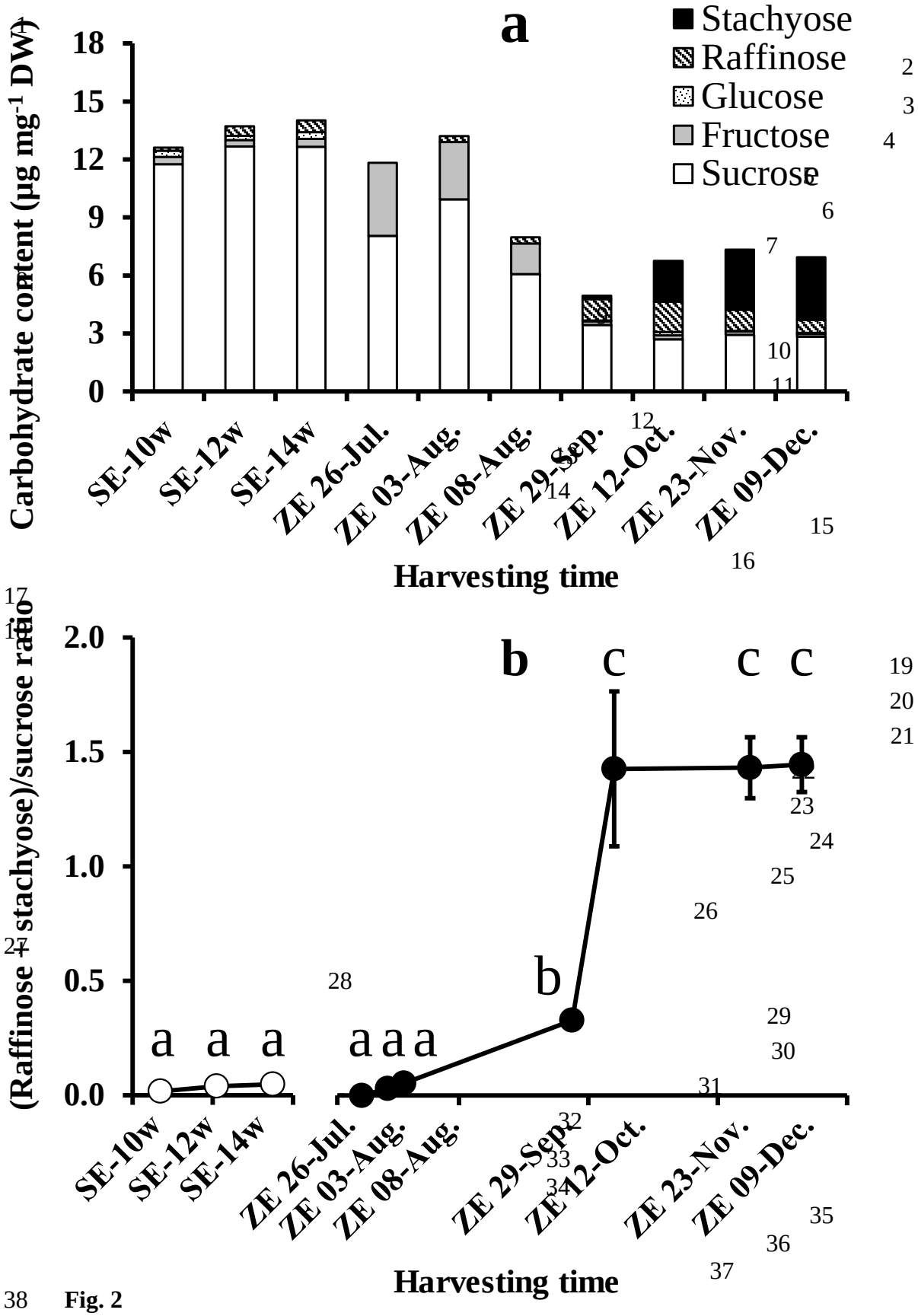

Fig. 2

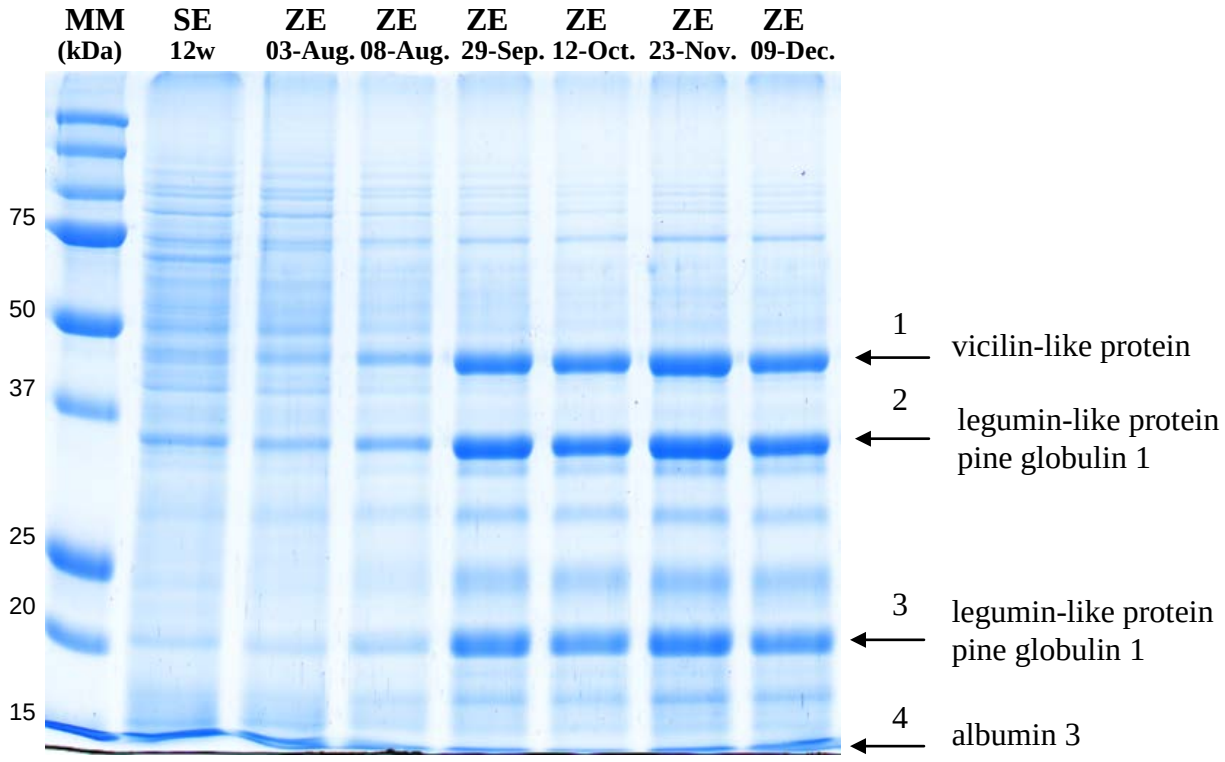

Fig. 3

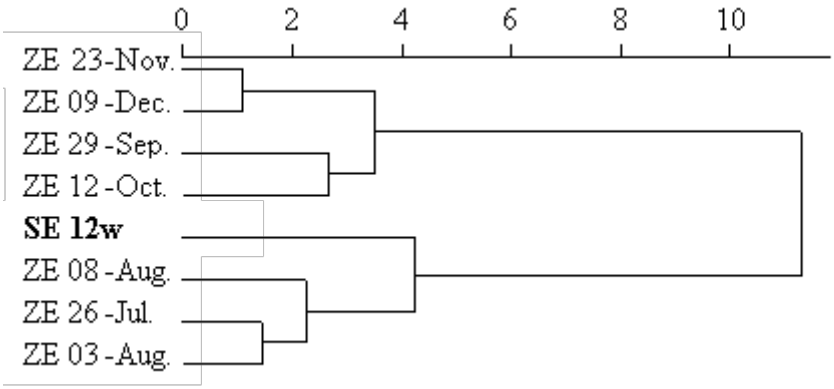

Fig 4

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 (µg 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		

1
2

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