Silva Lusitana ***13***(2): 203 - 216, 2005
© *EFN, Lisboa. Portugal*

Identification of Differentially Expressed Genes During Embryogenesis in Maritime Pine (*Pinus pinaster*)

Sónia Gonçalves*, John Cairney, Margarida Oliveira*** and Célia Miguel******
*PhD Student
Instituto de Tecnologia Química e Biológica. Quinta do Marquês, 2784-505 OEIRAS
**Associate Professor
School of Biology and Institute of Paper Science and Technology, 500 10th Street
Atlanta, NW 30332-0620, USA
***Assistant Professor
Faculdade de Ciências de Lisboa. Departamento Biologia Vegetal. Campo Grande
1749-016 Lisboa, Portugal/Instituto de Tecnologia Química e Biológica, Quinta do Marquês, 2784-505 OEIRAS
****Post-Doctoral Researcher
Instituto de Biologia Experimental e Tecnológica. Quinta do Marquês, 2784-505 OEIRAS

Abstract. Differential display of Poly (A^+) RNA isolated from zygotic embryos of maritime pine (*Pinus pinaster*) in 9 stages of development was used to isolate several cDNAs of differentially expressed genes during embryo development. Many cDNA sequences share low similarities to any known DNA or protein sequences but 5 transcripts are predicted to encode homologues of DNA-binding proteins, GTP-binding proteins, lipid transferases and dehydrogenases. A full-length cDNA clone was isolated for a gene (named *PpAAI-LTSS 1*) whose expression appears to be specific to middle stages of development. The full-length cDNA was also isolated for two clones (named *PpRab1* and *PpCR1*) specific to an early stage of embryo development.
Key words: pine; embryo development; gene expression

Identificação dos Genes Diferencialmente Expressos Durante a Embriogénese no Pinho Marítimo (*Pinus pinaster*)

Sumário. A técnica de "differential display" foi utilizada em RNA mensageiro extraído de embriões zigóticos de pinheiro-bravo em 9 estádios de desenvolvimento, de modo a isolar fragmentos de cDNA correspondentes a genes diferencialmente expressos durante o desenvolvimento do embrião. Muitos destes fragmentos possuem baixas homologias com sequências de DNA ou proteicas conhecidas, mas cinco transcritos codificam para homólogos de proteínas de ligação ao DNA, proteínas de ligação ao GTP, transferases lipídicas e desidrogenases. Um clone de cDNA de sequência completa foi isolado para um gene (denominado *PpAAI-LTSS1*), cuja expressão parece ser específica dos estádios intermédios de

---

1º Autor E-mail: soniag@itqb.unl.pt

desenvolvimento. A sequência completa de cDNA foi também obtida para outros dois clones (denominados *PpRab1* e *PpCR1*), específicos dos estádios iniciais de desenvolvimento do embrião.
Palavras-chave: pinheiro; desenvolvimento embrionário; expressão génica

Identification des Gènes Différentiellement Exprimés Pendant l'Embryogenèse en Pin Maritime (*Pinus pinaster*)

Résumé. Le criblage différentiel d'RNA messagers isolés d'embryons zygotiques de pin maritime (*Pinus pinaster*) à 9 stades différents du développement a été utilisé pour identifier des gènes dont l'expression varie au cours de l'embryogenèse. Une partie des cDNA isolés présentent une faible homologie avec des séquences connues. Cinq transcrits sont homologues à des gènes codant pour des protéines de liaison à l'DNA et au GTP, à des protéines de transfert de lipides et à des déshydrogénases. Un cDNA complet (nommé *PpAAI-LTSS 1*) dont l'expression semble être spécifique de la phase intermédiaire du développement a été cloné ainsi que deux autres cDNA complets (*PpRab1* et *PpCR1*) exprimés spécifiquement dans les phases précoces du développement.
Mots clés: pin; développement de l'embryon; expression des gènes

Introduction

Somatic embryogenesis (SE) has great potential for large-scale multiplication of elite genotypes. In addition, embryogenic tissue is readily cryo-preserved and revived, which makes somatic embryogenesis a highly desirable method for integration into forest tree breeding programs. However, there are still some limitations to the application of this technology in forestry. The number of plants obtained per embryogenic culture is still lower than expected and many genotypes are recalcitrant (ATTREE and FOWKE, 1993). The regeneration of viable plants from somatic embryos is also a common problem. The bottleneck in the regeneration process may occur at any of a number of stages, including maturation, germination, shoot apex elongation or acclimatisation (BECWAR and PULLMAN, 1995). In maritime pine, SE has been achieved (BERCETCHE and PÂQUES, 1995; LELU *et al.*, 1999; MIGUEL *et al.*, 2004). However, further progress in the application of this technology in pines and other gymnosperms has been limited due to the lack of basic knowledge of pine embryo development. Studies on hormone levels, metabolite pools (MINOCHA *et al.*, 1995; KAPIK *et al.*, 1995; FEIRER *et al.*, 1995), auxins biosynthesis, metabolism and transport in embryos (MICHALCZUK *et al.*, 1992; LIU *et al.*, 1993) have provided useful data but still little has been done to understand molecular processes and morphogenetic events underlying embryo formation. The comprehension of angiosperm embryogenesis has progressed through the study of mutants and functional analysis of the corresponding wild type genes (HARADA 1999, SCHRICK and LAUX, 2001). There are relatively few studies of genes specifically expressed during gymnosperm embryogenesis. Moreover, these studies have mainly focused on late stages of the embryogenesis process (reviewed by DONG and DUNSTAN, 2000). XU *et al.* (1997) described a technique that allowed differential display to be performed on minute amounts of tissue available from early stage pine embryos, thus allowing genes expressed at the earliest stages of

pine development to be identified and isolated (XU *et al*., 1997; CAIRNEY *et al*., 1999, 2000). A detailed functional and expression analysis of an early-expressed loblolly pine gene was thus possible (CIAVATTA *et al*., 2001, 2002). Following this line of investigation, the present study aims to provide useful molecular information about zygotic embryogenesis of maritime pine (*Pinus pinaster*).

Material and methods

Plant material

Maritime pine immature cones were collected from clones of one open-pollinated plus tree growing in the selected seed orchard at *Mata Nacional do Escaroupim*, Salvaterra de Magos, Portugal. The collection was performed weekly during the period of June 27th to July 31st, 2001. Cones were opened and seeds collected for isolation of embryos. The megagametophyte was opened and the dominant embryo or mass of embryos removed. Embryos were quickly observed through a dissecting microscope and evaluated for stage. Nine different stages of embryo development were considered (T0, T2, T3, T3B, T4, T4B, T5, T6 and T7), based on the staging system of PULLMAN and WEBB (1994) (Figure 1). Staged zygotic embryos were then placed in a cryostorage vial partially immersed in liquid nitrogen. Twenty to fifty similar-staged embryos were collected per vial. Frozen embryos were stored at –70ºC until analyses were performed.

mRNA extraction

Poly (A^+) RNA was extracted from zygotic embryos in different developmental stages using oligo (dT)-coated beads (Dynal Biotech ASA, Oslo, Norway), according to manufacturer's instructions. Messenger RNA (mRNA) concentration was then determined with RiboGreen® RNA quantitation reagent (Molecular Probes, Eugene, OR, USA), with the quantification performed at 480 /520 nm (excitation/emission).

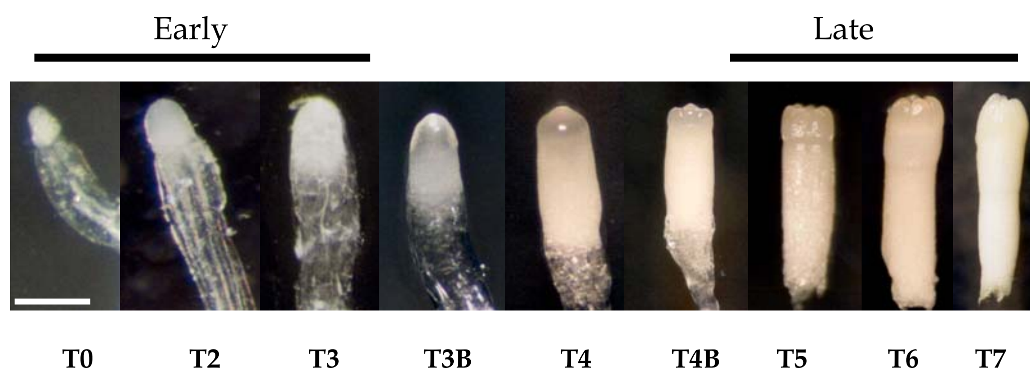

Figure 1 - Stages of embryo development used for poly (A^+) RNA extraction. Bar: T0, T2 = 300 µm; T3, T3B, T4 = 400 µm; T4B = 800 µm; T5, T6, T7 = 0.1 cm

Differential display (DDRT-PCR)

Differential display was carried out using the RNAimage® Differential Display Kit (GenHunter Corporation, Nashville, TN, USA) with some modifycations. A total of 2 nanograms of freshly diluted poly (A^+) RNA were reverse transcribed in a 20 μl reaction volume containing 2.5 mM dNTPs, 2 μM of $HT_{11}A$ or $HT_{11}G$ anchored primer in a buffer consisting of 25 mM Tris-Cl (pH 8.3), 37.6 mM KCl, 1.5 mM $MgCl_2$ and 5 mM DTT. One μL of MMLV reverse transcriptase was added after pre-incubation at 65° C for 5 minutes and at 37°C for 10 minutes. Reverse transcription was performed by incubation at 37°C for 50 minutes, and stopped by inactivating the MMLV reverse transcriptase at 75°C for 5 minutes. The products of reverse transcription were stored at 4°C for recent use or at –20°C for longer storage. Amplification of reverse transcribed RNA (cDNA) was performed in a total of 20 μL containing 250 μM dNTPs (Promega, Madison, WI, USA), 2 μM of the corresponding anchored primer used in the RT reaction, 2 μM of arbitrary primer (H-AP73, H-AP74, H-AP76 or H-AP77), 2 μL of cDNA, 1 U of Taq DNA polymerase (Qiagen, Valencia, CA, USA) in a buffer consisting of 100 mM Tris-Cl KCl (pH 8.4), 500 mM KCl, 15 mM $MgCl_2$ and 0.01% gelatin. Samples were amplified in 40 cycles of 94°C, for 30 sec, 40°C for 2 min and 72°C for 30 seconds. A final extension step of 5 min at 72°C completed the program. The amplified cDNA subpopulations of 3'-termini of mRNA generated by the pair of anchored and arbitrary primers were analyzed in 6% Polyacrilamide Pre-Cast Gels (Stratagene, La Jolla, CA, USA) after electrophoresis using CastAway Sequencing Device (Stratagene). Silver staining of the polyacrylamide gel was performed as described by the Silver Sequence DNA-Sequencing System (Promega). After staining, the gel was dried using a Cast Away GEL Dryer (Stratagene).

Cloning and sequencing of cDNA bands

The selection of the differential cDNA fragments was performed based on two expression patterns: transcript accumulation confined to specific stages of embryo development and up and down regulation in early and late embryogenesis. Bands were then recovered from the sequencing gel and eluted. The elutions were used as templates for reamplification using the same sets of primer pairs and conditions as for DDRT-PCR amplification. The reamplified fragments were electrophoresed in 1% agarose gels, stained with ethidium bromide and visualized under UV light. After excision from the gel and purification with Qiaquick PCR Purification Kit (Qiagen), the reamplified fragments were cloned into pGEM-Easy Vector (Promega). *E. coli* JM 109 competent cells were used for transformation. Transformants were analysed by PCR using M13 primers and plasmids containing an insert of the expected size were then isolated from bacterial cells with Wizard Miniprep DNA Purification Step (Promega) and sequenced with Sp6 and T7 primers. Computer-assisted searches for nucleotide sequences were carried out with the BLASTN and BLASTX programs (ALTSCHUL *et al*., 1997).

Isolation of full-length cDNA clones

The cDNA bands previously sequenced were used to design specific

primers for 5'end extension. Primers were designed to have a size of 23-28 base pairs (bp), GC content of 50-70% and Tm of at least 65ºC (Table I). All oligonucleotides were synthesized by Integrated Dna Technologies, Inc. (Skokie, IL, USA). The full-length clones were obtained using the 5'RACE (Rapid Amplification of cDNA Ends) Amplification kit from Clontech (Palo Alto, CA, USA), according to the manufacturer's instructions. Cloning and sequencing of the amplified cDNA fragment was performed as described above. Contig sequences were used to search for homologies in public databases.

Results and discussion

Differential Display

Differential display was performed to reveal differential accumulation of mRNA over zygotic embryo development in maritime pine. When cDNA molecules are radioactively labelled during differential display and subsequently separated by electrophoresis, misalignment of the autoradiographic image and the original gel can cause errors in cDNA isolation. In this work silver staining of gels was used instead of radioactive labelling because it is a simple, sensitive visualization technique that allows direct identification and isolation of DNA fragments of interest. For every reaction, a large number of bands of amplified cDNA were displayed in each lane. Several combinations of anchored and arbitrary primers were tested to reveal differential accumulation of mRNA in early and late embryogenesis and mRNA confined to specific stages of embryo development. From the combinations of two 3′ anchored primers ($HT_{11}A$ or $HT_{11}G$) and four arbitrary oligonucleotide primers (H-AP73, H-AP74, H-AP76 or H-AP77) used in these experiments, only the combination of $HT_{11}A$ or $HT_{11}G$ and H-AP74, H-AP76 or H-AP77 originated cDNA bands of interest. The selection of the cDNA bands was based on the expression patterns. Transcripts whose expression was confined to an early stage of development and transcripts with an up and down regulation over embryo development were selected (Figure 2).

Forty-nine cDNA bands of interest with a fragment length ranging from 100 to 600 bp were excised from the gel and successfully reamplified (using the same PCR conditions as for the DDRT-PCR amplification). Some researchers have reported difficulties in amplifying DNA excised from silver-stained gels, however, the protocols employed here have proven reliable and repeatable.

Table 1 – Sequences of the primers that allowed the determination of 5'-ends using RACE amplification

Clones	Sequences	Race Fragment (bp)
T0C11	5'-CAACTCCAGAAAAAACAAGCCCTCTCTCC-3'	**320**
T3Bx	5'-CCAGACAATACCACATTACCCCACTGCC- 3'	**538**
T2A2	5'-GCTTGTTATGGTCCTGAATCGCTCCTG-3'	**246**
T2E33	5'-TGGTAATGAAGCCTACTCCACCTGTCCATC-3'	**665**

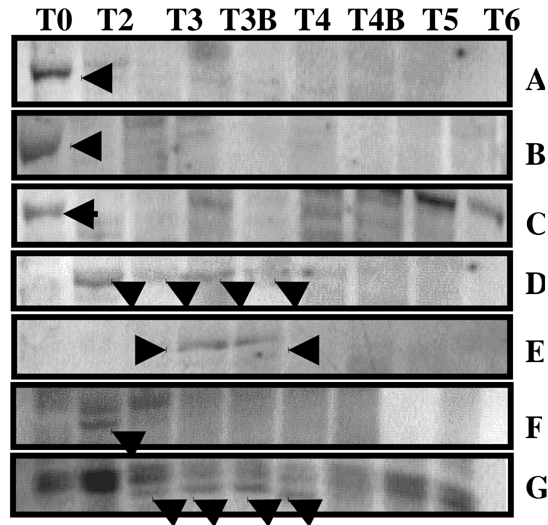

Figure 2 – Silver-stained cDNA fragments, obtained from differential display reactions run with poly (A)+ RNA from 9 stages of zygotic embryo development. **A** – Band T0D ($HT_{11}A$/H-AP74); **B** – band T0C ($HT_{11}G$/H-AP76); **C** – band T0F ($HT_{11}A$/H-AP74); **D** – bands T2E/T3E/T4E/T4BE ($HT_{11}A$/H-AP74); **E** - bands T3Bx and T4x ($HT_{11}G$/H-AP76); **F** – band T2A ($HT_{11}A$/H-AP74) and **G** – bands $T3_3$, $T3B_3$, $T4_3$ and $T4B_3$ ($HT_{11}A$/H-AP74). *Arrows indicate differentially displayed fragments*

After sequencing and searching through the databases, most sequences shared only low similarity to DNA or protein sequences in the Genbank. This may be due to the short length of most of the sequences obtained from DDRT-PCR, which contain mainly the non-conserved 3' UTR region. However, 5 sequences with strong similarities to proteins in public databases were identified (Table 2). These cDNAs are predicted to encode homologues of DNA-binding proteins, GTP-binding proteins, lipid transferases and dehydrogenases.

Fragments with a strong homology with sequences in the database or fragments showing an interesting expression pattern (such as stage-specific expression) were selected for further characterization.

Isolation and characterization of the full-length cDNA sequence

The 5'-RACE technique with specific primers was employed to obtain the full-length cDNA sequence of 6 fragments (T0C11, T0F 86, T3Bx, T3BB, T2A2 and T2E33 clones). Sequencing of the obtained PCR products revealed that only the fragments corresponding to clones T3Bx, T2A2 and T2E33 contained a complete full-length sequence. Database searches of contig sequences confirmed previously obtained homologies for DDRT-PCR sequences (data not shown).

Table 2 – Homologies obtained after Blast search for 5 clones

CLONE	**Seq. Length (bp)**	**Score / E-Value (EST database)**	**Tblastx**
T0C 11	323	545 /1 e-153	*Zea mays* pyruvate dehydrogenase (97 / 2e-38) Acession #AF069910
T3Bx 53	227	371 /1 e-100	*Arabidopsis thaliana* DIR1 protein (39 / 2e-05) Acession #AF342726
T2A2 2	666	609 / 1 e-171	*Solanum lycorpersicum* GTP-binding protein (272 / 6e-71) Acession # SLU38465
T0D 18	473	789 /0.0	*Zea mays* cytochrome b5 reductase (189/5e-46) Acession # AF0773772
T2E 33	455	771 / 0.0	*Zea mays* cytochrome b5 reductase (189/5e-46) Acession # AF0773772

Clone T3Bx is similar to a lipid transfer protein

The full-length of T3Bx 53, a clone with expression in middle stages of zygotic embryo development (Figure 2-E), was isolated. The gene corresponding to this clone was named *PpAAI-LTSS 1* (*Pinus pinaster* Alpha Amylase Inhibitor/ /Lipid Transfer/Seed Storage proteins No 1).

5'-RACE generated a full-length cDNA fragment with 647 bp, consisting of a 101 bp 5'-unstranslated region (UTR), a complete ORF of 315 bp encoding a polypeptide of 115 amino acids, followed by a 3'-UTR of 231 bp (Table 3).

The assembled sequence of *PpAAI-LTSS 1* contains an opening reading frame that encodes 88 amino acids with 41% identities and 60% similarities to protease inhibitor/seed storage/lipid transfer (LTP) protein family from *Arabidopsis thaliana* (MALDONALDO *et al.*, 2002) (Figure 3). By performing a conserved domain search (MARCHLER-BAUER *et al.*, 2003), an Alpha Amylase Inhibitor, Lipid Transfer and Seed Storage proteins domain (AAI-LTSS) was identified.

The T3B/T4 stage, where expression of *PpAAI-LTSS 1* was identified, morphologically corresponds to the pre-cotyledonary stage, where the embryo head becomes very dense and opaque. Since the formation of cotyledons is a critical phase of somatic embryo development, this gene will be further studied in detail.

Clone T2A2 encodes a GTP-binding protein

Clone T2A 2 is expressed in the early T2 stage of embryo development (Figure 2-F) and it is also the longest clone selected in this round of DDRT-PCR (666 bp). Due to its homologies with Rab proteins of GTP binding proteins family, this transcript was named *PpRab1* (*Pinus pinaster* Rab No1). The reconstituted full-length cDNA of *PpRab1* was 912 bp, consisting of a 352 bp 5'-unstranslated region (UTR), a complete ORF of 333 bp encoding a polypeptide of 120 amino acids, followed by a 3'-UTR of 227 bp (Table 4). Analysis of the deduced amino acid sequences revealed that this protein contained the G1-G5 conserved domains of the GTPase superfamily [GDSGVGKS ($Gx_4GK(S/T)$); DRYIST ($D-(X)_n)-T$); DTAG (DX_2G); NKCD (N/T)(K/Q)XD and FLETSAK (OOE(A/C/S/T)SA(K/L)], according to BOURNE *et al.*, 1991.

Table 3 – Nucleotide sequence of full-length cDNA and amino acid sequence of the predicted protein of PpAAI-LTSS 1

	Nucleotide [ϒ] / Amino Acid Sequence [ϒ1] (5'-3')
Full-length cDNA	TCTAATACGACTCACTATAGGGCAAGCAGTGGTAACAACGCAGAGTACGGGGGAATG CAGGAGCGCGCAGCACCCTACTGGTTCTCAAGTTTATTAGCAAG***ATGAAGGCGATGAA*** ***GATAATTATTGGCTTGGCAATAGTTACAGTAGCATTGCTCGCTAGCATTTCTTTGGGCAG*** ***AGGGGCCGTTCAGATCTGCAACGTGAGTCAGGATGACCTAATGCCATGTAAGCCCGCAG*** ***TCACACAGCCTCCTGCACAGCCTGTTGAAGCCTGTTGCAGCGTGTTAAGCACTGCAAATC*** ***TTACGTGCTTCTGTGAATTCGGGAATGACTATCCCTCCCTTCTTCGCATGTTTGGAATCGA*** ***TCCGGATCTTGCCAAGGCTTTGCCTGGAGAGTGCAAGTTAAATTCTCCTCCGGGCTGCTG*** ***A***ATGTTGAACTTGTTATAGTATGGAATAAGACGGCCTGTGGAGTCTGGCCATGAGGTCT CCATTGCTTATAATCTTGATTTTTCAGCGTTTATATGGCAGTGGGGTAATGTGGTATTGT CTGGTGTGGAGTACTAGCTTTTTCTGTAATGTAATAGCTCGCTGTATGCACTACCTATA ATATAGGTATAAAAAAATTAATGTTTCAGATTTCTACTCAAAAAAAAAAAAGCTT
Protein	MKAMKIIIGLAIVTVALLASISLGRGAVQICNVSQDDLMPCKPAVTQPPAQPVEACCSVLST ANLTCFCEFGNDYPSLLRMFGIDPDLAKALPGECKLNSPPGC

[ϒ] Open-reading frame is highlighted

```
gi|18422920|ref|NP_568699.1|    Expressed    protein;    protein    id:
At5g48485.1, supported by cDNA:13962., supported by cDNA: gi_17065561
[Arabidopsisthaliana]
 gi|17065562|gb|AAL32935.1|   Unknown protein [Arabidopsis thaliana]
 gi|18623490|gb|AAL76110.1|   DIR1 protein [Arabidopsis thaliana]
          Length = 102

 Score = 77.4 bits (189), Expect = 1e-13
 Identities = 41/108 (37%), Positives = 64/108 (59%), Gaps = 1/108
(0%)
 Frame = +3

Query:
93  ISKMKAMKIIIGLAIVTVALLASISLGRGAVQICNVSQDDLMPCKPAVTQP-PAQPVEAC 269
    ++  KA  +++ + IV +A+L   S+   A+ +C +SQD+L  CKPAV++  P  P + C
Sbjct:
1   MASKKAAMVMMAM-IVIMAMLVDTSV---AIDLCGMSQDELNECKPAVSKENPTSPSQPC 56

Query:
270 CSVLSTANLTCFCEFGNDYPSLLRMFGIDPDLAKALPGECKLNSPPGC 413
    C+ L  A+  C C + N     L  FG+DP+LA ALP +C L + P C
Sbjct:
57  CTALQHADFACLCGYKNS--PWLGSFGVDPELASALPKQCGLANAPTC 102
```

Figure 3 - Amino acid alignment of putative PpAAI-LTSS 1 protein with the DIR 1 protein from *Arabidopsis thaliana* (MALDONADO *et al.*, 2002)

A sequence similarity search of GenBank databases showed that *PpRab1* has 85% similarities and 77% identities to a *Pisum sativum* GTP-binding-protein related to Rab1 (NAGANO *et al.*,1993) (Figure 4). High similarities with *Petunia hybrida*, *Lotus corniculatus*, and *Arabidopsis thaliana* Rab1 GTP-binding protein were also obtained.

<u>Clone T2E33 has homologies with a cytochrome reductase</u>

Clone T2E 33 is expressed from the early T2 stage of embryo development until stage T4B (Figure 2-D). The analysis of the contig sequence revealed an 1127 bp full-length fragment consisting of a 217 bp 5'-unstranslated region (UTR), a complete ORF of 666 bp encoding a polypeptide of 242 amino acids, followed by a 3'-UTR of 244 bp (Table 5). This predicted protein has a high match (76% identity and 86% similarity) with a *Zea Mays* cytochrome b5 reductase isoform II (BAGNARESI *et al.*, 1999) (Figure 5). Several conserved domains were also identified within the sequence: Oxidoreductase NAD-binding (NAD-binding), Oxidoreductase FAD-binding domain (NAD-binding) and Flavodoxin reductases (ferredoxin-NADPH reductases). This clone was named *PpCR 1* (*Pinus pinaster* Cytochrome Reductase No 1).

Table 4 - Nucleotide sequence of full-length cDNA and amino acid sequence of predicted protein of PpRab1

	Nucleotide [ϒ] / Amino Acid Sequence [ϒ1] (5'-3')
Full-length cDNA	GTCCCCGAGTCTGCAAATCATTTCATTTCGTCAGCCATTTCCAGATCTGTGTACAGCTG AAGAGTGGCTTCAACATGACTACGGAATACGACTACTTGTTCAAGCTTTTGCTGATTGG TGATTCTGGGGTTGGAAAGTCATGTTTGCTACTGCGGTTTGCTGATGATTCATACGTTGG CAGTTACATCAGCACCATTGGTGTAGATTTCAAAATCAGAACAATTGAGTTGGATGGC AAGGCAATCAAGCTTCAAATTTGGGATACTGCAGGACAGGAGCGATTCAGGACCATA ACAAGCAGCTATTACCGCGGAGCTCACGGAATCATTATTGTTTATGATGTGACAGAT ***ATGGACAGTTTTAATAATGTCAAGCAATGGCTTAGTGAGATTGATAAATACGCAAGTGAC*** ***AATGTGAGCAAACTTCTAGTAGGAAACAAATGTGACTTGACAGAGAAACAAGCTGTTGA*** ***TCAGCAAATGGCAAAGTCCTTTGCAGATGAACTGGGGATCCCATTTCTAGAAACAAGTG*** ***CGAAAGATGCATCTAATGTAGAACAAGCATTTCTTACCATGGCTGGGGAAGTAAAGAGG*** ***AGGATGGCAAGTCAACTGACAGAGGGCAACAGGAAAGCGAACACTGTTGAAATGAAAG*** ***GACAGCCCCTTCAACAGAAAGGCGGCTGTTGTAGCTAG***AAATTTTTGTGGGTCTCATTTT GTGATTTATCTTAGGGTCTTCAAATTCTAAAAGGGATCAGCAAGATTTTCTTATTATTGT TCTACTGAGTTGAATATTGACATGGCATCAATACTTGATCCTCTCTGCTGGGAACTCTA GATCTTGCTAGTAAGCTGTTAACAATGTAGAATTGATGTATTTTATATTATTCTATCTTC TGTTGGAAAATAAAAAAAAAAAAGCTT
Protein	MTTEYDYLFKLLLI**GDSGVGKS**CLLLRFADDSYV**DRYIST**IGVDFKIRTIELDGKAIKLQXW**D** **TAG**QERFRTITSSYYRGAHGIIIVYDVTDMDSFNNVKQWLSEIDKYASDNVSKLLVG**NKCD**L TEKQAVDQQMAKSFADELGIP**FLETSAK**DASNVEQAFLTMAGEVKRRMASQLTEGNRKA NTVEMKGQPLQQKGGCCS

[ϒ] Open-reading frame in italic
[ϒ1] Conserved domains G1-G5 are highlighted

```
gi|541978|pir||S41430   GTP-binding protein, ras-like
                           (clone vfa-ypt1) - fava bean
 gi|303734|dbj|BAA02118.1|   GTP-binding protein [Pisum sativum]
 gi|452359|emb|CAA82707.1|     guanine  nucleotide  regulatory  protein
[Vicia faba]
 gi|738942|prf||2001457K   GTP-binding protein
 gi|1098293|prf||2115367A   small GTP-binding protein
                                                        Length = 202
 Score =  321 bits (823), Expect = 8e-87
 Identities = 157/197 (79%), Positives = 176/197 (89%)
 Frame = +2

Query:
92  DYLFKLLLIGDSGVGKSCLLLRFADDSYVDRYISTIGVDFKIRTIELDGKAIKLQXWDTA 271
    DYLFKLLLIGDSGVGKSCLLLRFADDSY+D YISTIGVDFKIRT+E DGK IKLQ WDTA
Sbjct:
6   DYLFKLLLIGDSGVGKSCLLLRFADDSYIDSYISTIGVDFKIRTVEQDGKTIKLQIWDTA 65

Query:
272 GQERFRTITSSYYRGAHGIIIVYDVTDMDSFNNVKQWLSEIDKYASDNVSKLLVGNKCDL 451
    GQERFRTITSSYYRGAHGIIIVYDVTD +SFNNVKQWLSEID+YASDNV+KLLVGNKCDL
Sbjct:
66  GQERFRTITSSYYRGAHGIIIVYDVTDEESFNNVKQWLSEIDRYASDNVNKLLVGNKCDL 125

Query:
452 TEKQAVDQQMAKSFADELGIPFLETSAKDASNVEQAFLTMAGEVKRRMASQLTEGNRKAN 631
    TE +AV  + AK+FADE+GIPF+ETSAKD++NVEQAF+ MA  +K RMASQ T  N +
Sbjct:
126 TENRAVPYETAKAFADEIGIPFMETSAKDSTNVEQAFMAMASSIKERMASQPT-NNARPP 184

Query: 632 TVEMKGQPLQQKGGCCS 682
           TV+++GQP+ QK GCCS
Sbjct: 185 TVQIRGQPVGQKSGCCS 201
```

Figure 4 - Amino acid alignment of putative PpRab1 protein with the RAB GTP-binding protein from *Pisum sativum* (NAGANO *et al.*, 1993)

Clone T0C11 has an incomplete orf

T0C 11 is a cDNA band of a mRNA differentially expressed in stage T0 (Figure 2-B). Additional sequence information of the TOC11 mRNA was obtained by 5'RACE and, after assembling the sequences of overlapping clones into a contiguous sequence, we observed an open-reading frame from the 5' end encoding 108 amino acids (data not shown). The predicted T0C11 protein is 89% identical and 95% similar to a pyruvate dehydrogenase from *Zea mays* (Figure 6) but, through analysis of the homologies, the orf corresponding to T0C 11 is still incomplete.

Clones T0D 18 and T0F 86

Clones T0D 18 and T0F 86 have a T0 stage specific expression (Figure 2-A; 2-C). For these clones we found no homologue in the databases, but their expression was specific of stage T0, and therefore they were included in this study. Until now, it was not possible to isolate the full-length of both T0D 18 and T0F 86. A large number of differentially expressed bands in this particular stage

of embryo development (as well in the T2 stage) were identified, implying that many genes are being specifically expressed at early stages of embryo development. It is possible that these transcripts play an important role in supporting the ongoing developmental processes, and further work to obtain more information on these clones will be performed.

Conclusions

The DDRT-PCR technique is a very useful approach for gene discovery, in this case, genes involved in zygotic-embryo development. In this paper we have described some preliminary work. Five differential expressed bands were selected for further studies, due to their particular expression patterns or homologies obtained in the databases. From these bands the full-length sequence was already obtained for three clones. The next step of this work is to confirm the differential expression using quantitative real-time PCR. This technique will be applied to all clones of both zygotic and somatic embryos, in order to compare their expression. Further studies will probe spatial expression by *in situ* hybridisation.

Table 5 - Nucleotide sequence of full-length cDNA and amino acid sequence of predicted protein of PpCR1

	Nucleotide [ϒ] / Amino Acid Sequence [ϒ1] (5'-3')
Full-length cDNA	GACTAAAGGAGGGAGTTGCACCAGAGTAGCTGTAAGATGGTGAGCGTTGAAGAACT CCTTCAGCCAGACAAGGCTCCTCTCTTGGGTGTGGCCTTGGCTGTTGTGGCTGTAGCT GTGGGAACTATATATTTTATTTCAACCAGAAAGCCCAAACCTTGCTTGGATCCAGAG AATTGGAAGAAGTTTAAATTGGTTCAGCGCACTCAATTGAGCCATA***ATGTGGCTAAAT*** ***TTCGCTTTGCACTCCCAACTCCGGACATCTGTGTTGGGGCTTCCAATTGGCCAGCATATC*** ***AGCTGCCAGGGAAGGGATGCTGAAGGTGAAGAAGTCATCAAACCCTATACTCCTACTA*** ***CCTTGGATTCTGATGTCGGCTACTTTGATCTGGTGATTAAGATGTATCCTTTAGGGGGA*** ***ATGTCTCACCATTTCCGCAAATTACAGGAGGGAGAATACTTGGCCGTAAGAGGTCCCA*** ***AGGGACGTTTCAAGTATCAGCCAGGTCAAGTAAGAGCATTTGGAATGCTTGCAGGAGG*** ***TTCTGGTATAACTCCAATGTTTCAGGTGACCAGAGCCATACTAGAAAACCCAAAGGACA*** ***ACACAAACGTATATCTCATCTATGGGAATGTGACTTATGATGATATTCTTCTCAAGGATG*** ***AACTTGATGACTTAGTGAAGACTTATCCAGKMMGCTTCCAAGTTTACCACGTTCTTAAT*** ***CAGCCACCTGAAGGATGGACAGGTGGAGTAGGCTTCATTACCAAGGAGATGATTGAGA*** ***CTCATTTCCCTGCACCTGCATCAGACATTCAAATCTTGAGATGTGGTCCACCACCAATGA*** ***ATAAAGCAATGGCAAGCCATATTGACGGACTTGGATACACTAAAGAGATGCAGTTCCA*** ***ATTTTGA***GACTTACGGATTCTGACAAATACTGCCCATCTGTTGTAGTTAACGTCTTGCA AGTTGTCTCTTCTGCTTGTTTGCTCTTTCTTGTTACCACACCAATAATACTGATATGATT TAGATTCCCACCTTACGATAAGCAACATTTTTATCAAAGTGTTTGTACTCGTGATTTCA ATATGAAGTTTATGTGGACAACATTAGGGTAAATGGGAATCTATAATCTTTGAACAA TAAAAAAAAAAAAGCTT
Protein	MWLNFALHSQLRTSVLGLPIGQHISCQGRDAEGEEVIKPYTPTTLDSDVGYFDLVIKMYPL GGMSHHFRKLQEGEYLAVRGPKGRFKYQPGQVRAFGMLAGGSGITPMFQVTRAILENPK DNTNVYLIYGNVTYDDILLKDELDDLVKTYPXXFQVYHVLNQPPEGWTGGVGFITKEMIET HFPAPASDIQILRCGPPPMNKAMASHIDGLGYTKEMQFQF

[ϒ] Open-reading frame is highlighted

```
gi|17221690|gb|AAL36459.1|AF205603_1      cytochrome   b5   reductase
isoform II [Zea mays]
                           Length = 279

Score =  343 bits (880), Expect(2) = e-108
Identities = 159/209 (76%), Positives = 181/209 (86%)
Frame = +2

Query:
254 TSVLGLPIGQHISCQGRDAEGEEVIKPYTPTTLDSDVGYFDLVIKMYPLGGMSHHFRKLQ 433
    TSVLGLPIGQHISC+G+DA GEEVIKPYTPTTLDSD+G F+LVIKMYP G MSHHFR+ +
Sbjct:
71  TSVLGLPIGQHISCRGQDASGEEVIKPYTPTTLDSDIGSFELVIKMYPQGRMSHHFRETK 130

Query:
434 EGEYLAVRGPKGRFKYQPGQVRAFGMLAGGSGITPMFQVTRAILENPKDNTNVYLIYGNV 613
     G+Y++V+GPKGRFKY PGQVRAFGM+AGGSGITPMFQVTRAILENPKDNT V+LIY NV
Sbjct:
131 VGDYMSVKGPKGRFKYLPGQVRAFGMVAGGSGITPMFQVTRAILENPKDNTKVHLIYANV 190

Query:
614 TYXXXXXXXXXXXXVKTYPXXFQVYHVLNQPPEGWTGGVGFITKEMIETHFPAPASDIQI 793
    TY            K YP  F++Y+VLNQPPE W GGVGF++KEMI+TH PAPA+DIQ+
Sbjct:
191 TYEDILLKEELDGMAKNYPDRFKIYYVLNQPPEVWDGGVGFVSKEMIQTHCPAPAADIQV 250

Query: 794 LRCGPPPMNKAMASHIDGLGYTKEMQFQF 880
           LRCGPPPMNKAMA+H+DGLGYTKEMQFQF
Sbjct: 251 LRCGPPPMNKAMAAHLDGLGYTKEMQFQF 279
```

Figure 5 - Amino acid alignment of putative PpCR 1 protein with the cytochrome b5 reductase protein from *Zea mays* (BAGNARESI *et al.*, 1999)

```
gi|3851003|gb|AAC72194.1|     pyruvate  dehydrogenase  E1  beta  subunit
isoform 3 [Zea mays]
                                  Length = 374

Score =  193 bits (490), Expect = 2e-48
Identities = 96/107 (89%), Positives = 102/107 (95%)
Frame = +1

Query:
1   GISAEVINLRSIRPLDRATINASVRKTSRLVTVEEGFPQHGIGAEICASVVEESFEYLDA 180
    GISAEVINLRSIRPLDRA INASVRKT+RLVTVEEGFPQHGIGAEIC SVVEESFEYLDA
Sbjct:
268 GISAEVINLRSIRPLDRAAINASVRKTNRLVTVEEGFPQHGIGAEICMSVVEESFEYLDA 327

Query: 181 PVERITGADVPMPYAANLERLAVPQVEDIIRASXRACYRAVPMSAVA 321
           PVERI GADVPMPYAANLER+AVPQV+DI+RA+ RACYRAVPM+A A
Sbjct:
       328 PVERIAGADVPMPYAANLERMAVPQVDDIVRAAKRACYRAVPMAAAA 374
```

Figure 6 – Amino acid alignments of putative T0C11 protein with pyruvate dehydrogenase from *Zea mays* (THELEN *et al.*, 1999)

Acknowledgements

Estação Florestal Nacional (EFN) is gratefully acknowledged for making plant material available. This research was supported by Fundação para a Ciência e Tecnologia (FCT) and III Framework Programme of the European Community through grants SFRH/BD//3135/2000 and SFRH/BPD/5645/2001.

References

ALTSCHUL, S.F., MADDEN, T.L., SCHAFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W., LIPMAN, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25** : 3389-402

ATTREE, S.M., FOWKE, L.C., 1993. Embryogeny of gymnosperms: advances in synthetic seed technology of conifers. *Plant Cell Tiss. Org. Cult* **35** : 1-35

BAGNARESI, P., THOIRON, S., MANSION, M., ROSSIGNOL, M., PUPILLO, P., BRIAT, J.F., 1999. Cloning and characterization of a maize cytochrome-b5 reductase with Fe3+-chelate reduction capability. *Biochemical Journal* **338**(2) : 499-505

BECWAR, M.R., PULLMAN, G.S., 1995. Somatic embryogenesis in loblolly pine (*Pinus taeda* L.). In: Somatic Embryogenesis in Woody Plants. Eds. Jain S, Gupta P, Newton R.J. Kluwer, The Netherlands). **3 :** 287-301

BERCETCHE, J., PAQUES, M., 1995. Somatic embryogenesis in maritime pine (*Pinus pinaster*). In: Jain S., Gupta P.K. & Newton R.J. (eds) *Somatic embryogenesis in woody plants*, Vol. 3 (pp. 221-242). Kluwer Academic Publishers, Dordrecht

BOURNE, H.R., DAVID, A.S., MCCORMICK, F., 1991. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349** : 117-127

CAIRNEY, J., XU, X., PULLMAN, G.S., CIAVATTA, V.T., JOHNS, B., 1999. Natural and Somatic Embryo Development in Loblolly Pine: Gene Expression Studies Using Differential Display and DNA Arrays. *Applied Biochemistry and Biotechnology* **77-79** : 5-17

CAIRNEY, J., XU, N., MACKAY, J., PULLMAN, J., 2000. Transcript profiling: a tool to assess the development of conifer embryos. *In Vitro Cellular & Developmental Biology - Plant*. **36** : 155-162

CIAVATTA, V., MORILLON, R., PULLMAN, G., CHRISPEELS, M., CAIRNEY, J., 2001. An aquaglyceporin is abundantly expressed early in the development of the suspensor and the embryo proper of loblolly pine. *Plant Physiol* **127** : 1557-1567

CIAVATTA, V.T., EGERTSDOTTER, U., CLAPHAM, D., VON ARNOLD, S., CAIRNEY, J., 2002. A promoter from the loblolly pine PtNIP1: 1 gene directs expression in an early-embryogenesis and suspensor-specific fashion. *Planta* **215** : 694-698

DONG, J.Z., DUNSTAN, D.I., 2000. Molecular biology of somatic embryogenesis in conifers. In S.M. Jain S.M., & S.C. Minocha S.C. (eds), Molecular biology of woody plants. Kluwer Academic Publishers. Amsterdam. 2000. pp 51-87

FEIRER, R.P., 1995. The biochemistry of conifer embryo development: amino acids, polyamines and storage proteins. In: Mohan Jain S, Gupta P, Newton R (Eds), Somatic Embryogenesis in Woody Plants. Vol. 1, Kluwer Academic Publishers: 317-336

HARADA, J.J., 1999. Signalling in plant embryogenesis. *Curr Op Plant Biol* **2** : 23-27

KAPIK, R.J., DINUS, R.J., DEAN, J.F.D., 1995. Abscisic acid during zygotic embryogenesis in *Pinus taeda*. *Tree Physiol* **15** : 405-409

KIYOSUE, T., YAMAGUCHI-SCHINOZAKI, K., HIGASHI, K., SATOH, S., KAMADA, H., HARADA, H., 1993. cDNA cloning of ECP40, an embryogenic-cell protein in carrot, and its expression during somatic and zygotic embryogenesis. *Plant Mol Biol* **21** : 1053-1068

LELU, M.-A., BASTIEN, C., DRUGEAULT, A., GOUEZ, M.-L., 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

LIU, C.-M., XU, Z.-H., CHUA, N.-H., 1993. Auxin polar transport is essential for the establishment of bilateral symmetry during early plant embryogenesis. *Plant Cell* **5** : 621-630

MALDONADO, A.M., DOERNER, P., DIXON, R.A., LAMB, C.J., CAMERON, R.K., 2002. A putative lipid transfer protein involved in systemic resistance signalling in Arabidopsis. *Nature* **419**(6905) : 399-403

MARCHLER – BAUER, A., ANDERSON, J.B., DEWEESE-SCOTT, C., FEDOROVA, N.D., GEER, L.Y., HE, S., HURWITZ, D.I., JACKSON, J.D., JACOBS, A.R., LANCZYCKI, C.J., LIEBERT, C.A., LIU, C., MADEJ, T., MARCHLER, G.H., MAZUMDER, R., NIKOLSKAYA, A.N., PANCHENKO, A.R., RAO, B.S., SHOEMAKER, B.A., SIMONYAN, V., SONG, J.S., THIESSEN, P.A., VASUDEVAN, S., WANG, Y., YAMASHITA, R.A., YIN, J.J., BRYANT, S.H., 2003. CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* **31** : 383-387

MICHALCZUK, L., COOKE, T.J., COHEN, J.D., 1992. Auxin levels at differents stages of carrot somatic embryogenesis. *Phytochem* **31** : 1097-1103

MIGUEL, C., GONÇALVES, S., TERESO, S., MARUM, L., MAROCO, J., OLIVEIRA, M.M., 2004. Somatic embryogenesis from 20 open-pollinated families of Portuguese plus tree of maritime pine. *Plant Cell Tissue and Organ Culture* **76** : 121-130

MINOCHA, R., MINOCHA, S.C., SIMOLA, L.K., 1995. Somatic embryogenesis and polyamines in woody plants. In: Mohan Jain S, Gupta P, Newton R (eds), Somatic Embryogenesis in Woody Plants. Vol. 1, Kluwer Academic Publishers: 337-359

NAGANO, Y., MURAI, N., MATSUNO, R., SASAKI, Y., 1993. Isolation and characterrization of cDNAs that encode eleven small GTP-binding proteins from *Pisum sativum*. *Plant Cell Physiol* **34**(3) : 447-455

PULLMAN, G.S., WEBB, D.T., 1994. An embryo staging system for comparison of zygotic and somatic embryo development. In: Proceedings of the TAPPI R&D Division Biological Sciences Symposium, October 3-6, Minneapolis, MN. Technical Association of the Pulp and Paper Industry Press, Atlanta, GA, pp 31-34

SCHRICK, K., LAUX, T., 2001. Zygotic embryogenesis: developmental Genetics. *In;* Current Trends in the Embryology of Angiosperms. Eds. SS. Bhojwani & WY Soh. Kluwer Academic Publishers, Dordrecht. ISBN. pp. 249-277

THELEN, J.J., MIERNYK, J.A., RANDALL, D.D., 1999. Molecular cloning and expression analysis of the mitochondrial pyruvate dehydrogenase from maize. *Plant Physiol* **119**(2) : 635-644.

XU, N., JOHNS, B., PULLMAN, G.S., CAIRNEY, J., 1997. Rapid and Reliable Differential Display from Minute Amounts of Tissue: Mass Cloning and Characterization of Differentially Expressed Genes from Loblolly Pine Embryos. *Plant Mol. Biol. Rep.* **15**(4) : 377-391

Entregue para publicação em Abril de 2004
Aceite para publicação em Março de 2005