

UNIVERSIDADE ESTADUAL DE CAMPINAS

CENTRO DE BIOLOGIA MOLECULAR E ENGENHARIA GENÉTICA

Jorge Hernandez Fernandez

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE PROTEÍNAS E
GENES EXPRESSOS DIFERENCIALMENTE DURANTE O
DESENVOLVIMENTO DO EMBRIÃO ZIGÓTICO DE
*Araucaria angustifolia***

este exemplar corresponde à redação final da tese defendida pelo (a) candidato (a) Jorge Hernandez Fernandez e aprovada pela Comissão Julgadora. Laura M. M. Ottoboni

Tese apresentada ao Instituto de
Biologia para obtenção do Título de
Doutor em Biologia Celular e Estrutural
na área de Biologia Celular

Orientadora: Profa. Dra. Laura Maria Mariscal Ottoboni

UNIDADE	BE
Nº CHAMADA	UNICAMP
	H431
V	1X
TOMADA	PGI 48941
PROD	16.837102
	C <input type="checkbox"/> D <input checked="" type="checkbox"/>
PREÇO	R\$11,00
DATA	10/05/02
Nº CPD	

CM001672B6-B

BIB ID 239953

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP**

Fernandez, Jorge Hernandez

Identificação e caracterização de proteínas e genes expressos diferencialmente durante o desenvolvimento do embrião zigótico de *Araucaria angustifolia*/Jorge Hernandez Fernandez. -- Campinas, S.P.: [s.n.], 2001.

Orientadora: Laura Maria Mariscal Ottoboni
Tese (doutorado) – Universidade Estadual de Campinas.
Instituto de Biologia.

1. Coníferas. 2. Embriogénese. 3. Semente. I. Ottoboni, Laura Maria Mariscal. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

F3911

H431

Data da Defesa: 16/01/2002

BANCA EXAMINADORA

Dra. Laura Maria Mariscal Ottoboni (Orientadora)

Laura m. m. Ottoboni
Assinatura

Dra. Maricilda Palandi de Mello

Maricilda Palandi de Mello
Assinatura

Dr. Carlos Augusto Colombo

Dr. Carlos Augusto Colombo
Assinatura

Dra. Mirian Perez Maluf

Assinatura

Prof.Dr. José Camillo Novello

Prof. Dr. José Camillo Novello
Assinatura

Dra. Alba Chiesse da Silva

Alba Chiesse da Silva
Assinatura

Dr. Márcio José da Silva

Assinatura

2002.01.16

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

*Se a fé remove até
montanhas, a persistência
é o que torna o irreal
possível.*

Anônimo

Agradecimentos aos familiares, amigos e colegas de trabalho.....

À Laura pela orientação, interesse, oportunidade e paciência de desenvolver um projeto de pesquisa sobre Araucária.

Aos Drs. Carlos Augusto Colombo, Miriam Perez Maluf e Maricilda Palandi de Mello pela leitura cuidadosa, sugestões e críticas ao manuscrito.

Ao Carlos (de novo) e ao Anderson pelos resultados do seqüenciamento das amostras.

Ao Dr. Gonçalo A. de Souza Filho (UENF) pelo seqüenciamento do N^o-terminal das proteínas.

Ao Santoro (Bioquímica-IB) pela disposição, ajuda e paciência na análise dos géis 2D.

À FAPESP pelo apoio financeiro.

À Sandra, Tânia, Fabiana, Patrícia (secretaria do CBMEG), à Lilian (secretaria do DBCE) e ao Márcio (CBMEG) que sempre mostraram disponibilidade e boa vontade de ajudar em tudo.

A Renata, Renato, Rogério, Edna e Wagner pelo companheirismo e amizade no dia a dia do laboratório.

Ao FUNDÃO (Lu, Ricardo, Ana e Bianca) pelo apoio constante e ótimo “microclima” de trabalho.

A Vanildo pela companhia nas coletas de material em Santa Catarina, consultas informais sobre o tema e pela amizade.

À Mada pelo ótimo café.

A Tessy, Pedro e Camilo pelo “suporte familiar direto”.

Ao meu pai, por ter orgulho do “cientista maluco” da família.

À minha família, que mesmo depois de tantas mudanças geopolíticas, ainda consegue (mal) se manter em contato.

Ao governo cubano por não ter interferido em nada.

À Stefy com o maior amor, por achar ainda que eu sou um bom pai.

À minha Li, pelo apoio, amor e companhia nos piores e melhores momentos.

ÍNDICE

RESUMO	9
SUMMARY	10
INTRODUÇÃO	11
1. <i>Araucaria angustifolia</i>	13
1.1. Importância econômica da <i>A. angustifolia</i>	14
1.2. A embriogênese zigótica em <i>A. angustifolia</i>	16
2. Síntese e deposição de proteínas durante a embriogênese em coníferas	19
2.1. Classificação das proteínas de reserva	20
2.2. Proteínas de reserva das coníferas	20
3. Expressão gênica durante o desenvolvimento do embrião zigótico de coníferas	22
4. Referências Bibliográficas	25
OBJETIVOS	29
1. Objetivo geral	30
2. Objetivos específicos	30
CAPÍTULO 1	
Caracterização das principais proteínas de reserva da semente de <i>Araucaria angustifolia</i>	31
1. Introdução	32
2. Materiais e Métodos	34
2.1. Coleta do material biológico	34
2.2. Extração das proteínas de reserva	34
2.3. Extração de proteína total	35
2.4. SDS-PAGE	36
2.5. 2D-PAGE	36
2.6. Transferência das proteínas para membrana de PVDF	37
2.7. Análise das seqüências do N ^o -terminal das proteínas	38
3. Resultados e discussão	38
3.1. Caracterização das proteínas de reserva de <i>A. angustifolia</i>	38
3.2. Análise da síntese de proteínas durante o desenvolvimento da semente de <i>A. angustifolia</i>	42
3.3. Análise através de 2D-PAGE das proteínas de reserva de <i>A. angustifolia</i>	45
3.4. As principais proteínas de reserva da semente de <i>A. angustifolia</i> pertencem ao grupo das vicilinas 7S	47
5. Referências Bibliográficas	50
ANEXO 1	53

CAPÍTULO 2

Identificação de genes expressos diferencialmente durante o desenvolvimento do embrião zigótico de *Araucaria angustifolia*

1. Introdução	57
2. Materiais e Métodos	59
2.1. Coleta de sementes	59
2.2. <i>Display</i> diferencial de mRNA	59
2.2.1. Isolamento de RNA e tratamento com DNase	59
2.2.2. Síntese de cDNA	60
2.2.3. PCR a partir do cDNA e eletroforese das amostras amplificadas	61
2.2.4. Isolamento, eluição e reamplificação das bandas diferenciais	61
2.3. Confirmação da expressão diferencial das bandas isoladas nos experimentos de <i>display</i> diferencial de mRNA	62
2.3.1. Síntese de cDNA visando a obtenção de sondas radioativas	62
2.3.2. Transferência das bandas isoladas dos géis de <i>display</i> diferencial de mRNA para membranas de nylon	62
2.3.2.1. Transferência do DNA para membranas de nylon pela técnica de <i>slot blot</i>	62
2.3.2.2. Transferência do DNA para membranas de nylon pela técnica de <i>macroarray</i>	63
2.3.3. Obtenção de sondas radioativas e hibridização das membranas	63
2.3.4. Análise dos resultados de hibridização	64
2.4. Clonagem e sequenciamento dos cDNAs diferenciais	65
2.5. Estratégia de sequenciamento	65
2.6. Análise das seqüências	65
3. Resultados e Discussão	66
3.1. <i>Display</i> diferencial de mRNA	66
3.1.1. Expressão diferencial de genes durante o desenvolvimento do embrião zigótico de <i>A. angustifolia</i>	66
3.1.2. Confirmação da expressão diferencial das bandas isoladas dos géis de <i>display</i> diferencial de mRNA através de <i>slot blot</i>	67
3.1.3. Confirmação da expressão diferencial das bandas isoladas dos géis de <i>display</i> diferencial de mRNA através de <i>macroarrays</i>	69
3.2. Expressão diferencial de genes no megagametófito e em diferentes regiões do embrião maduro	69
3.3. Clonagem e sequenciamento	70
3.4. cDNAs que apresentaram similaridade com seqüências presentes em bancos de dados	75

3.4.1. cDNAs com expressão mais acentuada no embrião maduro	75
3.4.2. cDNAs com expressão mais acentuada no embrião cotiledonar	79
3.4.3. cDNAs com expressão mais acentuada no embrião globular	81
3.4.4. cDNAs com expressão mais acentuada no megagametófito maduro	83
3.4.5. cDNAs com expressão mais acentuada no cotilédone do embrião maduro	84
3.4.6. cDNAs com expressão mais acentuada no eixo do embrião maduro	85
4. Referências Bibliográficas	87
ANEXO 2	91
ANEXO 3	109
CAPÍTULO 3	
Isolamento da sequência completa e caracterização dos cDNAs J6G-36, J5G-26 e J5G-20	117
1. Introdução	118
2. Isolamento da sequência completa e caracterização do cDNA J6G-36 – vicilina 7S de <i>Araucaria angustifolia</i>	118
2.1. Isolamento da sequência completa do cDNA J6G-36 através de RACE 3'	119
2.2. O gene de vicilina 7S de <i>A. angustifolia</i> é semelhante aos descritos para angiospermas e gimnospermas	119
2.3. As proteínas de 20-21 kDa e 28 kDa podem ser produtos da clivagem de um precursor comum	122
2.4. Expressão do gene de vicilina 7S no embrião e megagametófito durante o desenvolvimento e germinação da semente	123
3. Isolamento da sequência completa do gene <i>araCAF1</i> por RACE 5' a partir do cDNA J5G-26	128
4. Isolamento da sequência completa do gene da lipocalina	131
5. Referências Bibliográficas	135
CAPÍTULO 4	
Considerações finais sobre o desenvolvimento embrionário de <i>Araucaria angustifolia</i>	136
1. Desenvolvimento do embrião e depósito de reservas na semente de <i>A. angustifolia</i>	137
2. As principais proteínas de reserva de <i>A. angustifolia</i> são vicilinas 7S	139
3. Expressão gênica no desenvolvimento do embrião zigótico de <i>A. angustifolia</i>	140
4. Referências Bibliográficas	143
CONCLUSÕES	145

RESUMO

A *Araucaria angustifolia*, única espécie do gênero de ocorrência natural no Brasil, era encontrada principalmente nos Estados do Paraná (40% da superfície), Santa Catarina (31%) e Rio Grande do Sul (25%). Manchas esparsas eram observadas no sul de São Paulo (3%), sul de Minas Gerais e Rio de Janeiro, em áreas de altitude elevada (1%). A madeira serrada e laminada da Araucária foi, por um longo período, um dos produtos mais importantes de exportação brasileira. Atualmente, a *A. angustifolia* se encontra na lista das espécies brasileiras ameaçadas de extinção. As matas remanescentes correspondem a apenas 4,3% da área original. Estudos histológicos mostraram que durante o desenvolvimento, o embrião de *A. angustifolia* apresenta os estágios: (1) globular, (2) globular tardio, (3) torpedo, (4) cotiledonar e (5) maduro. O desenvolvimento do embrião de *A. angustifolia* difere em vários aspectos morfológicos do de *Pinus*, considerado como “modelo” das coníferas mais evoluídas. Este fato, torna interessante do ponto de vista acadêmico a realização de estudos moleculares sobre o desenvolvimento do embrião de *A. angustifolia*. No Capítulo 1, “Caracterização das principais proteínas de reserva de *A. angustifolia*”, os resultados obtidos indicaram que as principais proteínas de reserva da semente de *A. angustifolia* são solúveis em solução salina (0,25 M NaCl), não apresentam pontes bisulfídicas e têm peso molecular aparente de 20-21 kDa e 28 kDa. Estas características diferem das já descritas para outras proteínas de reserva de coníferas, e colocam as principais proteínas de reserva de *A. angustifolia* no grupo das vicilinas 7S. No Capítulo 2, “Identificação de genes expressos diferencialmente durante o desenvolvimento do embrião zigótico de *A. angustifolia*”, foi utilizada a técnica de *display* diferencial de mRNA para a identificação de genes específicos dos estágios globular, cotiledonar e maduro do embrião zigótico e no megagametófito da semente madura de *A. angustifolia*. Foram isolados 315 cDNAs com expressão diferencial. Entre os 189 cDNAs confirmados, 39 foram clonados, seqüenciados e as seqüências foram comparadas às depositadas em bancos de dados. No Capítulo 3, “Isolamento da seqüência completa e caracterização dos cDNAs J6G-36, J5G-26 e J5G-20”, três cDNAs que apresentam similaridade com seqüências disponíveis em bancos de dados, vicilina 7S, *araCAF1* e lipocalina, tiveram a respectiva seqüência completa caracterizada.

SUMMARY

Araucaria angustifolia, the only native conifer in Brazil, covered the States of Paraná (40%), Santa Catarina (31%) and Rio Grande do Sul (25%). Sparse patches were also found in areas with elevated altitudes, in the States of São Paulo (3%), Minas Gerais and Rio de Janeiro (1%). The *A. angustifolia* wood was, for a long time, one of the most important exportation products in Brazil. In the present days, the *A. angustifolia* is at risk of extinction, and the remaining areas correspond to 4.3% of the original. The *A. angustifolia* embryo development is divided in the following stages: (1) early globular stage, (2) late globular stage, (3) torpedo stage, (4) cotyledonar and (5) mature stage. The development of the *A. angustifolia* embryo differs in some morphologic aspects from other conifers. Therefore, it is interesting to study embryo development in *A. angustifolia*. In Chapter 1, “Characterization of the major storage proteins in *A. angustifolia*”, the characterization of the major storage proteins of *A. angustifolia* was described. These proteins are soluble in solutions with low-salt concentration (NaCl 0.25 M). They have apparent molecular weights around 20-21 kDa and 28 kDa, and there are no evidences of bisulfide bonds. All these characteristics diverged from those described for storage proteins in other conifers, and indicated that the major storage proteins of *A. angustifolia* are 7S vicilin-like. In Chapter 2, “Identification of differential expressed genes during development of the zygotic embryo in *A. angustifolia*”, the mRNA differential display technique was used to identify genes differentially expressed in the developing embryo, globular, cotyledonar and mature stages, and in the megagametophyte of mature seed. A total of 315 differentially expressed cDNAs were isolated. From these, 189 cDNAs had their differential expression confirmed and 39 were cloned and submitted to sequencing. Nine cDNAs showed similarity with GenBank sequences. In Chapter 3, “Isolation and characterization of the full length sequences of J6G-63, J5G-26 and J5G-20 cDNAs”, the isolation of full-length cDNA sequences of the 7S vicilin-like gene (J6G-63 cDNA), the lipocalin gene (J5G-20 cDNA) and the *araCAF1* transcription factor gene (J5G-26) is described.

INTRODUÇÃO

Um dos processos biológicos mais complexos e interessantes é a embriogênese. Durante este processo, uma célula ou zigoto usa todo o seu potencial genético em divisões celulares complexas e consecutivas que resultam no organismo pluricelular. O crescimento, divisão e diferenciação celular, durante a embriogênese, é aparentemente controlado por mecanismos moleculares evolutivamente conservados tanto em plantas como em animais. O estudo dos genes expressos nos diferentes estágios de desenvolvimento do embrião, assim como os mecanismos de regulação gênica envolvidos no controle da expressão temporal e tecido específico dos genes durante este processo de desenvolvimento é de fundamental importância.

Em plantas, o processo de desenvolvimento e diferenciação do embrião é controlado direta ou indiretamente por mudanças no padrão da expressão de genes (Goldberg *et al.*, 1989; Misra, 1994; Wobus & Weber, 1999; Girke *et al.*, 2000). Assim sendo, estes processos representam um modelo biológico interessante para estudos de organização, diferenciação celular e expressão diferencial de genes (Goldberg *et al.*, 1989; Zimmerman, 1993). Os mecanismos que levam à expressão temporal e tecido específico de genes durante o desenvolvimento da semente, vem sendo intensamente investigados pela comunidade científica nos últimos anos (Goldberg *et al.*, 1989; Zimmerman, 1993; Misra, 1994; Mullen *et al.*, 1996; Cairney *et al.*, 1999; Wobus & Weber, 1999; Girke *et al.*, 2000; White *et al.*, 2000).

Recentemente, progressos importantes foram feitos no que diz respeito à estrutura e modo de ação de genes expressos diferencialmente na semente de angiospermas. Utilizando técnicas avançadas de análise computacional de *microarrays*, foram identificados mais de 1000 genes com expressão diferencial na semente de *Arabidopsis thaliana* e *Brassica napus* (Girke *et al.*, 2000; White *et al.*, 2000), no que pode ser catalogado como o estudo mais completo de genes expressos na semente (<http://bpp.msu.edu/Seed/SeedArray.htm>).

Mecanismos de regulação da expressão gênica, durante o desenvolvimento do embrião de gimnospermas, muito parecidos aos descritos para angiospermas (Goldberg *et al.*, 1989; Zimmerman, 1993; Wobus & Weber, 1999; Girke *et al.*, 2000; White *et al.*, 2000), foram descritos por Mullen *et al.* (1996) e Dong & Dustan (1996, 1997, 1999). No entanto, pouco se sabe sobre o desenvolvimento da semente de coníferas. O estudo do

desenvolvimento deste tipo de semente é limitado pelo longo ciclo reprodutivo e a dificuldade em se obter sementes em desenvolvimento (Chattahai & Misra, 1998).

Embora timidamente, os estudos envolvendo o desenvolvimento e a germinação das sementes em gimnospermas estão aumentando (Newton *et al.*, 1992; Krassowski & Owens, 1993; Misra, 1994; Mullen *et al.*, 1996; Dong & Dustan, 1996, 1997, 1999; Cairney *et al.*, 1999; Sabala *et al.*, 2000). Contudo, os resultados obtidos e a quantidade de pesquisadores envolvidos em projetos nesta área ainda é proporcionalmente insuficiente.

Embora tenham ocorrido avanços notáveis na área de biotecnologia vegetal e biologia molecular de plantas, no Brasil são poucos os investimentos na área de biotecnologia de coníferas e mais especificamente, na obtenção de marcadores moleculares do processo de embriogênese e no desenvolvimento de técnicas de embriogênese somática e micropropagação. As poucas iniciativas registradas partem de grupos ilhados de setores particulares da indústria dificultando assim, o acesso aos dados obtidos nas pesquisas.

Em *Araucaria angustifolia*, a conífera nativa de maior importância econômica, estudos moleculares visando o desenvolvimento da semente ainda são inéditos. Estudos histológicos mostraram que durante o desenvolvimento, o embrião de *A. angustifolia* apresenta os estágios: (1) globular, (2) globular tardio, (3) torpedo, (4) cotiledonar e (5) maduro. O desenvolvimento do embrião de *A. angustifolia* difere em vários aspectos morfológicos do de *Pinus*, considerado como “padrão” das coníferas mais evoluídas. Este fato, torna interessante do ponto de vista acadêmico a realização de estudos moleculares sobre o desenvolvimento do embrião de *A. angustifolia*.

1. *Araucaria angustifolia*

Araucariaceae, do ponto de vista biogeográfico, é uma das famílias mais interessantes entre as gimnospermas. Ela está distribuída principalmente no hemisfério sul, ocorrendo na Nova Caledônia, Nova Guiné, Austrália, Nova Zelândia e América do Sul (Setoguchi *et al.*, 1998). Na América do Sul, as principais espécies são *Araucaria araucana* e *Araucaria angustifolia*.

A *Araucaria angustifolia* (*brasiliensis*) Otto Kuntze, também é conhecida como pinheiro-do-Paraná, pinheiro brasileiro, araucária ou simplesmente pinheiro. Ela é a única

espécie do gênero de ocorrência natural no Brasil (Shimizu & Oliveira, 1981). As árvores são monóicas, com 10 - 35 metros de altura e 50 - 120 centímetros de diâmetro, os troncos são retos e quase cilíndricos (Carvalho, 1994), raramente ramificados em dois, a casca é grossa e resinosa e a ramificação é racemosa apresentado verticilos quase horizontais. Os verticilos se ramificam em abundantes ramos secundários constituindo as grimpas que vão se adensar no ápice do caule, formando uma copa típica em forma de candelabro, umbela ou corimbo nas árvores adultas (Reitz *et al.*, 1978) (Figura 1A).

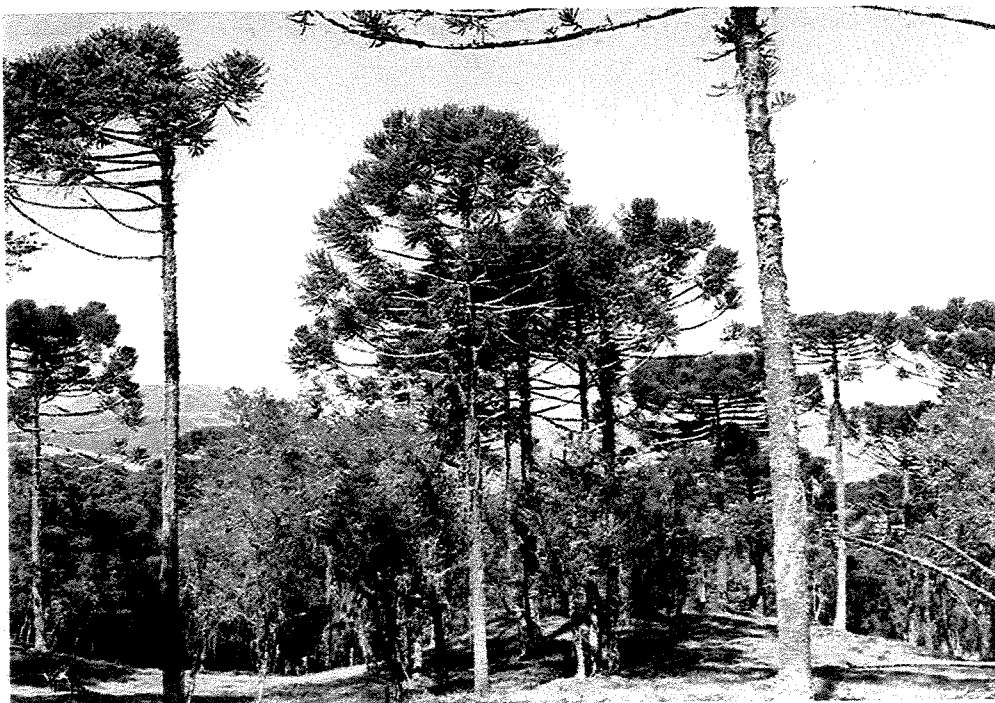
A *Araucaria angustifolia* ocorre naturalmente no Brasil e em pequenas manchas na Argentina (extremo nordeste, na província de Misiones) e no Paraguai (leste, no Departamento de Alto Paraná), ficando localizada entre as latitudes 19°15'S e 31°30'S e entre as longitudes 41°30'W e 54°30'W. No Brasil, a área original, de formato irregular, foi de cerca de 200.000 km². Esta espécie era encontrada principalmente nos Estados do Paraná (40% da superfície), Santa Catarina (31%) e Rio Grande do Sul (25%). Manchas esparsas eram observadas no Sul de São Paulo (3%), internando-se até o sul de Minas Gerais e Rio de Janeiro, em áreas de altitude elevadas (1%) (Carvalho, 1994).

1.1. Importância econômica da *A. angustifolia*

A madeira serrada e laminada da *A. angustifolia* foi, por um longo período, um dos produtos mais importantes na exportação brasileira. Esta madeira, de alta qualidade, era utilizada para construções em geral, trabalhos de caixotaria, fabricação de móveis, laminados, polpa de celulose, palitos de fósforo, lápis, formas para concreto e outros. Ela produz 58,3% de celulose de fibra longa e 28,5% de lignina, podendo assim ser utilizada na produção de papel de alta qualidade (Carvalho, 1994; Shimizu & Oliveira, 1981).

A resina, exsudada principalmente da casca, serve para a fabricação de vernizes, terebentina, acetona, ácido pirolenhoso e outros produtos químicos. Finalmente, o pinhão é muito apreciado na alimentação humana, de animais domésticos e da fauna silvestre que também são responsáveis pela sua dispersão (Figura 1B, C). Entre os animais da fauna silvestre que se alimentam da Araucária, destacam-se a gralha-azul (*Cyanocorax caeruleus*), o serelepe (*Sciurus aestuans*) e os porcos-do-mato (Mattos, 1994).

A:



B:



C:



Figura 1. A: *Araucaria angustifolia* fotografada na serra catarinense no município de Bom Retiro. B: Estróbilo (pinha) imaturo de 10-20 cm de diâmetro. C: Semente madura, mais conhecida como pinhão.

A exploração da *A. angustifolia* se intensificou a partir de 1934, tendo seu auge nas décadas de 50 a 70 com produção estável em torno de 2,8 milhões de metros cúbicos de madeira (Mattos, 1994). No Estado de São Paulo, as reservas foram exauridas entre 1930 e 1940 (Shimizu & Oliveira, 1981) e no Estado do Paraná, em 1977. As matas remanescentes correspondem a apenas 4,3% da área original (Shimizu & Oliveira, 1981).

O incremento em volume sem casca em povoamentos de *A. angustifolia* variou de 2,56 m³/ha ano em solos mais pobres até 14,48 m³/ha ano nos melhores solos. Isto indica um ótimo desempenho da *A. angustifolia*, sempre que cultivada em solos favoráveis. Mesmo assim, o plantio da araucária declinou de maneira drástica, sendo substituída por *Pinus elliottii*, *P. taeda* e *Eucaliptus sp.* que são menos exigentes quanto a fertilidade do solo. Segundo a Sociedade Brasileira de Silvicultura, na geração de energia e produção de celulose em 1999, foram consumidos 68 000 m³ de madeira de araucária. Isto representa menos de 1% do total de madeira consumida.

1.2. A embriogênese zigótica em *A. angustifolia*

A embriogênese em coníferas pode ser dividida em duas fases: (1) fase pró-embriônica: esta fase vai desde a fertilização até o rompimento da arquegônia pelo pró-embrião e (2) fase embrionária: a partir do rompimento da arquegônia até a maturação do embrião (Haines & Prakash, 1980).

O processo de fertilização em *A. angustifolia* não foi totalmente desvendado, mas algumas características importantes foram descritas (Mattos, 1994):

- I. o carpelo, para ser polinizado deve ter mais de um ano;
- II. o período da polinização à fertilização é de 12-14 meses;
- III. no desenvolvimento do pró-embrião há 3 possibilidades:
 - a. o pró-embrião inicia imediatamente o seu desenvolvimento;
 - b. o pró-embrião entra em repouso “hibernal”;
 - c. a arquegônia formado entra em repouso de cerca de 10 meses e, no fim deste período será fertilizado.
- IV. na última fase (relativamente rápida), o pró-embrião completa o seu desenvolvimento de novembro a maio.

(1) Fase pró-embrionária:

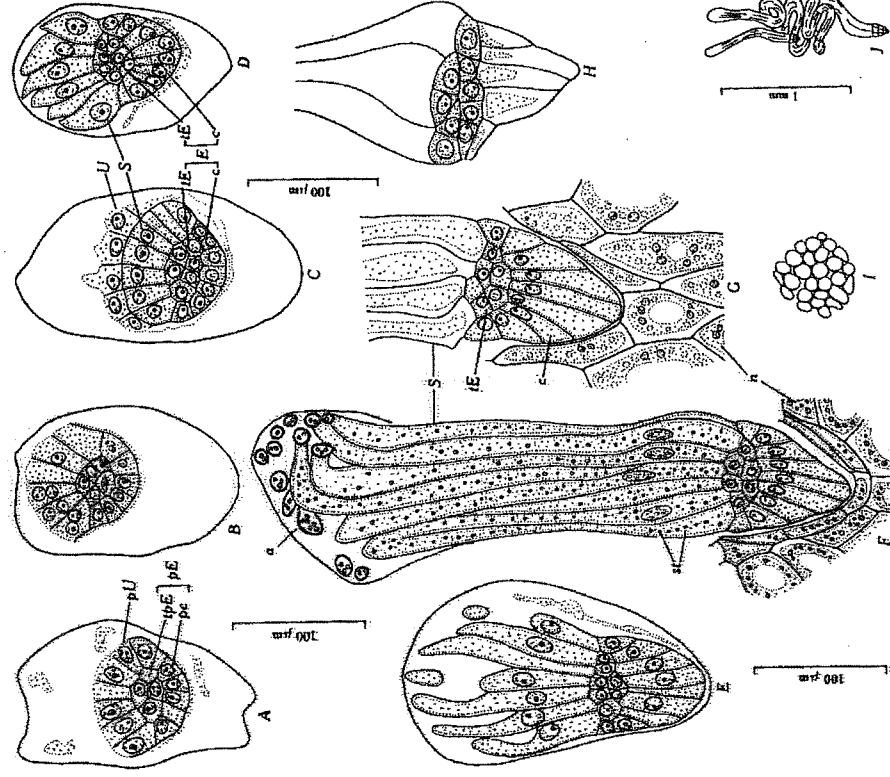
Em *Araucaria angustifolia*, após a fecundação, na arquegônia são formados entre 32 e 45 núcleos antes da formação das paredes celulares, consistindo este período na “fase de núcleo livre” (Mattos, 1994). Isto resulta na formação do pró-embrião primário que é formado por duas camadas celulares distintas: (i) o *primary embryonal tier* - **pE** que possui de 26 a 33 células (Haines & Prakash, 1980) e (ii) o *primary upper tier* - **pU** que possui de 31 a 38 células (Figura 2A-A) (Haines & Prakash, 1980). Comparativamente, as células que constituem o **pE** são menores, mais compactas e fisiologicamente mais ativas que as células que constituem o **pU**. Dois tipos de tecidos diferentes podem ser distinguidos no **pE**: (i) *primary cap* - **pC**, composto por 16-22 células e (ii) *true embryonal group* - **tpE**, composto por 9-13 células (Figura 2A-A) (Haines & Prakash, 1980). As células do **pC** arranjam-se ao redor do **tpE**. No decorrer do desenvolvimento, as células do **pC** dividem-se no sentido anticlinal formando a capa (**C**) que é constituída por uma única camada de células com função de proteção.

Enquanto as células do **tpE** sofrem divisões não sincronizadas em todos os planos (**tE**), as células do **pU** dividem-se e sofrem alongamentos dando origem ao suspensor primário (**S**) (Figura 2A-B a F). As células da região distal do suspensor sofrem espessamentos que possibilitam a ancoragem e a movimentação do pró-embrião através da zona de corrosão do megagametófito (Figura 2A-G).

(2) Fase embrionária:

Quando as células do pró-embrião chegam à zona central do megagametófito, as células do **tE** entram em um processo ativo de divisões mitóticas que resultam no crescimento acelerado da cabeça embrionária do embrião zigótico (Grupo embrionário - **tE**). Nesta etapa, a capa é quebrada e as células são desintegradas e absorvidas (Figura 2A-H, I). As células embrionárias sofrem um alongamento dando origem ao suspensor secundário. Com a formação do suspensor secundário, 3 tipos de células passam a ser reconhecidas no embrião: as células do ápice embrionário, as células intermediárias e as células do suspensor secundário. Quanto às células do suspensor primário, elas continuam conectadas ao suspensor secundário e a parte basal do megagametófito servindo como ponte direta para passagem de metabólitos para o embrião zigótico em desenvolvimento (Figura 2A-J).

A:



B:

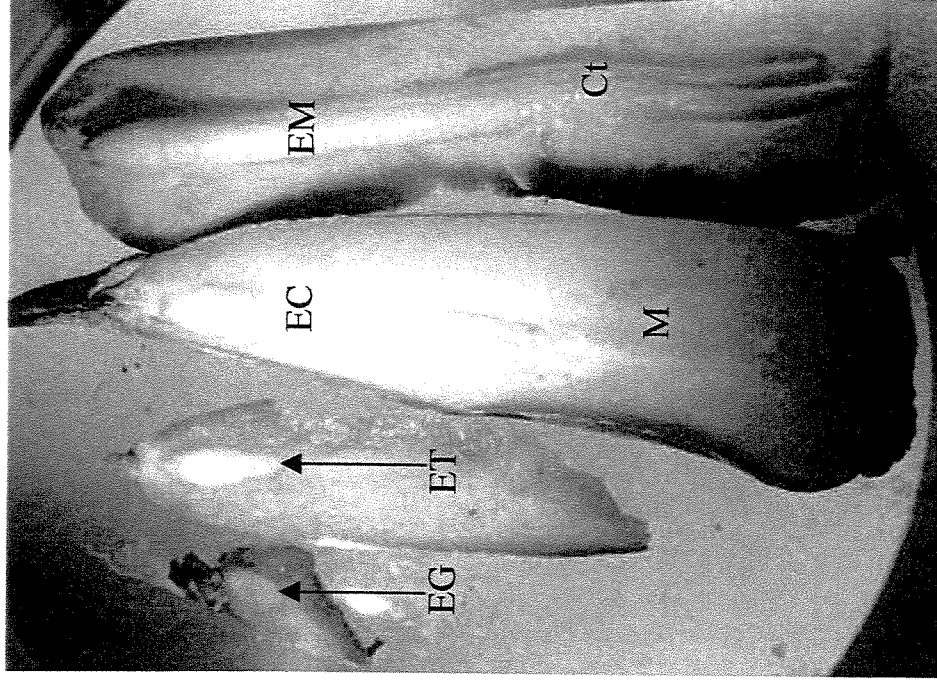


Figura 2. Representação esquemática do desenvolvimento do embrião zigótico de *A. angustifolia*. (A, B) pE- *primary embryonal tier*, pU - *primary upper tier*; pC - *primary cap*; tpE - *true embryonal group*; (C - G) tE - *grupo embrionário*; S- *suspensor primário*; C - *capa*, n - *células polinucleadas do megagametófito*. B: Fase embrionária. EG - *embrião globular*; ET - *embrião torpedado*, EC - *embrião cotiledonar*; EM - *embrião maduro*; M - *megagametófito*; Ct - *cotilédone*. Adaptado de Haines & Prakash (1980).

O embrião no decorrer do desenvolvimento, forma um grupo globular compacto que resulta na formação do embrião globular (EG) (Figura 2A-I, J). O embrião globular aumenta aproximadamente 2 vezes dando origem ao embrião globular tardio (Figura 2B). A parte apical do embrião globular tardio sofre um alongamento, resultando na formação do estágio de desenvolvimento embrionário conhecido como torpedo (ET) (Figura 2B). A seguir, com o alongamento posterior da região apical, são formados os cotilédones (Ct) dando origem à fase cotiledonar de desenvolvimento (EC) (Figura 2B). A seguir, o embrião aumenta de tamanho ocupando todo o eixo central do megagametófito (M); o ápice radicular se forma a partir da diferenciação de células próximas ao suspensor secundário e o ápice caulinar é formado a partir do ápice embrionário dando origem ao embrião maduro (EM) (Figura 2B).

Estudos morfológicos da fase pró-embrionária do desenvolvimento embrionário na *Araucaria* sugeriram que esta é uma das espécies mais primitivas de coníferas (Haines & Prakash, 1980). Estes autores assinalam (i) a fase de núcleo livre prolongada, (ii) a presença das células da capa e (iii) a ausência de clivagem do pró-embrião, como as principais diferenças encontradas no desenvolvimento do pró-embrião em *Araucaria*, se comparado com outras coníferas consideradas evolutivamente superiores (Cupressaceae, Pinaceae e Taxodiaceae) (Haines & Prakash, 1980).

2. Síntese e deposição de proteínas durante a embriogênese em coníferas

O desenvolvimento da semente como órgão especializado na proteção, dispersão e nutrição do embrião foi um passo importante na evolução das plantas terrestres (Shutov *et al.*, 1998). O acúmulo de compostos como lipídios, proteínas e amido na semente é de vital importância para suprir as necessidades do embrião durante a germinação e crescimento inicial da planta. Além de serem modelos interessantes para estudos biológicos, as sementes possuem também importância econômica. Estes fatores somados, vem motivando estudos sobre os processos de acumulação e utilização de reservas na semente, principalmente no que se refere às proteínas.

Segundo Bewley e Black (1990) as proteínas podem ser classificadas de acordo com a sua função nas seguintes categorias: (1) Proteínas estruturais - podem estar associadas a

complexos lipídicos de membranas extra e intra celulares e estão presentes no citoesqueleto, histonas e ribossomos; (2) Enzimas – estão envolvidas no metabolismo celular; (3) Proteínas de reserva - abundantes nas sementes e utilizadas como fonte de nutrientes durante a germinação e crescimento inicial da planta.

2.1. Classificação das proteínas de reserva

Osborne (1924) classificou as proteínas da semente com base na sua solubilidade em diferentes solventes em: (1) albuminas - proteínas solúveis em água, (2) globulinas – proteínas solúveis em soluções salinas, (3) prolaminas – solúveis em soluções álcool/água e (4) glutelinas – solúveis em soluções ácidas ou básicas diluídas. De acordo com Shewry *et al.* (1995) as principais proteínas de reserva das sementes são albuminas, globulinas ou prolaminas.

As proteínas de reserva também podem ser separadas com base no coeficiente de sedimentação ($S_{20,w}$). Os grupos mais conhecidos são as albuminas 2S, encontradas nas sementes das dicotiledôneas e evolutivamente relacionadas às prolaminas (presentes nas sementes das gramíneas), e as globulinas 7S tipo-vicilinas e 11S tipo-leguminas encontradas tanto em monocotiledôneas quanto em dicotiledôneas (Shewry & Casey, 1999).

2.2. Proteínas de reserva das coníferas

Nas sementes de coníferas, o embrião se encontra na cavidade de corrosão do megagametófito que é um tecido haplóide de origem materna (Hakman, 1993). No megagametófito são depositados a maior parte dos compostos de reserva, que são constituídos principalmente por proteínas e lipídios (Owens *et al.*, 1993). As reservas representam aproximadamente de 12 a 20% do peso fresco da semente, no caso de *Pinus taeda* L. (Groome *et al.*, 1991). As proteínas de reserva das coníferas são depositadas em corpos protéicos localizados no citoplasma das células do megagametófito e do embrião. Estudos histoquímicos mostraram que os corpos protéicos são formados por inclusões globóides e cristalóides em uma matriz amorfa (Green *et al.*, 1991; Krassowski & Owens,

1993; Hakman, 1993). O tamanho e o número dos corpos protéicos é aparentemente maior no parênquima do megagametófito do que nas células dos cotilédones do embrião, sugerindo que o megagametófito é o principal órgão de reserva de proteínas (Misra & Green, 1990; Misra, 1994).

As proteínas que formam os corpos protéicos foram divididas em proteínas solúveis (proteínas da matriz ou globóides) e proteínas insolúveis (cristalóides) (Gifford, 1988). Diversos estudos realizados com proteínas de reserva de coníferas mostraram que as cristalóides são as principais representantes desta classe de proteínas. Em sementes maduras de *Picea glauca*, *Picea abies*, *Pinus taeda* e *Pseudotsuga menziesii*, as cristalóides representam de 70 a 80% do total das proteínas de reserva. As cristalóides podem ser isoladas do embrião e megagametófito com soluções contendo SDS (2-4%) (Gifford, 1988) ou com soluções salinas concentradas (1 M) (Misra & Green, 1991; Hakman, 1993; Misra, 1994).

Misra & Green (1990) determinaram que as cristalóides no gênero *Picea* são constituídas principalmente por um grupo de proteínas com peso molecular aparente entre 55-57 kDa. Quando submetida à eletroforese em condições denaturantes, estas proteínas são subdivididas em 2 subunidades com peso molecular aparente ao redor de 34,5-35 kDa e 22,5-23 kDa (Misra & Green, 1990). As cristalóides começam a ser sintetizadas logo no início do desenvolvimento do embrião de *Picea glauca* e atingem um alto nível de acumulação durante a fase de maturação (Misra & Green, 1991).

Groome *et al.* (1991) observaram que em *Pinus taeda*, uma proteína de 60 kDa era a principal representante das cristalóides. Na presença de β -mercaptoetanol, esta proteína era subdividida em 2 subunidades de 37,5 e 22,5 kDa. Estudos baseados em experimentos com antígenos mostraram a existência de similaridade entre as cristalóides e as leguminas 11S das angiospermas (Misra & Green, 1994), ressaltando a relação evolutiva entre estas proteínas. Estes estudos foram reforçados por dados anteriormente obtidos através da análise das seqüências dos genes que codificam para estas proteínas (Newton *et al.*, 1992; Leal & Misra, 1993b).

3. Expressão gênica durante o desenvolvimento do embrião zigótico de coníferas

A caracterização da expressão gênica durante a embriogênese e o desenvolvimento e germinação da semente tem sido amplamente abordada, principalmente no que diz respeito às angiospermas (Goldberg *et al.*, 1989; Skriver & Mundy, 1990; Delseny *et al.*, 1993; Wobus & Weber, 1999; Girke *et al.*, 2000; White *et al.*, 2000). Estes estudos dividiram os genes em sete grupos principais: (1) genes constitutivos (expressos em todos os estágios de desenvolvimento), (2) genes com expressão específica no embrião, (3) genes expressos durante as etapas iniciais da embriogênese, (4) genes das proteínas de reserva, (5) genes expressos exclusivamente nas etapas finais da embriogênese (genes *lea*), (6) genes envolvidos nas etapas de maturação e início da germinação da semente (quebra de dormência) e (7) genes expressos durante a germinação (Goldberg *et al.*, 1989). Entre estes genes, os mais estudados e caracterizados são os correspondentes às proteínas de reserva, os genes *lea* (*late embryogenesis abundant*) e os genes expressos na quebra da dormência e durante a germinação da semente.

Ficou constatado que existem similaridades entre as populações de mRNA com padrão de expressão diferencial durante a germinação da semente em angiospermas e gimnospermas (Mullen *et al.*, 1996). Estes autores constataram o aumento em 7 vezes da quantidade de RNA no embrião de *Pinus taeda* durante a germinação. A mudança dos padrões de expressão das “populações” de mRNA na maturação e germinação das sementes de *Pinus taeda* foram estudadas neste trabalho. O padrão de expressão destas populações de mRNA no embrião e megagametófito foi muito similar ao descrito para angiospermas (Goldberg *et al.*, 1989, Wobus & Weber, 1999; Girke *et al.*, 2000). Os resultados obtidos por Mullen *et al.* (1996) sugerem que a regulação da expressão das diferentes populações de mRNA, durante o desenvolvimento e germinação da semente, é conservada do ponto de vista evolutivo.

Em coníferas, foram obtidos cDNAs das principais proteínas de reserva de *Picea glauca/engelmannii* (Newton *et al.*, 1992) e *Pseudotsuga menziesii* (Leal & Misra, 1993b). Clones de cDNA dos genes *lea* (*late embryogenesis abundant*), expressos nas etapas iniciais da germinação, foram isolados por Jarvis *et al.* (1996) em *Pseudotsuga menziesii*. Mudanças no padrão de expressão dos genes *lea* durante o desenvolvimento e germinação

das sementes foram detectadas em *Pinus taeda* (Mullen *et al.*, 1996; Cairney *et al.*, 1999), entre outros.

Um dos cDNAs de proteína de reserva de *Picea glauca/engelmannii* analisado representava uma proteína de 50 kDa, que apresentava 28-38% de identidade com as vicilinas 7S de angiospermas (Newton *et al.*, 1992). Experimentos de reconstrução do genoma mostraram que o cDNA isolado por Newton *et al.* (1992) era parte de uma família multigênica. Leal & Misra (1993b) em trabalhos com *Pseudotsuga menziesii* isolaram um cDNA (DF1) que apresentou identidade (29-38,5%) com as globulinas 11-12S das angiospermas. A hibridização do cDNA com membranas contendo RNA isolado do megagametófito em 2, 15, 29, 43, 57 e 71 DAF (dias após a fertilização) e do embrião em 43, 57 e 71 DAF mostrou que a expressão era mais abundante nos estágios iniciais e intermediários de desenvolvimento (Leal e Misra, 1993b).

Resultados parecidos foram obtidos com *Picea glauca/engelmannii* (Newton *et al.*, 1992), *Pinus pinaster* (Allona *et al.*, 1994) e *Pinus taeda* (King & Gifford, 1997), aonde foram isolados polipeptídios com 30-36% de identidade com as globulinas 7S das angiospermas. Foram isolados também, em *Pseudotsuga menziesii*, cDNAs contendo isoformas de proteínas de reserva similares as albuminas 2S (Chatthai & Misra, 1998).

Uma classe importante de genes expressos durante as fases finais do desenvolvimento da semente e início da germinação é a dos genes *lea* (*late embryogenesis abundant*). Jarvis *et al.* (1996) isolaram os primeiros genes *lea* de gimnospermas trabalhando com sementes maduras de *Pseudotsuga menziesii*. As 3 classes de genes isolados por Jarvis *et al.* (1996) DF65, DF6 e DF77 apresentaram similaridade com as classes 2, 1 e 3 de genes *lea* descritos em angiospermas, respectivamente.

Trabalhos envolvendo embriogênese somática em coníferas relataram genes com expressão diferencial no processo de crescimento e maturação, genes que respondiam ao choque térmico, genes induzidos por ABA e genes com expressão específica no embrião somático (Dong & Dustan, 1996, 1997, 1999; Filonova *et al.*, 2000). Em *Picea abies* (Pinheiro do Canadá), foi detectada uma proteína de transferência de lipídeos com expressão constitutiva no embrião somático. Esta proteína apresentou alta similaridade com proteínas descritas para angiospermas (Sabala *et al.*, 2000). Larsson *et al.* (1998) isolaram de uma biblioteca de cDNA de embrião de *Picea abies* o primeiro cDNA correspondente a

um gene *homeobox*, *HBK1*, em conífera. A sequência deste gene apresentou similaridade com os genes *homeobox KNOX* (*Knotted-like Homeobox*) encontrados em diferentes angiospermas. Os autores sugeriram que a função do gene *HBK1* é controlar a diferenciação celular no meristema apical (Larsson *et al.*, 1998).

Tranbarger & Misra (1996) isolaram um cDNA cuja expressão era abundante no megagametófito após a germinação de sementes de *Pseudotsuga menziesii*. Este cDNA codificava para uma protease de cisteína. Foi isolado também de *Pseudotsuga menziesii* um gene de metalotioneína (MT) cuja expressão ocorria principalmente no megagametófito durante a embriogênese (Chatthai *et al.*, 1997). A expressão deste cDNA era modulada por ABA (Chatthai *et al.*, 1997).

Na família *Araucariaceae*, estudos sobre a expressão de genes durante o desenvolvimento e germinação da semente ainda são inéditos.

4. Referências Bibliográficas

- Allona I., Casado R., Aragoncillo C. (1994) Biochemical genetics of a 7S globulin-like protein from *Pinus pinaster* seed. Theor. Appl. Genet. 88: 454.
- Bewley J. D. & Black M. (1990) Seed development and maturation. in Seed: Physiology of Development and Germination. New York, Plenum Press. p. 29.
- Cairney J., Nanfei X., Pullman G. S., Vincent T., Ciavatta V. T., Johns B. (1999) Natural and somatic embryo development in loblolly pine. Applied Biochem. and Biotech. 79: 5.
- Carvalho P.E.R. (1994) Espécies florestais brasileiras, recomendações silviculturais, potencialidades e usos da madeira. Embrapa-CWPF.Colombo PR. p. 70.
- Chatthai M., Kaukinen K. H., Tranbarger T. J., Grupta P. K., Misra S. (1997) The isolation of a novel metallothionein-related cDNA expressed in somatic and zygotic embryos of Douglas-fir: regulation by ABA, osmoticum and metal ion. Plant Mol. Biol. 34 (2): 243.
- Chatthai M., Misra S. (1998) Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziessi* [Mirb] Franco. Planta 206: 138.
- Delseny M. P., Gaubier G., Hull G., Saezunsquez J., Gallois P., Raynal M., Cooke R., Grellet F. (1993) Nuclear genes expressed during seed desiccation: relationship with responses to stress. In: A.A.Barsa (ed.). Stress induced gene expression. Harwood Academic Publishers.
- Dong J.Z., Dunstan D.I. (1996) Characterization of three heat-shock proteins genes and their development regulation during somatic embryogenesis in white spruce [*Picea glauca* (Moench) Voss]. Planta 200: 85.
- Dong J.Z., Dunstan D.I. (1997) Characterizations of cDNA representing five abscisic acid-responsive genes associated whit somatic embryogenesis in *Picea glauca*, and their responses to abscisic acid stereostructure. Planta 203: 448.
- Dong J. Z., Dunstan D. I. (1999) Cloning and characterization of six embryogenesis-associated cDNAs from somatic embryos of *Picea glauca* and their comparative expression during zygotic embryogenesis. Plant Mol. Biol. 39 (4): 859.

- Filonova L., Bozkov P., von Arnold S. (2000) Developmental pathway of somatic embryogenesis in *Picea abies* as revealed by time-lapse tracking. *J. Exp. Botany* 51 (343): 249.
- Gifford D.J. (1988) An electrophoretic analysis of the seed proteins from *Pinus monticola* and eight other species of pine. *Can. J. Bot.* 66: 1808.
- Girke T., Todd J., Ruuska S., White J., Benning C., Ohlrogge J. (2000) Microarray analysis of developing *Arabidopsis* seeds. *Plant Physiol.* 124: 1570.
- Goldberg R.B., Barker S.J., Perez-Grau L. (1989) Regulation of gene expression during plant embryogenesis. *Cell* 56: 149.
- Green M. J., McLeod J. K., Misra S. (1991) Characterization of Douglas fir protein body composition by SDS-PAGE and electron microscopy. *Plant Physiol. Biochem.* 29: 49.
- Groome M. C., Axler S. R., Gifford D. J. (1991) Hydrolysis of lipid and protein reserves in loblolly pine seeds in relation to protein electrophoretic patterns following imbibition. *Physiol. Plant* 83: 99.
- Haines R. J., Prakash N. (1980) Proembryo development and suspensor elongation in *Araucaria Juss.* *Aust. J. Bot.* 28: 511.
- 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.
- Jain S., Gupta P., Newton R. (1995) Somatic embryogenesis in wood plants. Kluwer Academic Publishers.
- Jarvis S. B., Taylor M. A., MacLeod M.R., Davies H. V. (1996) Cloning and characterization of the cDNA clones of three genes that are differentially expressed during dormancy breakage in the seeds of Douglas fir (*Pseudotsuga menziesii*). *J. Plant Physiol.* 147: 559.
- King J. E., Gifford D. J. (1997) Amino acid utilization in seed of Loblolly Pine during germination and early seedling growth. *Plant Physiol.* 113: 1125.
- Krassowski M. J., Owens J. N. (1993) The ultrastructural and histochemical post-fertilization megagametophyte and zygotic embryo development of interior spruce. *Can. J. Bot.* 71: 98.

- Larsson A. S., Svenson M., Liao H., Engstrom P. (1998) A homeobox gene with potential developmental control function in the meristem of the conifer *Picea abies*. *Plant Biol.* 95: 15118.
- Leal I. & Misra S. (1993) Molecular cloning and characterization of a legumin-like storage protein cDNA of Douglas fir seeds. *Plant Mol. Biol.* 21: 709.
- Mattos J. R. (1994) O pinheiro brasileiro. V.1. 2 ed. Artes Gráficas Princesa LTDA. p. 225.
- Misra S. (1994) Conifer zygotic embryogenesis, somatic embryogenesis and seed germination: Biochemical and molecular advances. *Seed Science Res.* 4: 357.
- Misra S., Green M. J. (1990) Developmental gene expression in conifer embryogenesis and germination. I. Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* [Moench] Voss. *Plant Sci.* 68: 163.
- Misra S., Green M.J. (1991) Developmental gene expression in conifer embryogenesis and germination. II. Crystalloid protein synthesis in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss. *Plant Sci.* 78: 61.
- Misra S., Green, M. J. (1994) Legumin-like storage polypeptide of conifer seed and their antigenic cross-reactivity with 11S globulins from angiosperms. *J. Exp. Botany* 45: 269.
- Mullen R. T., King J. E., Gifford D. J. (1996) Changes in mRNA populations during loblolly pine (*Pinus taeda*) seed stratification, germination and post-germinative growth. *Physiol. Plant.* 97: 545.
- Neutelings G., Domon J. M., Membré N., Bernier F., Meyer Y., Davia A., David H. (1998) Characterization of a germin-like protein gene expressed in somatic and zygotic embryos of pine (*Pinus caribaea* Morelet). *Plant Mol. Biol.* 38: 1190.
- Newton C. H., Flinn B. S., Sutton B. C. S. (1992) Vicilin-like seed storage proteins in the gymnosperm interior spruce (*Picea glauca/engelmannii*). *Plant Mol. Biol.* 20: 315.
- Osborne T. B. (1924) The vegetable proteins, 2nd ed. Longmans, Green and Co., London. p.154.
- Owens J. N., Morris S. J., Misra S. (1993) The ultrastructural, histochemical, and biochemical development of the post-fertilization megagametophyte and the zygotic embryo of *Pseudotsuga menziesii*. *Can. J. For. Res.* 23: 816.
- Reitz R., Klein R. M., Reis A. (1978) Projeto Madeira de Santa Catarina. Separata 28 de sellowia–Herbário Barbosa Rodrigues, Itajai.

- Sabala I., Elfstrand M., Farbos I., Clapham D., Arnold S.V. (2000) Tissue – specific expression of Pa18, a putative lipid transfer protein gene, during embryo development in Norway spruce (*Picea abies*). *Plant Mol. Biol.* 42: 461.
- Setoguchi H., Osawa T. A., Pintaud J., Jaffré T., Veillon J. (1998) Phylogenetic relationships within Araucariaceae based on *rbcL* gene sequences. *Am. J. of Botany* 85 (11): 1507.
- Shewry R., Casey R. (1999) Seed proteins. Kluwer Academic Publ.
- Shewry P. R., Napier J. A., Tatham A. S. (1995) Seed storage proteins: Structures and biosynthesis. *The Plant Cell* 7: 945.
- Shimizu J. Y., Oliveira Y. M. M. (1981) Distribuição da variação e usos dos recursos genética de araucária no Sul do Brasil. Embrapa-URPFCS, Documento 04, Curitiba-PR, p. 9.
- Shutov A. D., Braun H., Chesnokov Y. V., Horstmann C., Kakhovskaya I. A., Baumlein H. (1998) Sequence peculiarity of Gnetalean Legumin-like seed storage proteins. *J. Mol. Evol.* 47: 486.
- Skriver K., Mundy J. (1990) Gene expression in response to abscisic acid and osmotic stress. *Plant Cell* 2: 503.
- Tranbarger T. J., Misra S. (1996) Structure and expression of a developmentally regulated cDNA encoding a cysteine protease (pseudotzain) from Douglas fir. *Gene* 172: 221.
- White J. A., Newman T., Girke T., Focus N., Martinez de Ilarduya O., Jaworski J. G., Ohlrogge J., Benning C. (2000) A new set of *Arabidopsis* ESTs from developing seeds: The metabolic pathway from carbohydrates to seed oil. *Plant Physiol.* 124: 1582.
- Wobus U., Weber H. (1999) Seed maturation: Genetic programmes and control signals *Curr. Opin. in Plant Biology* 2: 33.
- Zimmerman J. (1993) Somatic embryogenesis: A model for early development in higher plants. *The Plant Cell* 5: 1411.

OBJETIVOS

1. Objetivo geral:

O presente trabalho tem como objetivo a identificação e caracterização de proteínas e genes expressos diferencialmente durante o desenvolvimento do embrião zigótico de *Araucaria angustifolia*.

2. Objetivos específicos:

- Separação e caracterização das principais proteínas de reserva do embrião e megagametófito de *A. angustifolia*;
- Isolamento de cDNAs expressos diferencialmente nos diferentes estágios de desenvolvimento do embrião (globular, cotiledonar e maduro);
- Confirmação da expressão diferencial dos cDNAs isolados;
- Amplificação, clonagem e seqüenciamento dos cDNAs com expressão diferencial confirmada;
- Comparação das seqüências obtidas com seqüências disponíveis em bancos de dados;
- Isolamento por RACE da seqüência completa de cDNAs de interesse;
- Seqüenciamento e caracterização dos cDNAs completos isolados por RACE.

CAPÍTULO 1

Caracterização das principais proteínas de reserva da semente de *Araucaria angustifolia*

1. Introdução

As proteínas são os biopolímeros mais versáteis que encontramos na natureza. Elas desempenham funções estruturais, enzimáticas, reguladoras e de reserva dentro da célula. Dada a sua importância econômica e abundância nas sementes, as proteínas de reserva de plantas sempre foram alvo de pesquisas. Embora a maioria das pesquisas sejam concentradas nas proteínas de reserva de angiospermas, nos últimos anos as proteínas de reserva de coníferas tem despertado o interesse da comunidade científica (Green *et al.*, 1991; Krassowski & Owens, 1993; Hakman, 1993; Misra, 1994; King & Gifford, 1997).

As proteínas de reserva das coníferas são depositadas em corpos protéicos localizados no citoplasma das células do megagametófito e do embrião. Estudos histoquímicos mostraram que os corpos protéicos são formados por inclusões globóides e cristalóides em uma matriz amorfa (Green *et al.*, 1991; Krassowski & Owens, 1993; Hakman, 1993). O tamanho e o número dos corpos protéicos é aparentemente maior no parênquima do megagametófito do que nas células dos cotilédones do embrião, sugerindo, que o megagametófito é o principal órgão de reserva de proteínas (Misra & Green, 1990).

As proteínas dos corpos protéicos foram divididas em proteínas solúveis (proteínas da matriz ou globóides) e proteínas insolúveis (cristalóides) (Gifford, 1988). Diversos estudos realizados com proteínas de reserva de coníferas mostraram que as cristalóides são as principais constituintes desta classe de proteínas. Em sementes maduras de *Picea glauca*, *Picea abies*, *Pinus taeda* e *Pseudotsuga menziesii*, as cristalóides representam de 70 a 80% do total das proteínas de reserva (Misra & Green, 1991; Hakman, 1993; Misra, 1994).

Misra & Green (1990) determinaram que as cristalóides no gênero *Picea* são constituídas majoritariamente por um grupo de proteínas com peso molecular aparente entre 55-57 kDa. Quando submetida à eletroforese em condições denaturantes, estas proteínas são subdivididas em 2 subunidades com peso molecular aparente ao redor de 34,5-35 kDa e 22,5-23 kDa (Misra & Green, 1990). As cristalóides começam a ser sintetizadas logo no início do desenvolvimento do embrião e atingem um alto nível de acumulação durante a maturação (Misra & Green, 1991). Estudos baseados em experimentos com antígenos mostraram homologia entre estas cristalóides e o grupo de

leguminas 11S das angiospermas (Misra & Green, 1994), ressaltando a relação evolutiva entre estas proteínas.

Resultados semelhantes foram obtidos em *Picea glauca/engelmannii* (Newton *et al.*, 1992), *Pinus pinaster* (Allona *et al.*, 1994) e *Pinus taeda* (King & Gifford, 1997), onde cDNAs isolados apresentaram 30-36% de identidade com as globulinas 7S e 11S de angiospermas. Foram isolados também, em *Pseudotsuga menziesii*, cDNAs contendo isoformas de proteínas de reserva similares às albuminas 2S (Chatthai & Misra, 1998).

A descrição dos mecanismos moleculares de evolução de globulinas em gimnospermas e angiospermas foi motivo de interesse de vários pesquisadores nos últimos anos (Hager *et al.*, 1995; Shutov *et al.*, 1995, 1998; Braun *et al.*, 1995). Após encontrarem similaridade entre as seqüências nucleotídicas e de aminoácidos Shutov *et al.* (1995), propuseram a equivalência estrutural dos domínios N'- e C'- terminal de vicilinas 7S e leguminas 11S de angiospermas e gimnospermas. Com base nestes resultados foi proposto um ancestral comum para as vicilinas 7S e leguminas 11S de angiospermas e gimnospermas (Bäumlein *et al.*, 1995; Shutov *et al.*, 1998). Segundo Bäumlein *et al.* (1995) e Shutov *et al.* (1998), as principais globulinas de reserva das sementes de angiospermas e gimnospermas tem como ancestral proteínas monoméricas de procariotos envolvidas na desidratação de esporos (esferulinas) (Bäumlein *et al.*, 1995). Esta evolução molecular teria ocorrido na forma de duplicação gênica nas gimnospermas menos evoluídas (Shutov *et al.*, 1998).

No presente capítulo, foram caracterizadas as principais proteínas de reserva da semente de *Araucaria angustifolia*. A caracterização das proteínas foi feita através de SDS-PAGE e 2D-PAGE. As proteínas mais abundantes foram transferidas para membranas de PVDF e tiveram o teminal amino seqüenciado. As seqüências obtidas foram comparadas em bancos de dados com o auxílio do algoritmo BLASTP.

2. Materiais e Métodos

2.1. Coleta do material biológico

Foram coletados cones femininos de 3 plantas localizadas em povoamentos naturais diferentes, na localidade de Bom Retiro, Santa Catarina, no período de novembro (1997) a março (1998), em intervalos de aproximadamente 15 dias. Nestas populações, as araucárias apresentam polinização cruzada. Isto possibilita a variabilidade genética entre as sementes de uma mesma pinha. As sementes (pinhões) foram transportadas para UNICAMP e armazenadas a -70°C .

Um total de 20 pinhões de cada cone coletado foi medido, a casca foi retirada e o peso fresco da semente foi determinado. As sementes foram a seguir colocadas em uma estufa a 80°C por 24 h e o peso seco da semente foi determinado. O mesmo procedimento foi realizado com 10 embriões.

2.2. Extração das proteínas de reserva

As proteínas de reserva de *A. angustifolia* foram isoladas pelo método de King & Gifford (1997) que utiliza SDS no tampão de extração e pelo método de Hakman (1993) que separa as diferentes classes de proteínas com base na solubilidade em NaCl. O método de King & Gifford (1997) foi utilizado na extração de proteínas do embrião e megagametófito nos diferentes estágios de desenvolvimento embrionário (globular, cotiledonar e maduro). O método de Hakman (1993) foi utilizado na extração de proteínas do embrião e megagametófito maduros.

Na extração das proteínas pelo método de King & Gifford (1997), 150 mg de tecido foram homogeneizados, em tubo eppendorf, em tampão Laemmli (Tris-HCl 65 mM, pH 6,8; SDS 2%; glicerol 10% e PMSF 1 mM) na proporção de 1/10 (g/ml). O homogeneizado foi centrifugado por 20 min a 4°C em microcentrífuga (10.000 rpm). O precipitado obtido após a centrifugação foi extraído novamente com tampão Laemmli e os sobrenadantes foram misturados. A seguir, foram realizadas duas extrações consecutivas com 1 volume de

fenol. As proteínas foram precipitadas da fase fenólica através da adição de 5 volumes de acetato de amônio 0,1 M em metanol e incubação por 2 h a -20°C. Após centrifugação por 15 min a 10.000 rpm, o precipitado foi lavado por 2 vezes com acetato de amônio 0,1 M em metanol e uma vez com acetona 80%. O precipitado (fração **SDS**) foi mantido a temperatura ambiente até secar, sendo então ressuspensionado em tampão Laemmli. As amostras foram mantidas a -20°C.

Na extração de proteínas pelo método de Hakman (1993), o tecido vegetal foi homogeneizado em Tris-HCl 65 mM, pH 8,5 e PMSF 1 mM na proporção de 1/10 (g/ml). O homogeneizado foi centrifugado por 10 min a 4°C em microcentrífuga (10 000 rpm). O precipitado obtido foi extraído com Tris-HCl 65 mM, pH 8,5; PMSF 1 mM; NaCl 1 M. Os sobrenadantes foram misturados, resfriados em gelo, diluídos 1:1 com água destilada e incubados por 1 h a 4°C. Após centrifugação por 20 min, 4°C, em microcentrífuga (10 000 rpm), o precipitado (fração **G1** que representa as cristalóides) foi ressuspensionado em 50-100 µl de NaCl 1 M e estocado a -20°C. O sobrenadante foi dialisado em 800-1000 volumes de água por 24-48 h, 4°C. As proteínas que precipitaram, após a diálise em água, foram coletadas por centrifugação (fração **G2** que representa as globóides) e ressuspensionadas em NaCl 1 M. As proteínas presentes no sobrenadante após a centrifugação (fração **Alb** que representa as albuminas) foram precipitadas com 7 volumes de acetona por 1-2 h, -20°C. Após centrifugação, o precipitado foi diluído em água e estocado a -20°C. As proteínas foram quantificadas pelo método de Bradford (1976).

2.3. Extração de proteína total

Foi extraída proteína total do embrião e megagametófito de sementes contendo embriões em diferentes estágios de desenvolvimento (globular, cotiledonar e maduro). A metodologia utilizada na extração foi a descrita por De Mot & Vanderleyden (1989). Aproximadamente 100 mg de tecido foi macerado em nitrogênio líquido e homogeneizado em 0,75 ml de tampão de extração (sacarose 0,7 M; Tris 0,5 M; HCl 30 mM; EDTA 50 mM; KCl 0,1 M; ditiotretitol 40 mM). O homogeneizado foi incubado a temperatura ambiente por 15 min e a seguir, foram realizadas duas extrações consecutivas com 1 volume de fenol. As proteínas foram precipitadas da fase fenólica através da adição de 5

volumes de acetato de amônio 0,1 M em metanol e incubação por 2 h a -20°C. Após centrifugação por 15 min a 7.000 rpm, o precipitado foi lavado por 2 vezes com acetato de amônio 0,1 M em metanol e uma vez com acetona 80%. O precipitado foi mantido a temperatura ambiente até secar, sendo então ressuscitado em 30-75 µl de tampão Laemmli. As amostras foram mantidas a -20°C. As proteínas foram quantificadas pelo método de Bradford (1976).

2.4. Separação de proteínas em SDS-PAGE

As proteínas de reserva e total foram analisadas através de eletroforese em gel de acrilamida 12% contendo SDS (Laemmli, 1970). O gel foi constituído por um gel de empacotamento (SDS 0,1%; acrilamida 5%; bis-acrilamida 0,25%; Tris-HCl 0,75 M, pH 6,8; TEMED 0,04% e persulfato de amônio 0,08%) e um gel de resolução (SDS 0,1%; acrilamida 12%; bis-acrilamida 0,06%; Tris-HCl 0,375 M, pH 8,8; TEMED 0,02% e persulfato de amônio 0,5%). A eletroforese foi efetuada em tampão Tris-glicina (Tris base 6,06 g/l, glicina 28,8 g/l e SDS 1 g/l), a 7,5 mA até a frente de azul de bromofenol atingir o gel de resolução e a 15-25 mA até o final do gel resolução. Para o cálculo do peso molecular aparente das proteínas, um padrão de peso molecular (*protein molecular weight standards range* 14.300 - 200.00/Gibco BRL) foi submetido à eletroforese juntamente com as amostras. Após a eletroforese, o gel foi corado com *Coomassie Brilliant Blue R-250* ou com prata. A análise densitométrica dos géis foi efetuada em um densitômetro a laser (UltraScan XL, LKB).

2.5. Separação de proteínas em 2D-PAGE

As proteínas totais foram analisadas através de eletroforese em duas dimensões, segundo modificações na metodologia descrito por De Mot & Vanderleyden (1989).

Aproximadamente, 100 µg de proteína foram submetidas à eletroforese no gel de primeira dimensão (1,5 mm de diâmetro e 13 cm de comprimento). O gel da primeira dimensão foi constituído por 3,6% de acrilamida; 0,21% de bis-acrilamida; 7,2% de anfólitos (BioRad) pH 5-8 e pH 3-10 na proporção de 5:1 (v/v); 2% de Nonidet P-40 e 55%

(p/v) de uréia. A focalização isoeétrica foi efetuada a 400 V por 18 h utilizando-se no reservatório superior NaOH 20 mM e no inferior H_3PO_4 10 mM.

O gel da primeira dimensão foi equilibrado e colocado sobre um gel de SDS-PAGE (Laemmli, 1970) para separação das proteínas de acordo com o seu peso molecular e tratado com solução de equilíbrio por 10 min.. O gel da segunda dimensão era constituído por um gel de empacotamento (SDS 0,1%; acrilamida 5%; bis-acrilamida 0,25%; Tris-HCl 0,75 M, pH 6,8; TEMED 0,04% e persulfato de amônio 0,08%) e um gel de resolução (SDS 0,1%; acrilamida 10%; bis-acrilamida 0,06%; Tris-HCl 0,375 M, pH 8,8; TEMED 0,02% e persulfato de amônio 0,5%). A eletroforese foi efetuada em tampão Tris-glicina (Tris base 6,06 g/l, glicina 28,8 g/l e SDS 1 g/l), a 7,5 mA até a frente de azul de bromofenol atingir o gel de resolução e a 15-25 mA até o final do gel de resolução. Para o cálculo do peso molecular aparente das proteínas, um padrão de peso molecular (*protein molecular weight standards range* 14.300 - 200.00/Gibco BRL) foi submetido à eletroforese juntamente com as amostras. Após a eletroforese, o gel foi corado com *Coomassie Brilliant Blue R-250* ou prata. A análise dos géis foi efetuada no programa *Image Master 2D v 3.01* (Amersham Pharmacia).

2.6. Transferência das proteínas para membrana de PVDF e seqüenciamento

As proteínas de reserva do embrião e megagametófito maduros, separadas através de 2D-PAGE, foram transferidas para membranas de PVDF (*Immobilon PSQ*, Millipore) utilizando-se a cuba de transferência TE62X *Transfor II* (Amersham Pharmacia). A transferência foi efetuada com tampão CAPS (CAPS 10 mM, pH 11,0; metanol 10%) por 1 h, 4°C, 650 mA. Após a transferência, as membranas foram coradas por 15 min com *Coomassie Brilliant Blue R 250* (*Coomassie Brilliant Blue R 250* 0,025%, metanol 40%) e descoradas durante a noite em metanol 40%. As membranas foram mantidas à temperatura ambiente até secar. As bandas de interesse foram cortadas das membranas, colocadas em tubo eppendorf e enviadas para a Universidade Estadual do Norte Fluminense (UENF) para seqüenciamento de aminoácidos na região N'-terminal das proteínas. O seqüenciamento foi realizado no seqüenciador Shimadzu PPSQ-10 *Automated Protein Sequencer* pela química de Edman.

Para evitar bloqueio do N^o-terminal, alguns cuidados foram tomados. São eles: (1) não foi utilizada uréia no tampão de amostra, (2) as amostras foram incubadas a 37°C por 15 min, (3) o gel de acrilamida da segunda dimensão foi feito com 18 h de antecedência, (4) foi adicionado no tampão do reservatório superior 11,4 mg/l de tioglicolato (Sigma) e (5) a transferência das proteínas para a membrana foi realizada com tampão CAPS a 4°C para evitar elevação da temperatura.

2.7. Análise das seqüências do N^o-terminal das proteínas

As seqüências do N^o-terminal das proteínas de reserva de *A. angustifolia* foram comparadas em bases de dados utilizando o serviço BLAST-NCBI (*National Center of Biotechnology*, www.ncbi.nlm.nih.gov:80/entrez/BLAST/), subprograma BLASTP. Para o alinhamento final das seqüências foi utilizado o programa Clustal W v. 1.8, e a apresentação gráfica foi feita com o auxílio do programa BOXADE v. 3.2.

3. Resultados e discussão

3.1. Caracterização das proteínas de reserva de *A. angustifolia*

Como descrito nos métodos, as proteínas de reserva foram isoladas do embrião e megagametófito de sementes maduras de *A. angustifolia*, utilizando-se SDS (Misra & Green, 1990; Krassowski & Owens, 1993; King & Gifford, 1997) ou NaCl (Hakman, 1993). Diferenças importantes foram observadas entre as diferentes frações protéicas isoladas com NaCl: **G1** (cristalóides), **G2** (globóides) e **Alb** (albuminas). Estas diferenças foram observadas tanto na quantidade de proteínas (Tabela 1) presente em cada uma das frações, quanto no perfil eletroforético em SDS-PAGE (Figura 1A).

A classe das globóides (**G2**) foi a mais abundante, representando 90% das proteínas extraídas do embrião e megagametófito da semente madura de *A. angustifolia* (Tabela 1). A análise eletroforética, em condições redutoras, das classes protéicas **G1**, **G2** e **Alb** seguida de análise densitométrica mostrou uma grande semelhança entre o perfil eletroforético de **G2** e das proteínas totais isoladas com SDS, tanto no embrião quanto no megagametófito (Figura 1B). Estes resultados confirmam os dados apresentados na Tabela 1, indicando ser

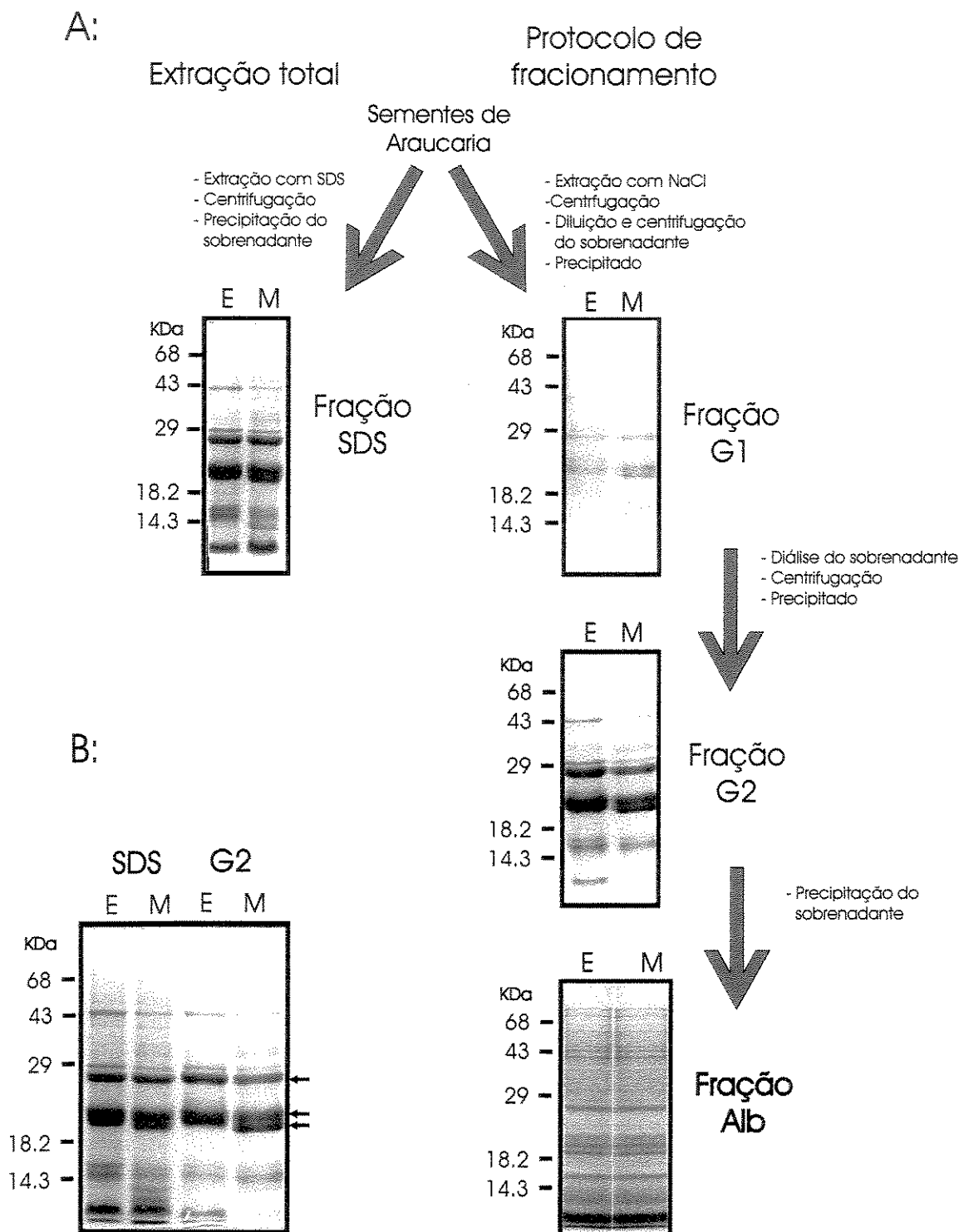


Figura 1. Principais proteínas de reserva de *A. angustifolia*. A: Esquema da estratégia de extração das diferentes frações protéicas e perfil eletroforético em SDS-PAGE 12% de 20 µg de proteínas de cada fração. B: Perfil eletroforético em SDS-PAGE 12% das principais proteínas de reserva do embrião (E) e megagametófito (M) da semente madura coincide nas frações SDS e G2. (indicados por setas os polipeptídeos de 28 e 20-21 kDa). Padrão de peso molecular indicado a esquerda em kDa. Os géis foram corados com *Coomassie blue*.

Tabela 1. Quantidade em mg e porcentagem estimada de proteínas nas frações G1, G2 e Alb, extraídas do megagametófito e embrião da semente madura de *A. angustifolia*. Os valores representam as médias de duas extrações com seis sementes cada. Os valores de proteína total foram obtidos a partir da soma dos valores obtidos para G1, G2 e Alb.

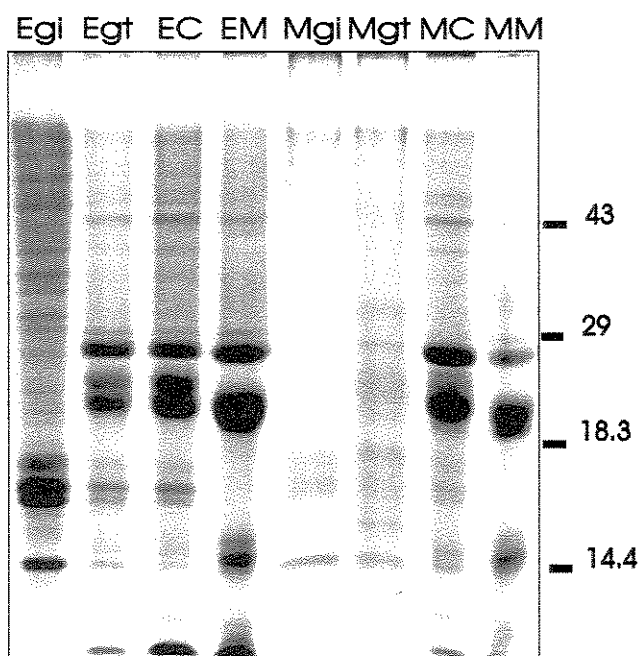
	Proteína Total (mg)	Fração G1 (mg)	Fração G1 (%)	Fração G2 (mg)	Fração G2 (%)	Fração Alb (mg)	Fração Alb (%)
Embrião	17.69 ± 2.62	1.36 ± 0.002	7.6	15.97 ± 2.6	90.2	0.52 ± 0.28	2.9
Megagametófito	24.95 ± 2.15	0.96 ± 0.09	3.8	22.51 ± 1.86	90.2	1.47 ± 0.29	5.9

G2 a principal classe de proteínas de reserva de *A. angustifolia*. Estes resultados diferem dos obtidos para outras coníferas. Isto é, outros autores demonstraram que as cristalóides (fração **G1**) são as principais proteínas de reserva em *Picea glauca* (Misra & Green, 1991), *Picea abies* (Hakman, 1993) e *Pinus taeda* L. (Groome *et al.*, 1991).

Os polipeptídeos de 20-21 e 28 kDa, encontrados em **SDS** e **G2** (Figura 1B), representam as principais proteínas de reserva do embrião e megagametófito de *A. angustifolia*. Na faixa de 20-21 kDa, 2 bandas foram observadas no megagametófito enquanto que no embrião, foi observada apenas uma banda que pode representar vários polipeptídeos com migração muito próxima (Figura 1B). Tanto no embrião como no megagametófito, estes polipeptídeos foram chamados de “grupo” de 20-21 kDa.

Através de análise densitométrica dos resultados mostrados na Figura 1B foi observado que no embrião, o grupo de 20-21 kDa representa 35% das proteínas presentes na fração **SDS** e 45% das proteínas presentes na fração **G2**. O polipeptídeo de 28 kDa representa 19% das proteínas presentes na fração **G2** e 13% das proteínas presentes na fração **SDS**. No megagametófito, o grupo de 20-21 kDa representa 31% das proteínas da fração **SDS** e 49% das proteínas da fração **G2**. O polipeptídeo de 28 kDa representa 13% das proteínas na fração **SDS** e 18% das proteínas da fração **G2**.

A:



B:

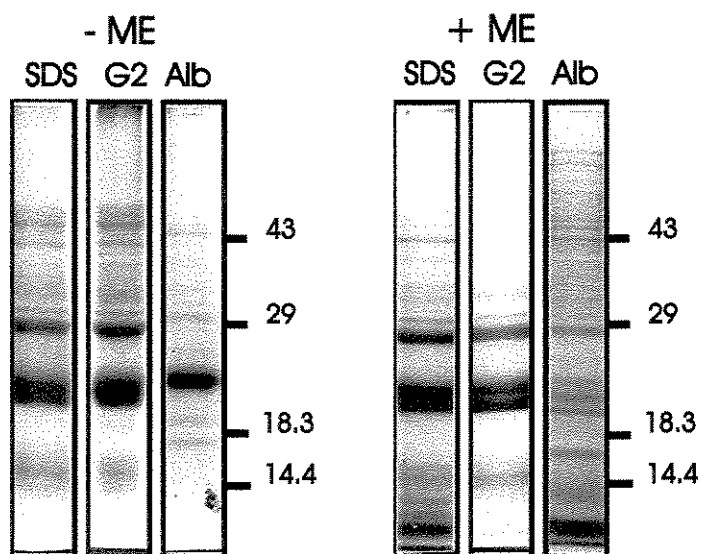


Figura 2. Perfil eletroforético em SDS-PAGE 12% das principais proteínas de reserva da semente de *A. angustifolia*. A: Proteínas extraídas com SDS do embrião (**E**) e megagametófito (**M**) nos diferentes estágios de desenvolvimento (**gi**- globular inicial, **gt**- globular tardio, **C**-cotiledonar e **M**- maduro). B: Mudança do padrão eletroforético das proteínas do embrião maduro das frações **SDS**, **G2** e **Alb** em SDS-PAGE em condições redutoras (**+ME**) e não redutoras (**-ME**). Padrão de peso molecular indicado a direita em kDa.

O perfil eletroforético em SDS-PAGE das proteínas totais extraídas com SDS do embrião e megagametófito de *A. angustifolia* nas diferentes fases do desenvolvimento embrionário mostrou que as proteínas acumuladas na maturação da semente tem peso molecular aparente de 12-13, 20-21 e 28 kDa (Figura 2A). Os polipeptídeos de 20-21 e 28 kDa acumulam-se no embrião desde o estágio globular tardio (EGt), coletado no início de fevereiro. Já no megagametófito, estes polipeptídeos aparecem só no estágio cotiledonar do embrião, coletado no final de fevereiro. Um grupo de proteínas de baixo peso molecular (12-13 kDa) é sintetizado no embrião e megagametófito desde o início do desenvolvimento do embrião de *A. angustifolia*. No embrião existem ainda, três polipeptídeos de 16, 17 e 18 kDa que aparecem nos estágios globular inicial (EGi), globular tardio (EGt) e cotiledonar (EC). Estes polipeptídeos desaparecem com o processo de maturação da semente (Figura 2A).

É importante mencionar que em condições não redutoras (-ME), as globóides (**G2**) não mudam o perfil eletroforético (Figura 2B). Isto sugere que a estrutura destas proteínas de reserva não apresenta pontes dissulfídicas intercadeias, como proposto para as cristalóides de *Picea* (Hakman, 1993) e *Pinus* (Groome *et al.*, 1991).

A quantidade de proteínas encontradas na fração **Alb** (albuminas) é baixa (Tabela 1). Esta classe de proteínas é composta por uma grande quantidade de polipeptídeos com peso molecular aparente variando de 10 a 100 kDa e o perfil eletroforético das albuminas do embrião é muito similar ao do megagametófito (Figura 1A). No entanto, em condições não redutoras, no embrião um polipeptídeo de 22 kDa é o mais abundante (Figura 2B). Em condições redutoras aparece um polipeptídeo com tamanho aproximado de 10-12 kDa tanto em **SDS** como em **Alb** (Figura 2B). Este resultado indica que o polipeptídeo de 22 kDa é composto por duas subunidades de 10-12 kDa unidas por pontes dissulfídicas. Estas características (dímeros formados por monômeros de 10-14 kDa unidos por pontes bissulfídicas) coincidem com as encontradas em outras proteínas de reserva descritas em angiospermas e que compõem o grupo das albuminas 2S (Shewry & Casey, 1999).

3.2. Análise da síntese de proteínas durante o desenvolvimento da semente de *A. angustifolia*

A eletroforese bidimensional é uma técnica amplamente utilizada no estudo da síntese protéica em mutantes (Damerval & Le Guilloux, 1998), em tecidos submetidos a condições de estresse (Costa *et al.*, 1998) e na caracterização das proteínas de reserva (Hirano, 1997). O seqüenciamento do N'-terminal das proteínas separadas através de 2D-PAGE é uma ferramenta poderosa para caracterização de proteínas (Allona *et al.*, 1994; Sun *et al.*, 1996).

Com o objetivo de aprofundar a análise das proteínas de *A. angustifolia*, 100 µg de proteína total de embrião (E) e megagametófito (M) de sementes nos estágios de desenvolvimento do embrião: globular (G), cotiledonar (C) e maduro (M) foram submetidas à eletroforese bidimensional e visualizadas através de coloração com prata (Figura 3). Os géis foram analisados no programa *Image Master 2D* v. 3.01 (Amersham Pharmacia). Os *spots* foram numerados e quantificados em cada gel. A seguir, os diferentes estágios de desenvolvimento do embrião e megagametófito foram comparados e os *spots* foram alinhados, na tentativa de se obter um mapa geral das mudanças da síntese protéica nestes tecidos durante o desenvolvimento da semente (Tabela 2).

Embora cada *spot* não represente necessariamente uma proteína diferente neste mapa (proteínas fosforiladas e não fosforiladas aparecem como *spots* de pI diferente, assim como a glicosilação e outras modificações mudam consideravelmente o peso molecular dos polipeptídeos) a análise dos géis reflete indiretamente a síntese de proteínas em cada tecido nos diferentes estágios de desenvolvimento (Mullen *et al.*, 1996).

No embrião e no megagametófito, a síntese protéica concentra-se nos estágios iniciais do desenvolvimento, diminuindo consideravelmente no estágio maduro. Isto coincide com a entrada da semente no período de dormência. Ao mesmo tempo, no estágio maduro aumenta a quantidade de proteínas sintetizadas especificamente neste estágio tanto no embrião (75%) como no megagametófito (53%). Isto possivelmente tem relação com a grande quantidade de genes *lea* (*late embryogenesis abundant*) expressos e ao acúmulo de proteínas de reserva. Estes dados podem servir de base para estudos futuros do genoma funcional de *A. angustifolia*. Os dados gerados nesta análise (Figura 3) possibilitaram a

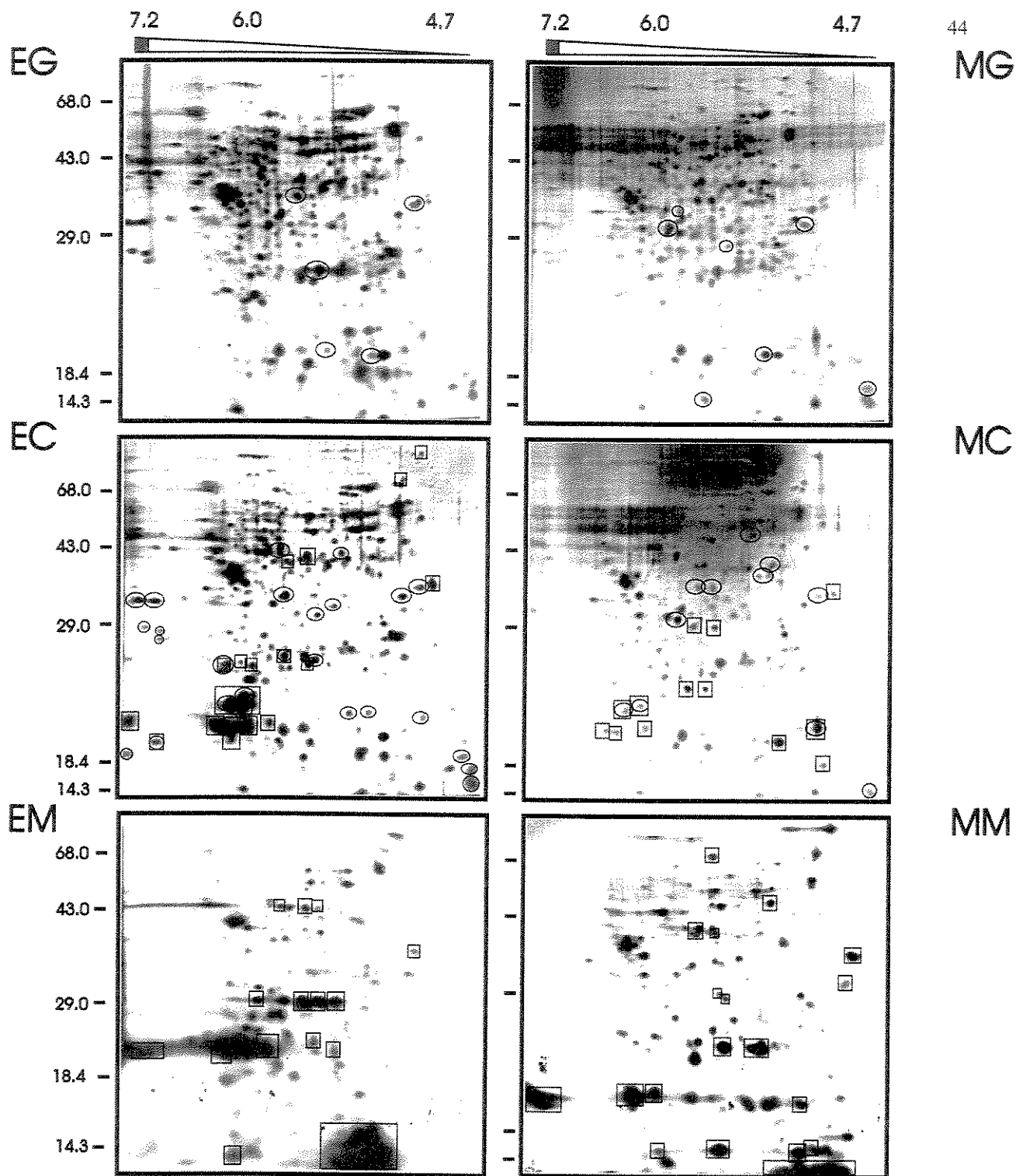


Figura 3. Análise através de 2D-PAGE das proteínas sintetizadas diferencialmente no estágio globular (G), cotiledonar (C) e maduro (M) do embrião zigótico (E) e megagametófito (M) de *A. angustifolia*. As proteínas foram separadas em grupos segundo a ausência (O) e presença (□) nos diferentes estágios de desenvolvimento.

Tabela 2. Análise do número de proteínas detectadas nos estágios globular, cotiledonar e maduro do embrião e megagametófito de *A. angustifolia*.

	Embrião			Total	Megagametófito			Total
	EG	EC	EM		MG	MC	MM	
Total de polipeptídeos por estágio	282	286	101	669	212	175	143	530
Polipeptídeos específicos por estágio	100	104	77		112	63	76	
(%)	35	35	75		47	36	53	

localização e classificação dos polipeptídeos que compõem os dois grupos principais de proteínas de reserva. Estes dados contribuíram para seleção dos polipeptídeos que foram seqüenciados. A detecção e numeração das proteínas destes géis, usada no alinhamento, é mostrada no Anexo 1.

3.3. Análise através de 2D-PAGE das proteínas de reserva de *A. angustifolia*

Em *A. angustifolia* as principais proteínas de reserva da semente (fração **G2**) são subdivididas, através de SDS-PAGE, em 2 grupos com peso molecular aparente de 20-21 kDa e 28 kDa, como descrito anteriormente (Figura 1A). Quando estas proteínas são submetidas a 2D-PAGE, cada grupo é subdividido em vários polipeptídeos (Figura 4). O grupo de 20-21 kDa é subdividido em 7-8 polipeptídeos no megagametófito e 8-9 no embrião, com pIs que variam de 5,6 a 7,2. A classe de 28 kDa é subdividida, no embrião e megagametófito, em aproximadamente 9-11 polipeptídeos, com pIs que variam de 5,3 a 6,2 (Figura 4). Outros grupos de proteínas são observados no embrião e megagametófito no gel bidimensional. Entre eles, está um grupo com 3-4 polipeptídeos de aproximadamente 45 kDa e pIs que variam de 6,0 a 5,6. Grupos de proteínas de baixo peso molecular (menos de 14 kDa) também são observados (Figura 4).

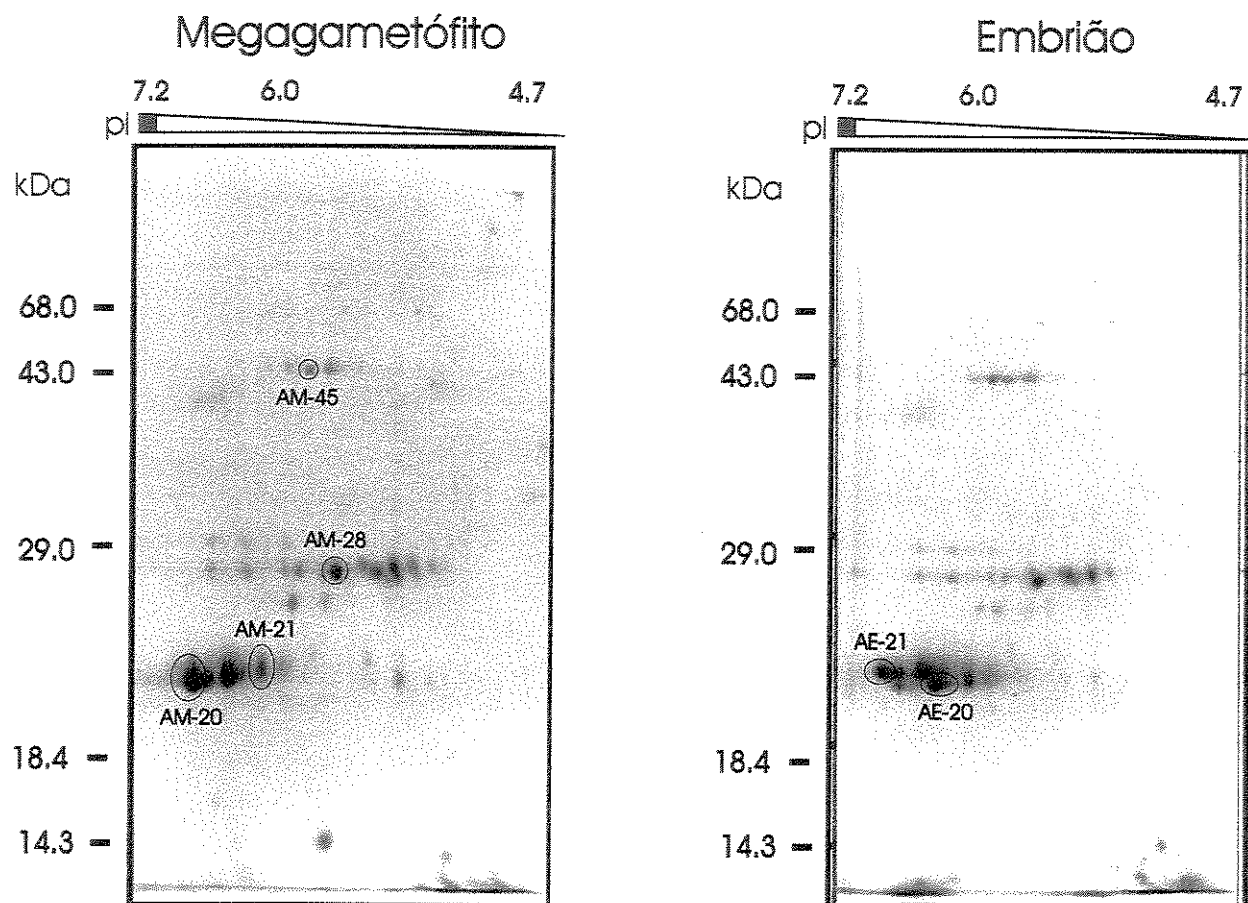


Figura 4. Separação através de 2D-PAGE das proteínas de reserva de *A. angustifolia* extraídas na fração G2. Os géis foram corados com *Coomassie blue* e as proteínas transferidas para membrana PVDF. As proteínas escolhidas para sequenciamento do N'-terminal estão assinaladas nos géis. Padrão de peso molecular indicado a esquerda em kDa.

Visando uma melhor caracterização, estas proteínas foram transferidas para membranas de PVDF e algumas selecionadas para seqüenciamento (Figura 4). Essas proteínas foram escolhidas com base no peso molecular aparente, entre as de maior quantidade e melhor separação no gel.

Poli-peptídeos representando as principais classes de proteínas de reserva no megagametófito (AM45, AM28, AM21 e AM20) e embrião (AE21 e AE20) (Figura 4) foram enviados para a Universidade Estadual do Norte Fluminense (UENF) para seqüenciamento do N'-terminal (Figura 4). Foram obtidas seqüências do N'-terminal das proteínas:

AE-21	RHGEEEGRNPHYVFKEEEQ
AM-20	EDGRNPHYVFKEEENNERSLATDA
AE-20	YRGPRFTYYFKEEHG

O alinhamento destas seqüências utilizando o programa Clustal W mostrou similaridade entre AE-21 e AM-20 (Figura 5A), sugerindo que elas fazem parte de uma mesma família de genes. Na tentativa de esclarecer a identidade destas proteínas, as seqüências foram comparadas em bancos de dados (NCBI/BLAST, subprograma *blastp*, banco *nr*). As seqüências AE-21 e AM-20 apresentaram similaridade com seqüências descritas para outras globulinas de gimnospermas como *Zamia furfuraceae* (Braun *et al.*, 1996) e *Picea glauca* (GenBank X63191) que por sua vez, são similares às vicilinas 7S de angiospermas (Figura 5B). Estes dados sugerem que AE-21 e AM-20 são proteínas pertencentes ao grupo das vicilinas 7S.

3.4. As principais proteínas de reserva da semente de *A. angustifolia* pertencem ao grupo das vicilinas 7S

Os resultados obtidos indicam que as principais proteínas de reserva do embrião e megagametófito da semente de *A. angustifolia* (fração **G2**) são solúveis em solução salina (0,25 M NaCl), não apresentam pontes dissulfídicas e tem peso molecular de 20-21 kDa e 28 kDa. Todas estas características diferem das já descritas para outras proteínas de reserva de coníferas (Green *et al.*, 1991; Krassowski & Owens, 1993; Hakman, 1993; Shewry & Casey, 1999).

A:

```

AE-21      1 RHGEEDGRNPYVFKEEE-----
AM-20      1 ----EDGRNPYVFKEEENERSLATDA
AE-20      1 ---YRGPRFTYVFKEEHH-----
consensus      ...*...*...*...

```

B:

```

AM-20      1 -----GRNPYVFKEEENERSLATDA
AE-21      1 --RHGEEDGRNPYVFKEEEO-----
P. glauca  42 GREEEREENPYVFESSTRASSEAG
Zamia      20 ACFSELLEIPYVEDRSIVITVQCKAG
consensus      ...*...*...*...

```

Figura 5. Alinhamento das seqüências obtidas no seqüenciamento do N'-terminal de AM-20, AE-21 e AE-20 (A) e comparação do N'-terminal das proteínas AM-20 e AE-21 de *A. angustifolia* com seqüências do N'-terminal de globulinas similares as vicilinas 7S de angiospermas (B). *P. glauca*- Genbank número X63191 e *Zamia furfuraceae* (Braun *et al.*, 1996). Na linha *consensus*, os resíduos idênticos são representados com asteriscos, e os resíduos conservados por um ponto. Os números indicam a posição ocupada pelo aminoácido que inicia a seqüência no polipeptídeo original maduro.

Nas análises através de 2D-PAGE, estas proteínas apresentaram valores de pI que variavam de 6,2 a 5,4, e as seqüências do N'-terminal de AM-20 e AE-21, apresentaram similaridade com proteínas similares as vicilinas 7S de angiospermas (Braun *et al.*, 1996). Foi detectado no N'-terminal a presença do motivo NPYVF, conservado em AE-21 e AM-20. Este motivo coincide com o motivo NP(Y/F)VF, altamente conservado nas vicilinas 7S (Lawrence *et al.*, 1994; Braun *et al.*, 1996). Segundo Lawrence *et al.* (1994), este motivo permite a formação de uma “ponte” entre os domínios N' e C' das vicilinas 7S, estabilizando a sua interação (Lawrence *et al.*, 1994). A conservação deste motivo é observada tanto na estrutura tridimensional descrita para cristais de Phaseolina (Lawrence *et al.*, 1994) e Canavalina (Ko *et al.*, 1993), quanto no alinhamento de seqüências do grupo das vicilinas 7S de angiospermas e gimnospermas (Lawrence *et al.*, 1994).

Segundo alguns autores, as principais globulinas de reserva das sementes de angiospermas e gimnospermas tem como ancestral comum, proteínas monoméricas de procariotos envolvidas na desidratação de esporos (esferulinas) (Bäumlein *et al.*, 1995). Shutov *et al.* (1995) propuseram a equivalência estrutural dos domínios N' e C'-terminal de vicilinas 7S e leguminas 11S de angiospermas e gimnospermas ao encontrar similaridade entre as seqüências destas proteínas. Com base nestes resultados, foi proposto um ancestral comum para as vicilinas 7S e leguminas 11S de angiospermas e gimnospermas (Bäumlein *et al.*, 1995; Shutov *et al.*, 1995, 1998). Esta evolução molecular teria ocorrido na forma de duplicação do gene nas gimnospermas menos evoluídas (Shutov *et al.*, 1998).

As propriedades das vicilinas de reserva de *A. angustifolia* descritas neste trabalho, sugerem que estas proteínas de reserva podem ser globulinas 7S na forma monomérica (20-21 kDa), ou polipeptídios resultantes da clivagem interna de um precursor de aproximadamente 45 kDa (20-21 e 28 kDa) (Figura 4). Estas perguntas serão respondidas no capítulo 3 com a análise do seqüenciamento do gene completo da vicilina de *A. angustifolia*.

5. Referências Bibliográficas

- Allona I., Casado R., Aragoncillo C. (1994) Biochemical genetics of a 7S globulin-like protein from *Pinus pinaster* seed. *Theor. Appl. Genet.* 88: 454.
- Bäumlein H., Braun H., Kakhovskaya I. A., Shutov A. D. (1995) Seed storage proteins of spermatophytes share a common ancestor with desiccation proteins of fungi. *J. Mol. Evol.* 41: 1070.
- Bradford M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248.
- Braun H., Czihal A., Shutov A. D., Bäumlein H. (1996) A vicilin-like seed protein of cycads: similarity to sucrose binding proteins. *Plant Mol. Biol.* 31:35.
- Chatthai M., Misra S. (1998) Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziessi* [Mirb] Franco. *Planta* 206: 138.
- Damerval C., Le Guilloux M. (1998) Characterizations of novel proteins affected by the *o2* mutation and expressed during maize endosperm development. *Mol. Gen. Genet.* 257: 354.
- De Mot R., Vanderleyden J. (1989) Application of two-dimensional protein analysis for strain fingerprinting and mutant analysis of *Azospirillum* species. *Can. J. Microbiol.* 35: 960.
- Gifford D. J. (1988) An electrophoretic analysis of the seed proteins from *Pinus monticola* and eight other species of pine. *Can. J. Bot.* 66: 1808.
- Green M. J., McLeod J. K., Misra S. (1991) Characterization of Douglas fir protein body composition by SDS-PAGE and electron microscopy. *Plant Physiol. Biochem.* 29: 49.
- Groome M. C., Axler S. R., Gifford D. J. (1991) Hydrolysis of lipid and protein reserves in lobolly pine seeds in relation to protein electrophoretic patterns following imbibition. *Physiol. Plant* 83: 99.
- Hager K. P., Braun H., Czihal A., Mullen B., Baumlein H. (1995) Evolution of seed storage protein genes: Legumin genes of *Ginkgo biloba*. *J. Mol. Evol.* 41: 457.

- 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.
- Hirano H. (1997) Screening of rice genes from the cDNA catalog using the data obtained by protein sequencing. *J. Prot. Chem.* 16 (5): 533.
- King J. E., Gifford D. J. (1997) Amino acid utilization in seed of Loblolly Pine during germination and early seedling growth. *Plant Physiol.* 113: 1125.
- Ko T. P., Ng J. D., McPerson A. (1993) The three-dimensional structure of canavalin from jack bean. *Plant Physiol.* 101: 729.
- Krassowski M. J., Owens J. N. (1993) The ultrastructural and histochemical post-fertilization megagametophyte and zygotic embryo development of interior spruce. *Can. J. Bot.* 71: 98.
- Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 247: 32.
- Lawrence M. C., Izard T., Beuchat M., Blagrove R. J., Colman P. M. (1994) Structure of phaseolin at 2.2 Å resolution: Implications for common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* 238: 748.
- Misra S. (1994) Conifer zygotic embryogenesis, somatic embryogenesis and seed germination: Biochemical and molecular advances. *Seed Science Res.* 4: 357.
- Misra S., Green, M. J. (1990) Developmental gene expression in conifer embryogenesis and germination. I. Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* [Moench] Voss. *Plant Sci.* 68: 163.
- Misra S., Green, M. J. (1991) Developmental gene expression in conifer embryogenesis and germination. II. Crystalloid protein synthesis in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss. *Plant Sci.* 78: 61.
- Mullen R. T., King J. E., Gifford D. J. (1996) Changes in mRNA populations during loblolly pine (*Pinus taeda*) seed stratification, germination and post-germinative growth. *Physiol. Plant.* 97: 545
- Shewry R., Casey R. (1999) Seed proteins. Kluwer Academic Publishers.
- Shewry P. R., Napier J. A., Tatham A. S. (1995) Seed storage proteins. Structures and Biosynthesis. *The Plant Cell* 7: 945.

- Shutov A. D., Braun H., Chesnokov Y. V., Horstmann C., Kakhovskaya I. A., Baumlein H. (1998) Sequence peculiarity of Gnetalean Legumin-like seed storage proteins. *J. Mol. Evol.* 47: 486.
- Shutov A. D., Kakhovskaya I. A., Braun H., Baumlein H., Muntz K. (1995) Legumin-like and Vicilin-like seed storage proteins: Evidence for a common single-domain ancestral gene. *J. Mol. Evol.* 41: 1057.
- Sun J., Nakaguama H., Karita S., Ohmiya K., Hattori T. (1996) Rice embryo globulins: amino-terminal amino acid sequences, cDNA cloning and expression. *Plant Cell Physiol.* 37 (5): 612.

ANEXO 1

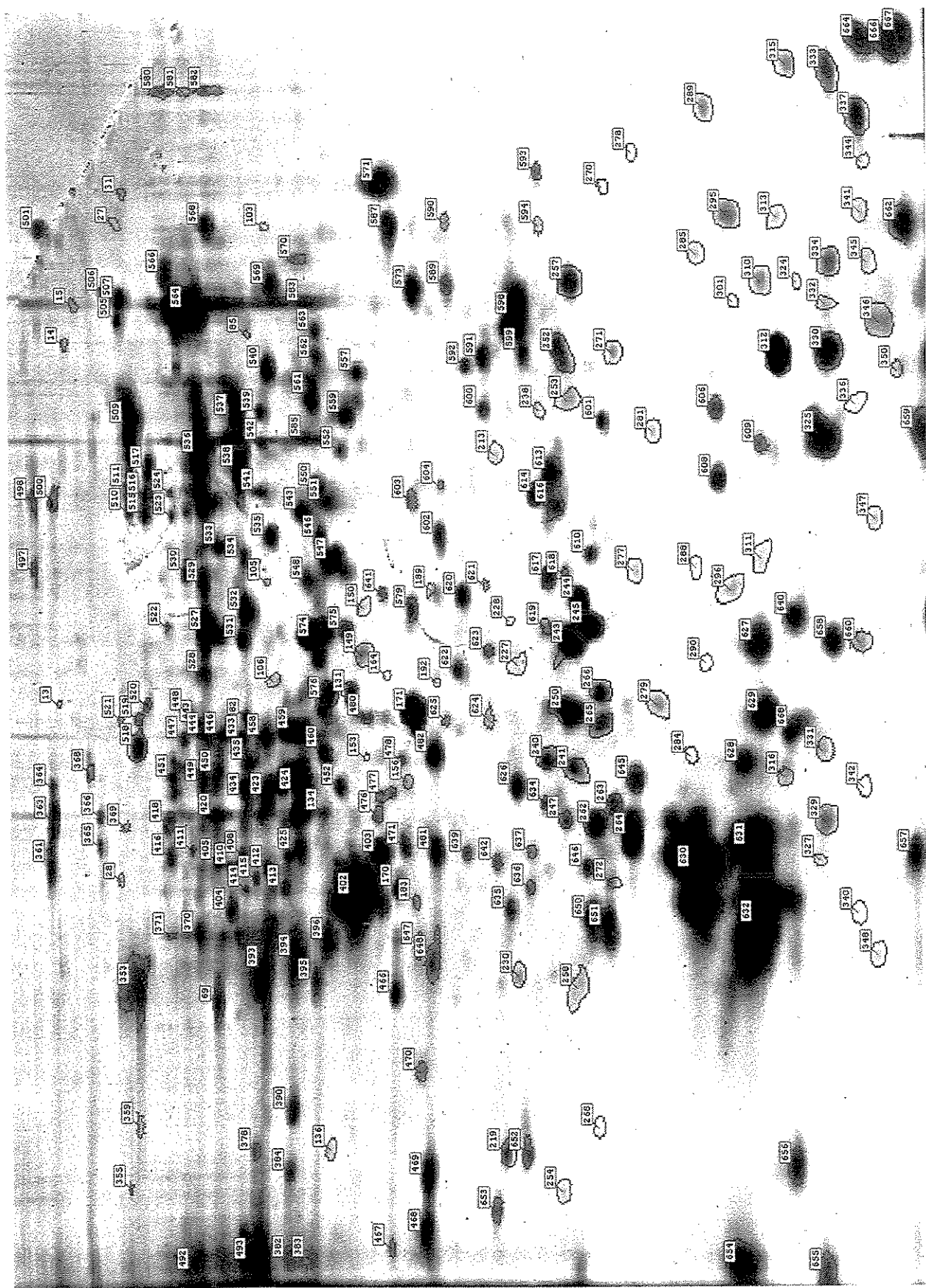


Figura 2. Gel de 2D-PAGE corado com prata, contendo proteínas extraídas do embrião zigótico de *A. angustifolia* no estágio cotiledonar. Este gel foi tomado como "referência" no alinhamento e quantificação das proteínas do embrião nos estágios globular, cotiledonar e maduro.

CAPÍTULO 2

Identificação de genes expressos diferencialmente durante o desenvolvimento do embrião zigótico de *Araucaria angustifolia*

1. Introdução

O acúmulo de reservas como lipídios, proteínas e carboidratos, assim como os mecanismos de regulação que levam à expressão temporal e tecido específica de genes durante o desenvolvimento das sementes em angiospermas (Goldberg *et al.*, 1989; Zimmerman, 1993;; Wobus & Weber, 1999; Girke *et al.*, 2000) e gimnospermas (Misra, 1994; Mullen *et al.*, 1996; Cairney *et al.*, 1999) vem sendo intensamente investigados pela comunidade científica.

Na última década, as pesquisas envolvendo o desenvolvimento do embrião e a expressão gênica em coníferas, vem aumentando, principalmente em *Picea* (Misra & Green, 1991; Leal & Misra, 1993a; Krassowski & Owens, 1993; Misra, 1994), *Pinus* (Mullen *et al.*, 1996; Neutelings *et al.*, 1998; Cairney *et al.*, 1999) e *Pseudotsuga* (Chatthai & Misra, 1998). Trabalhos envolvendo embriogênese somática em coníferas (Jain *et al.*, 1995) relataram genes com expressão diferencial no processo de crescimento e maturação, genes que respondiam ao choque térmico, genes induzidos por ABA e genes com expressão específica no embrião somático (Dong & Dustan, 1996, 1997, 1999). Cairney *et al.* (1999) compararam os padrões de expressão gênica no desenvolvimento embrionário do embrião zigótico e somático de *Pinus taeda*, na tentativa de implementar uma metodologia de controle de marcadores moleculares nos protocolos de embriogênese somática em *Pinus*. Jarvis *et al.* (1996) isolaram os primeiros genes *lea* de gimnospermas trabalhando com sementes maduras de *Pseudotsuga menziesii*. Foram isolados os cDNAs das proteínas de reserva de *Picea glauca/engelmannii* (Newton *et al.*, 1992) e *Pseudotsuga menziesii* (Leal & Misra, 1993b). Mullen *et al.* (1996) estudaram o padrão de expressão de populações de mRNA no embrião e megagametófito em *Pinus taeda* e concluíram que é muito similar ao descrito para angiospermas (Goldberg *et al.*, 1989; Zimmerman, 1993; Girke *et al.*, 2000). Contudo, os resultados obtidos e a quantidade de pesquisadores envolvidos em projetos nesta área ainda são insuficientes.

Em *Araucaria angustifolia*, a conífera nativa de maior importância, estudos visando o desenvolvimento do embrião mostraram que este difere em vários aspectos morfológicos do desenvolvimento do embrião de outras coníferas como *Pinus*, *Picea* (Pinheiro do Canadá) e *Pseudotsuga* (ver Introdução, Figura 1). Este fato torna interessante, do ponto de

vista acadêmico, a realização de estudos moleculares sobre o desenvolvimento do embrião de *A. angustifolia*.

Estudos sobre a expressão diferencial de genes durante o desenvolvimento do embrião de *A. angustifolia* ainda são inéditos. Assim sendo, com o objetivo de isolar e caracterizar genes com expressão diferencial durante o desenvolvimento do embrião zigótico de *A. angustifolia*, foi extraído RNA do embrião nos estágios globular (EG), cotiledonar (EC) e maduro (EM). Foi isolado também, RNA do megagametófito da semente madura (MM). Estes RNAs foram utilizados em experimentos de *display* diferencial de mRNA. A expressão diferencial dos cDNAs isolados foi confirmada através de experimentos de *slot blot* e *macroarray*. Os cDNAs que apresentaram os maiores coeficientes de expressão foram clonados e sequenciados. As seqüências obtidas foram comparadas em bancos de dados. Os resultados obtidos são relatados neste capítulo.

2. Materiais e Métodos

2.1. Coleta de sementes

Cones femininos foram coletados em povoações naturais da localidade de Bom Retiro, Santa Catarina, entre janeiro e maio de 1999 e 2000. As sementes (pinhões) foram desinfetadas, congeladas em nitrogênio líquido e armazenadas a -70°C.

Coletas :

fevereiro - sementes contendo embrião no estágio globular

março - sementes contendo embrião no estágio cotiledonar

abril - sementes contendo embrião no estágio maduro

2.2. *Display* diferencial de mRNA

Nos experimentos de *display* diferencial, RNA foi isolado de vários (5-20) embriões (para minimizar a variabilidade genética dada pelo fato de embriões de um mesmo cone terem pais diferentes) nos estágios de desenvolvimento: (1) globular tardio, (2) cotiledonar e (3) maduro. Foi isolado também, RNA do megagametófito maduro e dos cotilédones, eixo e zona radicular do embrião maduro. A partir destes RNAs, foram sintetizados cDNAs com os *primers* ancoradores: T₁₂VC e T₁₂VG. Os *primers* arbitrários (Operon Technologies Inc.) utilizados nas reações de amplificação por PCR foram: OPF-03 (CCTGATCACC), OPJ-02 (CCCGTTGGGA), OPJ-03 (TCTCCGCTTG), OPJ-05 (CTCCATGGGG), OPJ-06 (TCGTTCCGCA), OPJ-07 (CCTCTCGACA), OPJ-08 (CATACCGTGG) e OPJ-09 (TGAGCCTCAC).

2.2.1. Isolamento de RNA e tratamento com DNase

O isolamento de RNA foi feito segundo metodologia descrita por Preccott & Martin (1987), com algumas modificações.

Para o isolamento de RNA, o tecido vegetal foi macerado em nitrogênio líquido e homogeneizado em: Tris-HCl 50 mM, pH 9,0; EDTA 5 mM, pH 8,0; SDS 5%;

LiCl 150 mM na proporção de 1/10 (g/ml). A seguir, foi realizada uma extração com fenol: clorofórmio: álcool isoamílico (25:24:1, v/v/v) e uma com clorofórmio: álcool isoamílico (24:1, v/v). O RNA foi precipitado com 0,4 volumes de LiCl 8 M a -20°C por 2 h. Após centrifugação (10 000 rpm), o RNA foi ressuspensionado em 50 µl de água DEPC. Foi efetuada nova precipitação com 0,1 volume de acetato de sódio 3 M, pH 5,0 e 2,5 volumes de etanol absoluto. O RNA foi incubado a -20°C durante a noite. Após centrifugação (10000 rpm), o RNA foi lavado com etanol 70% e ressuspensionado em 80 µl de água DEPC.

O RNA foi tratado com DNase através da adição de: 0,5 µl de *RNA guard* (37,8 U/µl, Gibco BRL); 5 µl de Tris-HCl 0,1 M, pH 8,3; 5 µl de KCl 0,5 M; 5 µl de MgCl₂ 15 mM; 1 µl de DNase (10 U/µl); H₂O q.s.p. 100 µl. A reação foi mantida a 37°C por 1 h. Adicionou-se 100 µl de água e foi realizada uma extração com fenol: clorofórmio (3:1). O RNA foi precipitado com acetato de sódio 3 M, pH 5,2 (1/10V) e etanol absoluto (2,5V). Após centrifugação (10 000 rpm), o RNA foi ressuspensionado em água DEPC e quantificado em gel. A concentração foi ajustada para 1 µg/µl.

2.2.2. Síntese de cDNA

A síntese de cDNA foi efetuada com transcriptase reversa a partir de *primers* ancoradores do tipo T₁₂VN (Promicro). Para isto, em um tubo eppendorf foram misturados 2 µl do *primer* T₁₂VN (T₁₂VC ou T₁₂VG) (10 µM), 1 µl de RNA tratado com DNase (1 µg/µl) e 7 µl de água. A mistura foi incubada a 68°C por 5-10 min. A seguir, foram adicionados 9 µl da mistura: Tris-HCl 10 mM, pH 8,3; KCl 75 mM; MgCl₂ 3 mM; *Human placental RNase inhibitor* 20 U; DTT 10 mM; dNTPs 1 mM. Após 10 min de incubação à temperatura ambiente, foi adicionado 1 µl de transcriptase reversa M-MLV (20 U/µl) e a reação foi incubada por 50 min a 37°C. A enzima foi inativada em banho a 95°C por 5 min e os cDNAs foram armazenados a -20°C.

Para síntese de cDNA foram utilizadas também *beads* do kit *Ready to Go RT-PCR Beads* (Amersham Pharmacia). Para a utilização do kit, as *beads* foram colocadas em gelo. A seguir, foram adicionados RNA tratado com DNase (1µg/µl), 5 µl do *primer* T₁₂VN (10 µM) e água q.s.p. 50 µl. Após 5 min de incubação no gelo, a reação foi mantida

a 42°C por 30 min. A transcriptase reversa foi inativada a 95°C por 5 min. As reações foram armazenadas a -20°C. A utilização das *beads* permitiu resultados rápidos e eficazes. Assim sendo, foi dada a preferência à utilização das *beads*.

2.2.3. PCR a partir do cDNA e eletroforese das amostras amplificadas

Os *primers* ancoradores (T₁₂VC e T₁₂VG, Promicro) e os *primers* arbitrários (OPF-03, OPJ-02, OPJ-03, OPJ-05, OPJ-06, OPJ-07, OPJ-08 e OPJ-09; Operon) foram utilizados nas amplificações dos cDNAs em um termociclador (Perkin Elmer 2400 ou 9700). As reações foram feitas em duplicata e eram constituídas por: 2 µl do cDNA, 2 µl do *primer* ancorador (10 µM), 2 µl do *primer* arbitrário (10 µM) e 14 µl da mistura: 2 µl de tampão de amplificação 10x (Gibco BRL), 0,5 µl de MgCl₂ 50 mM, 2 µl de dNTP (20 µM de cada um), 0,5 µl de [α -³³P]dATP, 8,7 µl de água e 0,3 µl de *Taq* polimerase (5 U/µl, Gibco BRL). As condições de amplificação foram: 94°C (3 min), seguido de 40 ciclos de 94°C (30 s), 40°C (2 min) e 72°C (30 s). A extensão final foi a 72°C por 15 min.

Após a amplificação, 7 µl do produto de cada reação de PCR foram misturados com 5 µl do tampão: formamida deionizada 95% (v/v); EDTA 10 mM, pH 8,0; xileno cianol 0,09%; azul de bromofenol 0,09% (Xu *et al.*, 1997). As amostras foram incubadas a 80°C por 4 min e 10-12 µl foram aplicados no gel denaturante de acrilamida (6%). A corrida foi a 50 W por aproximadamente 4 h e 30 min. Após a corrida, o gel foi transferido para papel Whatman 3 MM e seco a vácuo. Os produtos de amplificação foram visualizados através de exposição a filme de raio-X (Hyperfilm, Amersham Pharmacia).

2.2.4. Isolamento, eluição e reamplificação das bandas diferenciais

Após ser revelado, o filme de raio-X foi posicionado sobre o gel, na mesma posição em que foi exposto, e as bandas diferenciais foram cortadas do gel. O DNA das bandas diferenciais foi eluído em 100-200 µl de água por 15 min a 95°C. A seguir, o DNA foi precipitado com 0,1 volume de acetato de sódio 3 M, pH 5,2 e 2 volumes de etanol absoluto, durante a noite a -20°C. Após centrifugação, o DNA foi ressuspensionado em 10 µl de água.

Quatro μl do DNA eluído foi amplificado novamente. Para isto, foram adicionados: Tris-HCl 2 mM, pH 8,3; KCl 5 mM; MgCl_2 4 mM; 1 μM do *primer* arbitrário; 1 μM do *primer* ancorador; 20 μM de dNTP e 2,5 U de *Taq* polimerase (Gibco BRL). O volume final da reação foi de 40 μl . As condições de amplificação foram as mesmas do item 2.2.3. O produto das reações de amplificação foram separados através de eletroforese em gel de 1,4% de agarose em tampão TAE 1x. As bandas de interesse foram cortadas do gel e o DNA foi purificado com o kit *Concert Rapid Gel Extraction System* (Gibco BRL).

2.3. Confirmação da expressão diferencial das bandas isoladas nos experimentos de *display* diferencial de mRNA

2.3.1. Síntese de cDNA visando a obtenção de sondas radioativas

O cDNA foi sintetizado a partir de RNA isolado de embriões zigóticos de *A. angustifolia* nos estágios globular, cotiledonar e maduro; cotilédone embrionário, eixo embrionário e parte radicular do embrião maduro; e megagametófito maduro. Para síntese de cDNA foram utilizadas *beads* do kit *Ready to Go RT-PCR Beads* (Amersham Pharmacia).

2.3.2. Transferência das bandas isoladas dos géis de *display* diferencial de mRNA para membranas de nylon

As membranas foram preparadas utilizando duas metodologias diferentes: *slot blot* (Xu *et al.*, 1997), inicialmente, e em uma segunda etapa, a metodologia de *dot blot* de alta densidade ou *macroarray* como descrito por Nanfei *et al.* (1997).

2.3.2.1. Transferência do DNA para membranas de nylon pela técnica de *slot blot*

Dez ng de DNA de cada banda isolada, em um volume final de 150 μl , foram denaturados por 10 min a 95°C. A seguir, adicionou-se 150 μl de SSC 20x filtrado. As amostras foram transferidas para membranas de nylon (Hybond-N, Amersham Pharmacia) com o auxílio de um equipamento próprio para *slot blot* (Hybri-Slot, Gibco BRL). Após a transferência das amostras, os *blots* foram lavados com 300 μl de SSC 10x

filtrado. As membranas foram secadas à temperatura ambiente e o DNA foi fixado a 80°C por 2 h. Todo o experimento foi realizado em duplicata.

2.3.2.2. Transferência do DNA para membranas de nylon pela técnica de *macroarray*

Trinta µl da reação de reamplificação das bandas isoladas por *display* diferencial (aproximadamente 600-800 ng) foram transferidos para uma microplaca. A seguir, foram adicionados 8 µl de NaOH 2 N e 42 µl de H₂O Milli-Q. A denaturação foi à temperatura ambiente por 15 min. Um µl desta mistura (8-10 ng) foi transferido para membrana Hybond-N com um carimbo manual de 96 amostras (V&P Scientific). As membranas foram incubadas por 2 h a 80°C e guardadas à temperatura ambiente, envolvidas em papel Whatman 3MM.

2.3.3. Obtenção de sondas radioativas e hibridização das membranas

Os cDNAs sintetizados a partir de RNA isolado dos embriões globular, cotiledonar e maduro; cotilédone, eixo e parte radicular do embrião maduro; e megagametófito maduro de *A. angustifolia*, foram amplificados, quantificados e utilizados como sonda nos experimentos de *slot blot* ou *macroarray* para confirmação da expressão diferencial das bandas isoladas por *display* diferencial. A quantificação foi feita em gel de 2% de agarose através de comparação com o padrão de massa e peso molecular $\phi\chi 174/Hae$ III (Gibco BRL). Aproximadamente 30 ng dos cDNAs foram marcados com o kit *Megaprimer DNA labelling system* (Amersham Pharmacia), seguindo-se as especificações do fabricante.

As membranas foram pré-hibridizadas a 42°C por 4-6 h em: formamida deionizada 50%; SSC 5x; Denhardt's 10x; Tris-HCl 20mM, pH 7,5; SDS 1% e DNA de esperma de salmão denaturado 100 µg/ml. As membranas foram hibridizadas durante a noite a 42°C em: formamida deionizada 50%; SSC 5x; SDS 1%; sulfato de dextran 5%; Denhardt's 2x; Tris-HCl 20 mM, pH 7,5; DNA de esperma de salmão denaturado 100µg/ml e 10^7 - 10^8 cpm/ml da sonda previamente denaturada a 98-100°C por 5 min. As membranas

foram lavadas à temperatura ambiente por 2 vezes, 15 min cada vez, em solução contendo SSC 2x e SDS 0,1%. A seguir, as membranas foram lavadas por 2 vezes a 60°C, 30 min cada lavagem, em solução contendo SSC 0,1x e SDS 0,1%.

As membranas foram expostas por 3-5 h a filmes de raio-X (Hyperfilm MP, Amersham Pharmacia) e quantificadas no programa *CrazyQuant* v. 1.3, ou expostas em IP Bas-MS e quantificadas no *Bio Imaging Analyzer* FLA-3000 (Fujifilm) utilizando o programa *ArrayGauge* v. 1.0.

2.3.4. Análise dos resultados de hibridização

Os valores obtidos para os sinais de hibridização em cada membrana foram normalizados através da divisão do valor da intensidade de cada amostra pela mediana de todas as amostras da membrana (Gill *et al.*, 1999). Após este passo, foi calculada a média das duas repetições de cada experimento. Para obter o valor relativo de expressão dos genes comparados em cada experimento, foi utilizada a fórmula $CER = (\text{Estágio analisado} / (\text{Estágio 1} + \text{Estágio 2} + \dots + \text{Estágio N})) \times N$. N representa o número total de estágios analisados na quantificação da expressão do gene.

Esta fórmula fornece a expressão relativa do gene (CER - coeficiente de expressão relativa). Valores próximos de 1 representam os genes cuja expressão diferencial não foi confirmada. Genes induzidos apresentam CERs maiores que 1 e genes reprimidos apresentam CERs menores que 1. A partir dos CERs, os cDNAs foram agrupados através da utilização dos programas Gene Cluster (<http://rana.stanford.edu/software>) e TreeView v. 1.45 ([www.rana.stanford.edu/software](http://rana.stanford.edu/software)). O padrão de expressão foi representado nas Figuras obtidas através de cores, onde o vermelho representa a expressão mais acentuada e o amarelo ausência de expressão. Expressões intermediárias entre a máxima (vermelha) e a ausência (amarela) são mostradas nas Figuras através da combinação das duas cores (Anexo II).

2.4. Clonagem e seqüenciamento dos cDNAs diferenciais

As bandas que tiveram a expressão diferencial confirmada e apresentaram os maiores CERs foram clonadas no vector pGEM-T *Easy* (Promega), segundo as especificações do fabricante. Pelo menos dois clones de cada banda foram seqüenciados. Nas reações de seqüenciamento utilizou-se o kit DNA *sequencing Big Dye*TM (Perkin Elmer). As reações foram constituídas de 400 ng de DNA, 2 µl de *mix Big Dye* e 1 µl (5 µM) de *primer* (direto ou reverso). As condições de amplificação foram: 96°C (10 seg), 50°C (5 seg) e 60°C (4 min) por um total de 25 ciclos. As amostras, foram precipitadas com isopropanol e mantidas a -20°C até o momento do seqüenciamento. O seqüenciamento foi realizado em um seqüenciador automático ABI PRISMTM 377 (Perkin Elmer).

2.5. Estratégia de seqüenciamento

A estratégia de seqüenciamento foi a seguinte:

- Foram seqüenciados no mínimo 2 sub-clones (um com o *primer* direto e outro com o reverso) de cada banda diferencial.
- As seqüências obtidas com os *primers* direto e reverso foram alinhadas com o auxílio do Clustal W ou C.A.P. (*Contig Assembly Program*).
- Caso fossem observadas discrepâncias no alinhamento dos clones, ou parte da seqüência com leitura duvidosa, mais clones (1-2) da banda diferencial eram seqüenciados.
- Foi confeccionada uma base de dados em formato FASTA com todos os fragmentos seqüenciados. Para detectar possíveis fragmentos redundantes na base de dados, foi utilizado o programa Clustal W.

2.6. Análise das seqüências

O alinhamento das sub-seqüências (direto e reverso) dos clones foi feito através do serviço BCM *Search Launcher* (<http://searchlauncher.bcm.tmc.edu/>) no sub-programa Clustal W v. 1.8 ou no C.A.P. Estas seqüências foram comparadas com seqüências disponíveis no banco de dados *nr* utilizando o serviço BLAST-NCBI (*National Center of Biotechnology*, www.ncbi.nlm.nih.gov:80/entrez/BLAST/) no sub-programa BLASTX.

3. Resultados e Discussão

3.1. *Display* diferencial de mRNA

A técnica de *display* diferencial de mRNA foi utilizada com o objetivo de isolar genes com expressão diferencial durante o desenvolvimento do embrião zigótico. Com este objetivo, foram utilizados nos experimentos RNA isolado de embriões globular (tardio) (E5), cotiledonar (E6) e maduro (E7). Foram utilizados também, RNA isolado do megagametófito (M) e de 3 regiões do embrião maduro: região radicular (ER), cotiledonar (EC) e eixo (parte central do embrião) (EE).

3.1.1. Expressão diferencial de genes durante o desenvolvimento do embrião zigótico de *A. angustifolia*

Com o objetivo de se isolar genes com expressão diferencial durante o desenvolvimento do embrião zigótico de *A. angustifolia*, foi extraído RNA dos estágios de desenvolvimento do embrião: globular (E5), cotiledonar (E6) e maduro (E7). O cDNA correspondente foi sintetizado com os *primers* ancoradores T₁₂VC e T₁₂VG. Na amplificação dos cDNAs foram utilizados os *primers* ancoradores T₁₂VC e T₁₂VG e os *primers* arbitrários: OPJ-06, OPJ-07, OPJ-08 e OPJ-09 (Operon). Um total de 246 bandas diferenciais foram detectadas nos géis de *display* (Anexo 2). Estas bandas foram subdivididas em 5 grupos: (1) bandas com expressão mais acentuada no estágio globular, (2) bandas com expressão mais acentuada no estágio cotiledonar, (3) bandas com expressão mais acentuada no estágio maduro, (4) bandas presentes nos 3 estágios com aumento da expressão no decorrer do desenvolvimento e (5) bandas presentes nos 3 estágios com diminuição da expressão no decorrer do desenvolvimento. Estes dados estão resumidos na Tabela 1.

Das 246 bandas isoladas nos experimentos de *display* diferencial de mRNA (Tabela 1) foram escolhidas 202, com base no tamanho e intensidade de expressão, para serem utilizadas em experimentos de confirmação da expressão diferencial. Estas bandas foram transferidas para membranas de nylon com o auxílio de *slot blot* ou *macroarray*. As membranas foram hibridizadas com sondas de cDNA sintetizados a partir de RNA de cada estágio de desenvolvimento. Os experimentos foram realizados em duplicata. As bandas

Tabela 1. Bandas isoladas através de *display* diferencial de mRNA. Foram isoladas 35 bandas do estágio globular (EG), 26 do estágio cotiledonar (EC), 47 do estágio maduro (EM), 74 que apresentavam um aumento no decorrer do desenvolvimento (A) e 64 que diminuíam no decorrer do desenvolvimento (D).

	Quantidade de bandas	%
EG	35	14.2
EC	26	10.6
EM	47	19.1
A	74	30.1
D	64	26.0
Total	246	100

amplificadas com um mesmo par de *primers* foram colocadas na mesma membrana (Figura 1).

3.1.2. Confirmação da expressão diferencial das bandas isoladas dos géis de *display* diferencial de mRNA através de *slot blot*

Devido ao grande número de bandas diferenciais isoladas, optou-se inicialmente pelo método de *slot blot* para confirmação da expressão diferencial destas bandas. As bandas isoladas com os pares de *primers*, T₁₂VC - OPJ-08, T₁₂VG - OPJ-08, T₁₂VC - OPJ-09, T₁₂VG - OPJ-09 (Anexo 2) foram transferidas para membranas. Após hibridização, foram consideradas confirmadas as bandas que apresentaram no *slot blot* expressão diferencial igual à observada nos experimentos de *display* diferencial. Bandas com comportamento diferente foram descartadas. As bandas com expressão diferencial mais acentuada foram escolhidas para posterior clonagem e sequenciamento. No anexo 2 são mostrados em detalhes os resultados obtidos com os experimentos de *slot blot*. Estes resultados estão resumidos na Tabela 2.

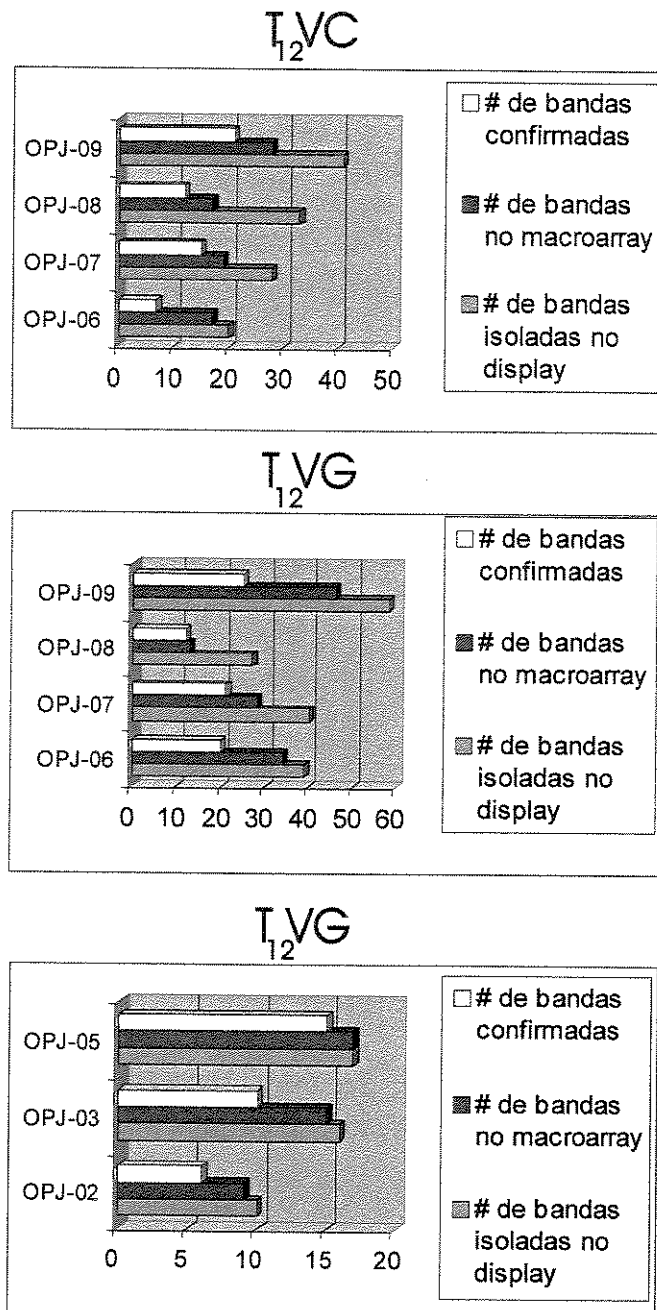


Figura 1. Número total de bandas isoladas dos diferentes estágios de desenvolvimento de *A. angustifolia*, hibridizadas e com expressão diferencial confirmada. Os *primers* utilizados foram: ancoradores $T_{12}VC$ e $T_{12}VG$ e arbitrários OPJ-02, OPJ-03, OPJ-05, OPJ-06, OPJ-07, OPJ-08 e OPJ-09.

Tabela 2. Número de bandas com expressão diferencial confirmada nos experimentos de *slot blot*, com cada combinação de *primers*.

Arbitrários	T ₁₂ VC		T ₁₂ VG	
	# total de bandas	# de bandas confirmadas	# total de bandas	# de bandas confirmadas
OPJ-08	17	12	13	12
OPJ-09	28	21	46	25

3.1.3. Confirmação da expressão diferencial das bandas isoladas dos géis de *display* diferencial de mRNA através de *macroarrays*

A confirmação da expressão diferencial das bandas isoladas dos géis de *display* diferencial de mRNA utilizando *macroarrays* (Nanfei *et al.*, 1997), mostrou-se mais rápida que *slot blot*. Este método permite uma confirmação rápida das bandas diferenciais. As bandas isoladas dos experimentos de *display* diferencial de mRNA com os pares de *primers* T₁₂VC - OPJ-06, T₁₂VG - OPJ-06, T₁₂VC - OPJ-07 e T₁₂VG - OPJ-07 foram transferidas para membranas de *nylon* para confirmação da expressão diferencial. Foram consideradas confirmadas as bandas que apresentaram no *macroarray* a mesma expressão diferencial observada nos géis de *display* diferencial. Bandas com comportamento diferente foram descartadas. As bandas com expressão diferencial mais acentuada foram escolhidas para posterior clonagem e seqüenciamento. No anexo 2 são mostrados em detalhe os resultados obtidos com os experimentos de *macroarray*. Estes resultados estão resumidos na Tabela 3.

3.2. Expressão diferencial de genes no megagametófito e em diferentes regiões do embrião maduro

Com o objetivo de isolar genes com expressão diferencial no megagametófito (M) e no embrião: região cotiledonar (EC), eixo (EE) e região radicular (ER), foram realizados experimentos de *display* diferencial com RNA isolado destes tecidos da semente madura

Tabela 3. Número de bandas com expressão diferencial confirmada nos experimentos de *macroarray*, com cada combinação de *primers*.

Arbitrários	T ₁₂ VC		T ₁₂ VG	
	# total de bandas	# de bandas confirmadas	# total de bandas	# de bandas confirmadas
OPJ-06	17	7	34	20
OPJ-07	19	15	28	21

de *A. angustifolia*. Nestes experimentos, foram utilizados os *primers* ancoradores T₁₂VG e T₁₂VC e os *primers* arbitrários OPF-03, OPJ-02, OPJ-03 e OPJ-05. Foram detectadas 69 bandas diferenciais sendo 21 do megagametófito, 27 do eixo do embrião (EE), 11 da região cotiledonar (EC) e 10 da região radicular (ER). Destas 69 bandas, 56 (21 do megagametófito, 25 do eixo do embrião e 10 da região cotiledonar) foram escolhidas para análise tomando como base o tamanho e a intensidade das bandas.

As bandas foram transferidas para membranas de nylon com o auxílio de um replicador manual. As membranas foram hibridizadas com sondas de cDNAs sintetizados a partir de RNA de cada tecido. O experimento foi realizado em duplicata. As bandas obtidas com um mesmo par de *primers* foram colocadas em uma mesma membrana. Foram consideradas confirmadas as bandas que apresentaram no *macroarray* a mesma expressão diferencial observada nos géis de *display*. Bandas com comportamento diferente foram descartadas. As bandas com expressão diferencial mais acentuada foram escolhidas para clonagem e seqüenciamento. No anexo 2 são mostrados em detalhe os resultados obtidos com os experimentos de *macroarray*. Estes resultados estão resumidos na Tabela 4.

3.3. Clonagem e seqüenciamento

A expressão das bandas confirmadas, que apresentaram os maiores CERs, é mostrada na Figura 2. Assim sendo, a expressão diferencial de 58 bandas (23 do estágio globular, 11 do cotiledonar e 24 do maduro) isoladas do embrião zigótico é mostrada na

Tabela 4. Número de bandas com expressão diferencial confirmada nos experimentos de *macroarray*, com cada combinação de *primers*.

	T ₁₂ VG		T ₁₂ VC	
Arbitrário	# total de bandas	# de bandas confirmadas	# total de bandas	# de bandas confirmadas
OPF-03	-	-	15	6
OPJ-02	9	6	-	-
OPJ-03	15	10	-	-
OPJ-05	17	15	-	-

Figura 2A. A expressão diferencial das bandas do megagametófito (7 bandas), região cotiledonar (3 bandas) e eixo (9 bandas) do embrião maduro é mostrada na Figura 2B. Dentre as bandas apresentadas na Figura 2, 39 foram selecionadas para clonagem e seqüenciamento. Assim, foram escolhidas as bandas J6C-08, J6G-21, J6G-22, J7C-66, J9C-66, J9C-67, J9C-68, J9C-80, J9C-83 e J9C-84 do embrião globular; J6G-36, J7G-93, J7G-95, J7G-96, J7G-97, J8C-12, J8C-16, J8C-17 do embrião no estágio cotiledonar; J6G-53, J6G-57, J7C-104, J7C-105, J7G-113, J7G-114, J7G-115, J7G-120, J7G-123, J7G-125, J7G-126 do embrião maduro; J5G-20, J5G-29, J5G-32 do cotilédone do embrião maduro; J5G-18, J5G-19, J5G-21, J5G-22, J5G-23 do eixo do embrião maduro; e J5G-26, J5G-33 do megagametófito maduro. A expressão diferencial das bandas selecionadas é mostrada na Figura 3.

Após serem clonados, os 39 cDNAs foram seqüenciados e as seqüências obtidas foram comparadas com seqüências disponíveis em bancos de dados utilizando o serviço BLAST - NCBI (*National Center of Biotechnology*, www.ncbi.nlm.nih.gov:80/entrez/BLAST/) e os subprogramas blastn e blastx. Os resultados obtidos estão resumidos na Figura 4. As seqüências estão agrupadas no anexo 3.

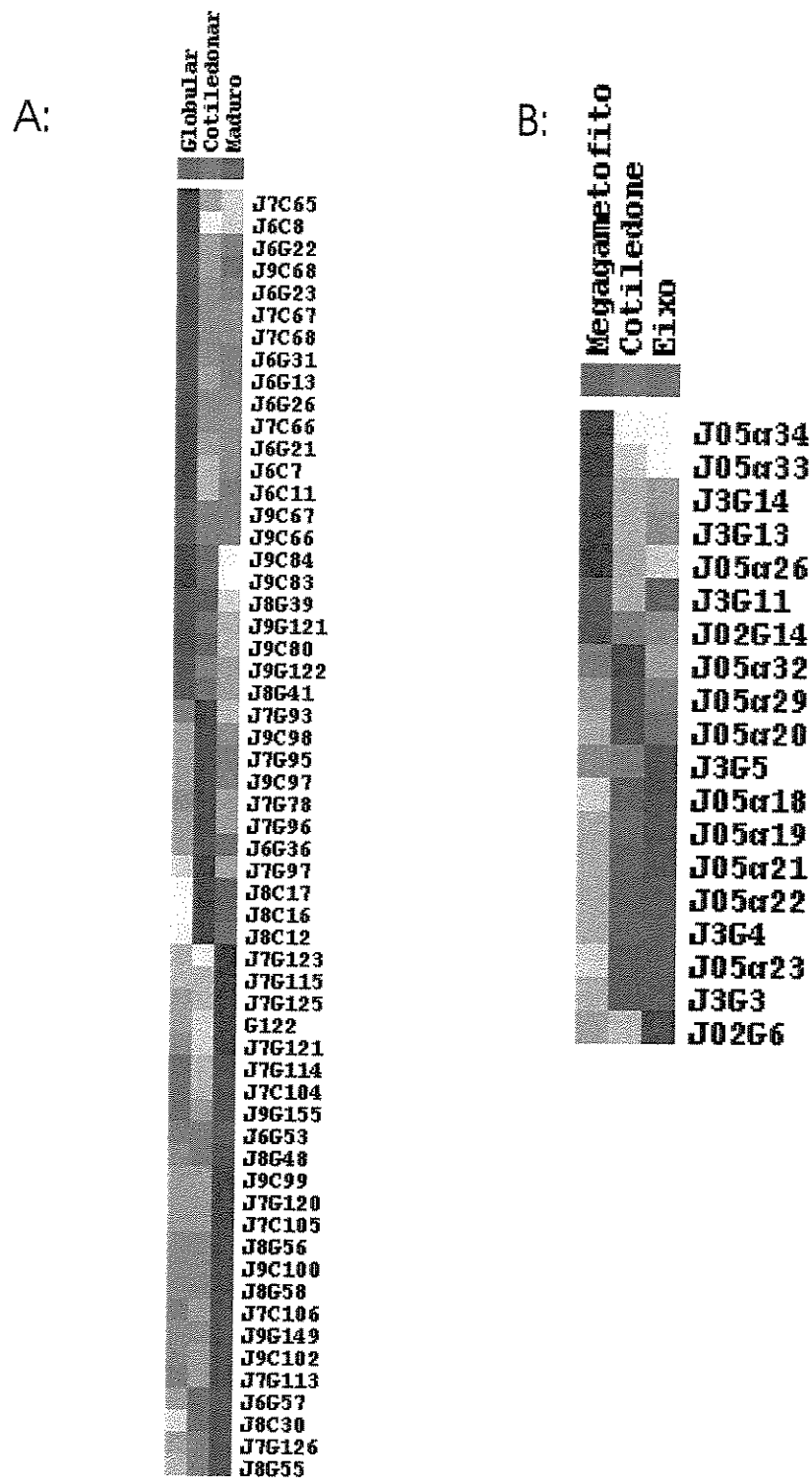


Figura 2. Expressão das bandas diferenciais confirmadas nos experimentos de *slot blot* e *macroarray*. A: bandas com expressão diferencial durante o desenvolvimento do embrião zigótico de *A. angustifolia* (embrião no estágio globular, cotiledonar e maduro). B: bandas com expressão diferencial no megagametófito e embrião zigótico: região cotiledonar e eixo. O agrupamento foi realizado tomando como base o CER dos cDNAs. A expressão mais acentuada é indicada pela cor vermelha, e a ausência de expressão pela cor amarela.

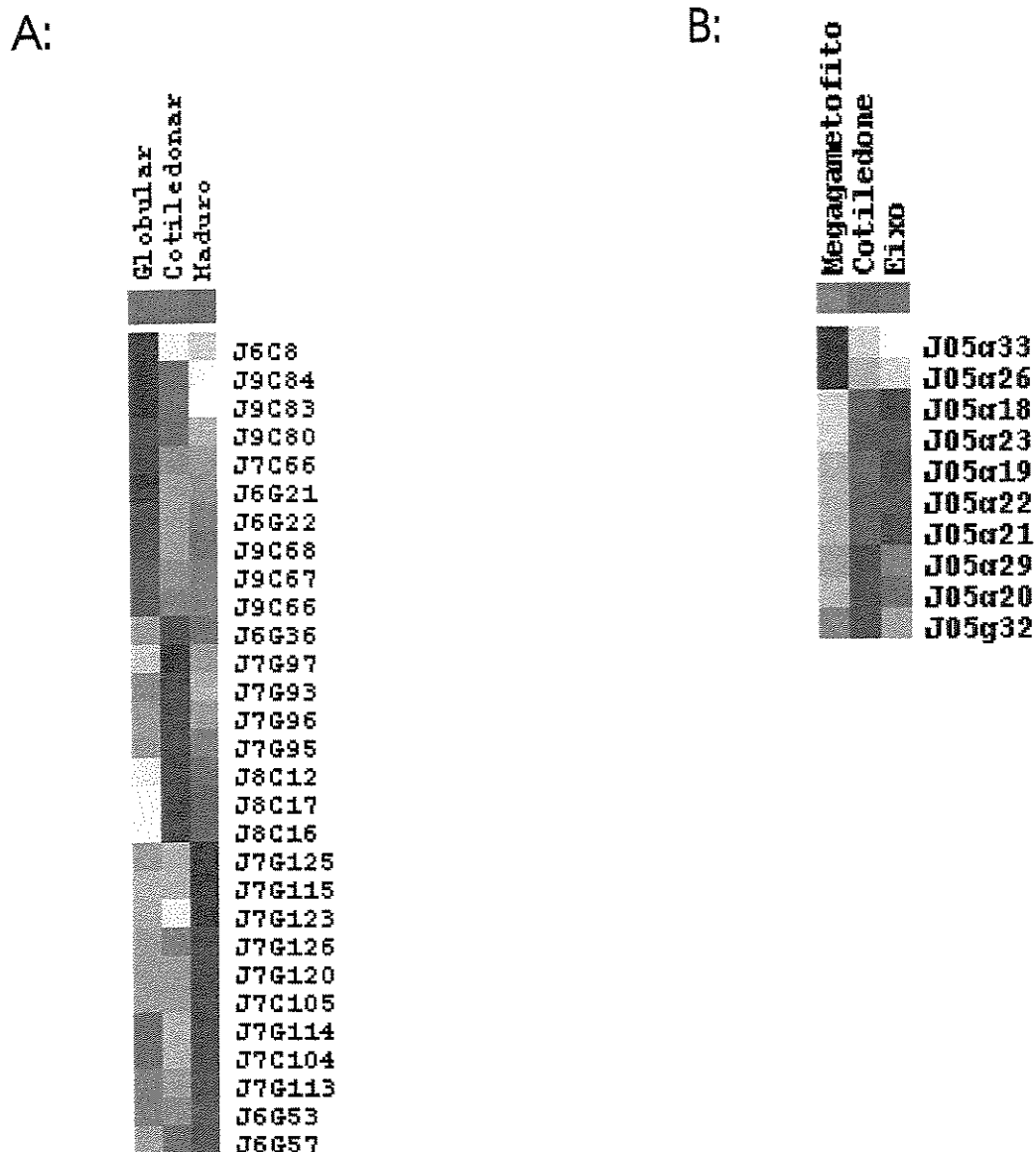


Figura 3. Expressão diferencial das 39 bandas clonadas e seqüenciadas. A: bandas com expressão diferencial durante o desenvolvimento do embrião zigótico de *A. angustifolia* (embrião globular, cotiledonar e maduro). B: bandas com expressão diferencial no megagametófito e embrião zigótico: região cotiledonar e eixo. O agrupamento foi realizado tomando como base o CER dos cDNAs. A expressão mais acentuada é indicada pela cor vermelha, e pela cor amarela, a ausência de expressão.

A:

Fragmento	b.p	EG	CER	Gene similar	Espécie	E	I	S	Possível função
J6C08	301	2,519	0,194	Nenhuma similaridade					
J9C84	162	1,871	1,042	Nenhuma similaridade					
J9C83	171	1,998	1,002	Nenhuma similaridade					
J9C80	252	1,516	1,073	Nenhuma similaridade					
J7C66	325	1,668	0,690	Nenhuma similaridade					
J6G21	265	1,831	0,512	Putative retroelement gypsy/Ty3	A. thaliana	3E-5	34%	52%	?
J6G22	260	1,528	0,559	Nenhuma similaridade					
J9C68	363	1,459	0,565	Nenhuma similaridade					
J9C67	629	1,317	0,796	Putative LRR receptor protein kinase	A. thaliana	4E-27	93%	92%	Sinalização
J9C66	558	1,245	0,839	Nenhuma similaridade significativa					
J6G36	493	0,618	1,406	Nenhuma similaridade					
J7G97	330	0,328	2,146	Vicilin-like protein	Picea glauca	1E-8	38%	55%	Reserva
J7G93	538	0,795	1,805	Nenhuma similaridade					
J7G96	335	0,607	1,763	Nenhuma similaridade					
J7G95	421	0,605	1,598	Nenhuma similaridade significativa					
J8C12	235	0,154	1,804	Nenhuma similaridade					
J8C16	181	0,061	1,801	Nenhuma similaridade					
J7G115	445	0,408	0,419	Nenhuma similaridade					
J7G123	297	0,458	0,160	Putative SP75					
J7G126	242	0,515	0,801	Nenhuma similaridade	A. thaliana	1E-50	64%	81%	Fator de transcrição
J7C105	238	0,622	0,609	Putative retroelement Osvaldo	Drosophila buzzatii	2E-13	43%	65%	?
J7G114	511	1,041	0,375	Nenhuma similaridade					
J7C104	260	0,992	0,370	Nenhuma similaridade					
J7G113	511	0,817	0,528	Putative retroelement gag/pol	A thaliana	1E-20	53%	69%	?
J6G53	290	0,895	0,794	Nenhuma similaridade significativa					
J6G57	159	0,533	1,035	Vicilin-like protein	Picea glauca	2E-8	38%	55%	Reserva
				Nenhuma similaridade					

B:

Fragmento	b.p	MM	CER	Gene similar	Espécie	E	I	S	Possível função
J05g33	305	2,565	0,323	Nenhuma similaridade					
J05g26	296	2,209	0,460	Putative CCR4-associated factor	A. thaliana	5E-32	78%	87%	Fator de transcrição
F03C14	411	2,260	0	Nenhuma similaridade					
F03C2	331	2,379	0,702	Nenhuma similaridade					
J05g29	374	0,587	1,485	Nenhuma similaridade significativa					
J05g20	668	0,478	1,496	Outher membrane lipoprotein-like	A. thaliana	5E-59	77%	86%	Reserva ?
J05g32	382	0,880	1,606	Nenhuma similaridade					
F03C6	405	0	2,108	Nenhuma similaridade significativa					
J05g18	762	0,230	1,116	Nenhuma similaridade					
J05g23	495	0,311	1,328	Nenhuma similaridade					
J05g19	763	0,453	1,055	Nenhuma similaridade					
J05g22	495	0,468	1,185	Nenhuma similaridade					
J05g21	523	0,498	1,110	Nenhuma similaridade					
F03C8b	262	0,185	1,612	Ribonucleic acid-binding protein	Zea mays	2E-33	69%	87%	Fator de transcrição?
				Nenhuma similaridade significativa					

b.p- pares de bases

CER- Coeficiente de expressão relativa

EG- Embrião cotiledonar

E- Embrião cotiledonar

I- Identidade

S- Similaridade

E- E-value

FEM- Megagametófito maduro

EG- Embrião cotiledonar

FEM- Megagametófito maduro

I- Identidade

S- Similaridade

E- E-value

Figura 4. Resultados obtidos com o sequenciamento dos cDNAs clonados. A. Resumo dos resultados obtidos com a comparação das seqüências dos cDNAs isolados do embrião nos estágios globular, cotiledonar e maduro em bancos de dados. B. Resumo dos resultados obtidos com a comparação das seqüências dos cDNAs isolados do megagametófito, cotiledone e eixo do embrião maduro em bancos de dados.

3.4. cDNAs que apresentaram similaridade com seqüências presentes em bancos de dados

Os resultados obtidos com a comparação das seqüências dos cDNAs com seqüências disponíveis em bancos de dados foram divididos em três grupos: resultados com *E-value* maiores de 1 – seqüências sem similaridade, resultados com *E-value* entre 1 e $1e^{-5}$ – seqüências com similaridade não significativa e resultados com *E-value* menor de $1e^{-5}$ – seqüências que apresentaram similaridade significativa com seqüências disponíveis em bancos de dados (Figura 4).

3.4.1. cDNAs com expressão mais acentuada no embrião maduro

J7G-115:

O cDNA J7G-115 foi amplificado com o *primer* OPJ-07 e possui 445 pb. A análise da seqüência deduzida de aminoácidos deste fragmento apresentou 81% de similaridade sobre 147 resíduos da proteína putativa Spt5 de *Arabidopsis thaliana*, também descrita para *Drosophila*, *Mus musculus*, *Homo sapiens*, e outros (Figura 5). As seqüências com similaridades relevantes são mostradas na Tabela 5.

```

J7g115: 1  LSTLLANRKKGHFLKGDVAVVVRGDLKNLMGWVEKVEDDNIHVKPKKEGIQGTCLICKENE 60
          LSTL ANRKKGHF+KGDVAVV++GDLKNL GWVEKV+++N+ ++ + KG+ L E E
Spt5 1: 414 LSTLFANRKKGHFMKGDVAVIVIKGDLKNLKGWVEKVDEENVLIRSEVKGLPDPLAVNERE 473

J7g115: 182 LCKFFKPGDHVKVVSNGTHQATGMVVKVENHVLIIIVSDTTKEDIRVFADNVVESSEITSG 361
          LCK+F+PG+HVKKVVSNGTH+GATGMVVKV+ HV RVFAD+VVESSE+T+G
Spt5 1: 474 LCKYFEPGNHVKVVSNGTHGATGMVVKVDQHV-----RVFADHVVESSEVTTG 521

J7g115: 61 VTKLGDYELHDLVQLDHTSFGIIVRVE 87
          VTK+GDYELHDLV LD+ SFG+I+R+E
Spt5 1: 522 VTKIGDYELHDLVLLDNLSFGVIIRLE 548

```

Figura 5. Comparação entre a seqüência deduzida de aminoácidos do cDNA J7G-115 e a seqüência de aminoácidos da proteína putativa Spt5 de *A. thaliana*. A proteína possui uma seqüência de 990 aminoácidos. As letras mostram os resíduos idênticos e as cruzeis mostram os similares.

Tabela 5. Proteínas que apresentaram similaridade com o cDNA J7G-115, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Arabidopsis thaliana</i>	-hypothetical protein T28D5.40 (pir T14189). -putative Spt5 like protein (AL109819)	$1e^{-50}$	(95/147) 64%	(120/147) 81%
<i>Drosophila melanogaster</i>	-Spt5 gene product(AE003796) -Dspt5 (AF222864)	$6e^{-23}$	(53/136) 38%	(87/136) 63%
<i>Mus musculus</i>	-chromatin structural protein homolog Supt5hp (U88539)	$1e^{-22}$	(57/140) 40%	(87/140) 61%
<i>Homo sapiens</i>	-DSIF p160 (AB000516)	$2e^{-22}$	(56/140) 40%	(87/140) 62%

A proteína Spt5, em *Saccharomyces cerevisiae* forma um complexo com Spt4. Este complexo está envolvido em processos de repressão da transcrição, modulando possivelmente a estrutura da cromatina. Foi demonstrado que a proteína homóloga funcional descrita para *Homo sapiens*, é fosforilada reversivelmente durante a mitose (Stachora *et al.*, 1997). Este homólogo, chamado de DSIF, é composto por subunidades de 160 kDa (p160) e 14 kDa (p14), homólogas a Spt5 e Spt4 respectivamente (Wada *et al.*, 1998). O complexo DSIF associa-se à RNA Polimerase II estimulando “*in vitro*” o processo de transcrição em condições limitantes de nucleotídeos. Este complexo, também reprime o processo de transcrição na falta de outros fatores de supressão (Wada *et al.*, 1998). Estudos com linhagens mutantes em *Saccharomyces cerevisiae* mostraram que Spt4 e Spt5 são fundamentais na transcrição, interagindo com a RNA Pol II “*in vivo*” (Hartzog *et al.*, 1998). Foi demonstrado que Spt5 interage com a RNA Pol II no domínio C-terminal (CTD) não fosforilado. O CTD fosforilado afeta esta interação. O complexo Spt5-RNA Pol II interage diretamente com a enzima guanilil transferase, estimulando indiretamente o processo de adição do cap no mRNA (Wen & Shatkin, 1999).

O possível mecanismo mediante ao qual Spt5 regula de maneira positiva ou negativa o processo de transcrição foi estudado em detalhe em humanos. Foi observado, que Spt5 contém diferentes domínios incluindo um N-terminal ácido, quatro repetições KOW homólogas ao regulador de transcrição NusG de *E. coli* (Wada *et al.*, 1998), e dois domínios C-terminais com elementos repetitivos denominados de CTR1 e CTR2 (Ivanov *et*

al., 2000). Estes elementos repetitivos são ricos em resíduos de serina e treonina, potenciais sítios de fosforilação para as quinases (Ivanov *et al.*, 2000).

O domínio CTR1 está envolvido na inibição da transcrição mediante o contato com o inibidor DRB (5,6-dichoro-1- β -D-ribofuranosylbenzimidazole), e ao mesmo tempo na ativação mediante contato com a proteína Tat (Ivanov *et al.*, 2000). Este processo é totalmente reversível. Ao mesmo tempo, CTR1 e CTR2 podem ser fosforilados, afetando o processo de interação com a RNA Pol II, ou alterando a afinidade do complexo por outros fatores de transcrição.

Embora o cDNA J7G-115 apresente alta similaridade com o fator de transcrição SPT5, só estudos posteriores podem atribuir uma função específica para este cDNA no desenvolvimento embrionário de *A. angustifolia*.

J7C-104:

O fragmento J7C-104, amplificado com os *primers* OPJ-07 e T₁₂VC, possui 260 pb. A análise da sequência deduzida de aminoácidos deste cDNA revelou 69% de similaridade sobre 84 resíduos da poliproteína gag/pol de um retroelemento de *Arabidopsis thaliana*, também descrito para *Vicia faba*, *Oryza sativa*, e outros (Figura 6). As sequências que apresentaram similaridades mais relevantes são mostradas na Tabela 6.

```
J7C104 : 87 PLDTYAGHEMFSLMDGFSGYNQI*INPEDQHKIAFTTPWGTFCYKVVPFGLKNASATYQR 27
          P+++ G EM S MD FSGYNQI +NP+DQ K +F T GT+ YKV+PFGLKN ATYQ
Gag-pol: 439 PIESTTGQEMLSFMDAFSGYNQIMNPDDEKTSFITERGTYYYKVMFPGLKNVGATYQL 498

J7C104 : 26 AMTYIFHDYMHKIIEDYVDDLLAK 3
          + +F D + K +E Y+DD+L K
gag-pol: 499 LVNIMFKDLLGKTMEVYIDDMLVK 522
```

Figura 6. Comparação entre a sequência deduzida de aminoácidos do cDNA J7C-104 e a sequência de aminoácidos da poliproteína gag/pol de um retroelemento de *A. thaliana*. As letras mostram os resíduos idênticos e as cruzes mostram os similares.

Os retroelementos (retrotransposons e retrovirus) têm dois tipos de genes em comum: genes *gag* que codificam proteínas estruturais da partícula viral, e genes *pol* que codificam proteínas catalíticas requeridas para a replicação. Entre os retrotransposons mais

Tabela 6. Proteínas que apresentaram similaridade com o cDNA J7C-104, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Arabidopsis thaliana</i>	-putative retroelement gag/pol polyprotein (AC007045)	$1e^{-20}$	(45/84) 53%	(59/84) 69%
<i>Vicia faba</i>	-reverse transcriptase-like protein (AB007466)	$7e^{-18}$	(41/79) 51%	(54/79) 67%
<i>Oryza sativa</i>	-Similar to Sorghum bicolor Gypsy-Ty3 type retrotransposon (AP001366)	$4e^{-15}$	(37/85) 43%	(54/85) 63%

estudados, e semelhantes aos retrovirus, estão os do grupo LTR de *Drosophila*. Entre eles, a família mais numerosa e dispersa é a dos elementos *gypsy-like* (Marín & Lloréns, 2000).

A estrutura interna dos elementos *Ty3-gypsy* é muito similar a dos retrovirus e são encontrados na seguinte ordem: genes pol – protease – RT - ribonuclease H -integrase. Foram descritos elementos da família *Ty3-gypsy* que possuíam uma ORF (*open reading frame*) que poderia codificar para a proteína *env*. Na taxonomia atual, os retroelementos LTR são classificados em duas famílias: Pseudoviridae (subgrupo *Ty1/Copia*) e Metaviridae (subgrupo *Ty3/Gypsy*). Metaviridae é dividida de acordo com a presença do gene *env* (gênero *Errantivirus*) ou ausência (gênero *Metavirus*) (Marín & Lloréns, 2000).

Os retroelementos da família *Ty3-gypsy* estão distribuídos nas angiospermas e gimnospermas. Eles foram descritos em uma grande quantidade de organismos. Em coníferas, já foi descrito para *Pinus radiata* o elemento Ifg7 (Marín & Lloréns, 2000).

J7G-126:

O cDNA J7G-126 foi amplificado com o *primer* OPJ-07 e possui 242 pb. A análise da sequência deduzida de aminoácidos deste cDNA em bancos de dados revelou 65% de similaridade sobre 66 resíduos do retrotransposon Buzzatii-osvaldo de *Drosophila* também descrito para *Zea mays*, *Oryza sativa*, e outros (Figura 7). As sequências que apresentaram similaridades mais relevantes são mostradas na Tabela 7.

Os cDNAs J7C-104 e J7C-66 também apresentaram similaridade com os retroelementos (Figura 4).

J7G126: 81 PLDTLTRKGVQFKWTEEHQEAFFGLKG*LVKAPILVYPNWEKEFYIHVDASRFAIGCALA 21
 PL+ L +KGV+++WTEEHQ AFE +K L ++P+L P++ K F + DAS + +G L
 Buzzat: 3859 PLNALLKKGVKWEWTEEHQRAFETVKAKLTESPVLACPDFSKPFCLQTDASNYGLGAILT 4038

J7G126: 20 QEDQYG 15
 Q + G
 Buzzat: 4039 QTSEEG 4056

Figura 7: Comparação entre a sequência deduzida de aminoácidos do fragmento J7G-126 e a sequência de aminoácidos do retroelemento *Buzzatii-osvaldo* de *Drosophila*. O retroelemento possui uma sequência de 9085 nucleotídeos. As letras mostram os resíduos idênticos e as cruzes mostram os similares.

Tabela 7. Proteínas que apresentaram similaridade com o cDNA J7G-126, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Drosophila buzzatii</i>	- <i>Osvaldo retrotransposon</i> (gi 4539019)	$2e^{-13}$	(29/66) 43%	(44/66) 65%
<i>Oryza sativa</i>	- <i>genomic DNA, chromosome 4, BAC clone: H0512B01</i>	$8e^{-13}$	(28/62) 45%	(44/62) 70%
<i>Zea mays</i>	- <i>Ty3/gypsy-type retrotransposon reverse transcriptase/integrase polyprotein pseudogene gi 2655227</i>	$2e^{-11}$	(28/62) 45%	(41/62) 65%
<i>Sorghum bicolor</i>	- <i>22 kDa kafirin cluster</i> (AF061282)	$4e^{-10}$	(27/62) 43%	(37/62) 59%

3.4.2. cDNAs com expressão mais acentuada no embrião cotiledonar

J6G-36:

O fragmento J6G-36 foi amplificado com o *primer* OPJ-06 e possui 493 pb. A análise da sequência deduzida de aminoácidos deste fragmento em bancos de dados revelou 55% desimilaridade sobre 58 resíduos de proteínas similares as vicilinas 7S de *Picea glauca*, *Vicia faba*, e outros (Figura 8). As sequências que apresentaram similaridades mais relevantes são mostradas na Tabela 8.

A presença de proteínas similares as vicilinas 7S de reserva no embrião de *A. angustifolia* foi descrita no Capítulo 1. O cDNA J6G-53 apresentou a mesma sequência do cDNA J6G-36, com uma pequena variação na parte do N'-terminal. Esta variação leva o cDNA J6G-53 a ser 133 bases menor que o J6G-36 (Figura 9). As sequências da parte N'-

J6G-36: 79 RHGEEGRNPYVFKEEQERRLATDAGEIRAVPLFREFSTLVRDLENYELNFFHMKPDAF 139
 R EE NPYVF + R +++AGEIRA+P F E S L+ + + F MKP
 Picea : 45 RREEEREENPYVFHSDSFRTRASSEAGEIRALPNFGEVSELLGGIRKFRVTFFIEMKPKTV 104

J6G-36: 140 MRHHYSGA 147

M HY A

Picea : 105 MLPHYIDA 112

Figura 8. Comparação entre a sequência deduzida de aminoácidos do cDNA J6G-36 e a sequência de aminoácidos da *vicilin-like storage protein* de *Picea glauca*. A proteína possui uma sequência de 450 aminoácidos. As letras mostram os resíduos idênticos e as cruzes mostram os similares.

Tabela 8. Proteínas que apresentaram similaridade com o cDNA J6G-36, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Picea glauca</i>	<i>vicilin-like storage protein</i>	$5e^{-08}$	(28/58) 41%	(37/58) 54%
<i>Vicia faba</i>	<i>putative sucrose binding protein</i>	$1e^{-07}$	(28/72) 38%	(42/72) 57%
<i>Pisum sativum</i>	<i>62K sucrose-binding protein homolog</i>	$7e^{-07}$	(28/92) 30%	(48/92) 51%
<i>Cucurbita maxima</i>	<i>protein PV100</i>	$2e^{-05}$	(24/87) 27%	(44/87) 49%
<i>Glycine max</i>	<i>Sucrose binding protein (SBP)</i>	$2e^{-04}$	(21/59) 35%	(35/59) 58%

terminal deduzidas destes cDNAs, mostraram ser idênticas às já sequenciadas no Capítulo 1. Isto sugere que os cDNAs J6G-36 e 53 são fragmentos dos transcritos correspondentes a AE-21 e AM-20 respectivamente (Figura 9).

Embora o cDNA J6G-36 seja expresso principalmente no embrião cotiledonar, o cDNA J6G- 53 é expresso principalmente no embrião maduro (Figura 4). Com o objetivo de caracterizar a proteína de reserva de *A. angustifolia*, o cDNA completo de J6G-36 foi isolado por RACE e sequenciado. Estes resultados são apresentados no Capítulo 3.

```

>J6G-36      TCGTTCCGCATCGTTCCGCATCGTTCCGCAGAGGTAATGCATCCACGAATTATCTTTC
>J6G-36      TTCATGAATTTCATATTCCTTTTGTTCACACTGACTTTCAGAGTTGATCAATGCAGGTGAGGTGGTGAACGTG

      M A N G R F P L L Y L V L V L V G F V C A I R
>J6G-36      GCTCATGGCGAATGGTAGGTTCCCGTTGCTTTACTTGGTGCTTGTCTTGTAGGTTTTGTGTGTGAATCCGT

> AE-21      C N E V D H Y K V H R H G E E E G R N P Y V F K
>J6G-36      TGCAATGAGGTTGATCACTACAAGGTTTCATCGACATGGTGAAGAAGAGGGTCGAAATCCATACGTTTAAAGG
      *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
>J5G-53      CGTTCCGCAGAGGTGAGGTGGTGAACGTGCTCATGGCGAAGATGAGGGTCGAAATCCATACGTTTAAAGG
> AM-20      R S A E V R W S N C A H G E D E G R N P Y V F K

      E E E Q E R R L A T D A G E I R A V P L F R E F S
>J6G-36      AGGAGGAGCAAGAGAGGAGACTCGCTACAGATGCAGGCGAGATCCGAGCGGTTCCCTCTTCAGAGAGTTCTC
      *****
>J5G-53      AGGAGGAGCAAGAGAGGAGACTCGCTACAGATGCAGGCGAGATCCGAGCGGTTCCCTCTTCAGAGAGTTCTC
      E E E Q E R R L A T D A G E I R A V P L F R E F S

      T L V R D L E N Y E L N F F H M K P D A F M R H
>J6G-36      CACCCTCGTAAGAGATCTCGAGAACTATGAACATCAATTTCTTTCACATGAAGCCCGATGCGTTCATGCGTCAT
      *****
>J5G-53      CACCCTCGTAAGAGATCTCGAGAACTATGAACATCAATTTCTTTCACATGAAGCCCGATGCGTTCATGCGTCAT
      T L V R D L E N Y E L N F F H M K P D A F M R H

      H Y S G A D H L S L F Y K E R Q E F S V C G T
>J6G-36      CATTACAGCGGTGCAGATCATTATCGTTGTTCTACAAGGAAAGGCAAGAATTCAGTGTGTGCGGAACGA
      *****
>J5G-53      CATTACAGCGGTGCAGATCATTATCGTTGTTCTACAAGGAAAGGCAAGAATTCAGTGTGTGCGGAACGA
      H Y S G A D H L S L F Y K E R Q E F S V C G T

```

Figura 9. Sequências dos clones J5G36 e 53 alinhadas no programa Clustal W e as respectivas seqüências deduzidas de aminoácidos. Em negrito são mostradas as seqüências do N^o-terminal das proteínas AE-21 e AM-20 obtidas no Capítulo 1. As diferenças na parte N^o-terminal das seqüências, justificam a presença de dois polipeptídeos com seqüências N^o-terminais similares, mas não idênticas como já foi visto no Capítulo 1.

3.4.3. cDNAs com expressão mais acentuada no embrião globular

J9C-68:

O fragmento J9C-68 foi amplificado com os *primers* OPJ-09 e T₁₂VC e possui 363 pb. A análise da seqüência deduzida de aminoácidos deste fragmento em bancos de dados revelou 82% de similaridade sobre 79 aminoácidos da *putative LRR receptor protein kinase* de *Arabidopsis thaliana* (figura 10). As seqüências que apresentaram similaridades mais relevantes são mostradas na Tabela 10.

J9C-68: 6 DTLSLDWITRLKIAIGAAEGLSFLHHECSPPLVHRDVQASSILLDDKFEVRLGSLSEAC 66
 D L+SLDWITRLKIA+GAAEGLS+LHHECSPPLVHRDVQASSILLDDKFEVRLGSLSEA
 LRR : 621 DGLKSLDWITRLKIALGAAEGLSYLHHECSPPLVHRDVQASSILLDDKFEVRLGSLSEAY 680

J9C-68: 67 SQEGEVHQSVIARLLRLSQ 85
 +Q G+ +QS I+RLLRL Q
 LRR : 681 AQ-GDAYQSRISRLRLRPQ 698

Figura 10. Comparação entre a sequência deduzida de aminoácidos do cDNA J9C-68 e a sequência de aminoácidos da *putative LRR receptor protein kinase* de *Arabidopsis thaliana*. A proteína possui uma sequência de 925 aminoácidos. As letras mostram os resíduos idênticos e as cruzes mostram os similares.

Tabela 10. Proteínas que apresentaram similaridade com o cDNA J9C-68, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Arabidopsis thaliana</i>	<i>putative LRR receptor protein kinase</i>	$8e^{-30}$	(66/79) 83%	(73/79) 91%
<i>Arabidopsis thaliana</i>	<i>receptor protein kinase-like protein</i>	$4e^{-29}$	(61/94) 64%	(75/94) 78%
<i>Oryza sativa</i>	<i>ESTs similar to receptor protein kinase, ERECTA</i>	$1e^{-13}$	(33/65) 7350%	(48/65) 73%
<i>Arabidopsis thaliana</i>	ERECTA	$2e^{-13}$	(32/50) 64%	(43/50) 86%

Os receptores do tipo LRR (*leucine-rich repeat*) fazem parte do grupo dos receptores proteína-quinase com domínios transmembrana (Walker, 1994). A proteína ERECTA (Torii *et al.*, 1996) foi descrita como um receptor transmembrana com propriedades de proteína-quinase no domínio C'-terminal. Segundo Torii *et al.* (1996) esta proteína estaria envolvida no processo de alongamento celular, mediante interação do seu domínio extracelular com fatores responsáveis pela ativação do domínio quinase citoplasmático da proteína. Vários genes similares a esta proteína foram descritos em angiospermas contudo, em gimnospermas genes similares só foram descritos em *Pinus silvestris* (AJ250467, não publicado segundo NCBI, 08/2001).

A similaridade apresentada pelo cDNA J9C-68 foi com o domínio proteína-quinase. Assim sendo, não é possível determinar se a proteína apresenta a mesma similaridade com os domínios transmembrana e com o domínio de ligação extracelular de ERECTA. Estas

características só serão determinadas com o seqüenciamento do gene completo. Estes resultados permitirão a atribuição de uma função para esta proteína no desenvolvimento do embrião de *A. angustifolia*.

3.4.4. cDNAs com expressão mais acentuada no megagametófito maduro

J05G-26:

O fragmento J50G-26 foi amplificado com os *primers* OPJ-05 e T₁₂VG e possui 296 pb. A análise da seqüência deduzida de aminoácidos deste cDNA em bancos de dados mostrou 87% de similaridade sobre 65 aminoácidos do *putative CCR4-associated factor* de *A. thaliana* na porção C-terminal (figura 11). As seqüências que apresentaram similaridades mais relevantes são mostradas na Tabela 11.

```

J5G26: 1  LHGGLNKLAEILLDVKRIGVCHQAGSDSLTSTFRKLREGFFNGSTEKYAGVLYGLGXE 60
          LHGGLNKLAEILLDV+R+G+CHQAGSDSLTSTFRKL+E FF GS EKY+GVLYGLG E
CCR4: 188 LHGGLNKLAEILLDVERVGICHQAGSDSLTSTFRKLQENFFIGSMEKYSGLVLYGLGVEN 247

J5g26: 61 SENGH 65
          + H
CCR4: 248 GQIVH 252

```

Figura 11. Comparação entre a seqüência deduzida de aminoácidos do cDNA J50G-26 e a seqüência de aminoácidos do *putative CCR4-associated factor* de *A. thaliana*. A proteína de *A. thaliana* possui uma seqüência de 252 aminoácidos. As letras mostram os resíduos idênticos e as cruzes mostram os similares.

Tabela 11. Proteínas que apresentaram similaridade com o cDNA J05G-26, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Arabidopsis thaliana</i>	- <i>putative CCR4-associated factor</i> (AC006223)	5e ⁻³²	(51/65) 78%	(57/65) 87%
<i>Mus musculus</i>	- <i>catabolite repressor protein (CCR4)-associative factor 1</i>	3e ⁻¹²	(37/68) 54%	(46/68) 67%
<i>Saccharomyces cerevisiae</i>	-POP2 homolog (NP035265.1)	3e ⁻¹²	(37/68) 54%	(46/68) 67%

O gene *CAF1* (*CCR4 associated factor1*) foi isolado em vários organismos. A proteína CAF1 junto com a proteína Lex A funciona como ativador da transcrição (Draper *et al.*, 1995). Em levedura, este gene foi clonado e denominado de *POP2*. Ele interage com o CCR4 (*carbon catabolite repressor protein*) e está fisicamente associado a este complexo. O POP2 é um fator de transcrição que pode desempenhar vários papéis incluindo repressão da expressão gênica. Da mesma maneira que o CCR4, POP2 é requerido em diversos processos incluindo os relacionados ao metabolismo de carbono (Draper *et al.*, 1995).

O cDNA J5G-26 foi escolhido para seqüenciamento completo. Estes resultados são apresentados no Capítulo 3.

3.4.5. cDNAs com expressão mais acentuada no cotilédono do embrião maduro

J05G-20:

O fragmento J05G-20 foi amplificado com os *primers* OPJ-05 e T₁₂VG e possui 668 pb. A análise da seqüência deduzida de aminoácidos deste fragmento em bancos de dados revelou 86% de similaridade sobre 75 aminoácidos da *outer membrane lipoprotein-like* de *A. thailana* (Figura 12). Este cDNA pode representar o gene completo desta lipoproteína para *A. angustifolia*. As seqüências que apresentaram similaridades mais relevantes são mostradas na Tabela 12.

```

J5G20: 5  SEQKSLEVVKGLDLERYMGRWYEIASMPSRFQPKNGVNTRATYALNDDGTVHVLNETWVE 65
          +E+K +EVVKGL++ERYMGRWYEIAS PSRFQPKNGV+TRATY LN DGT+HVLNETW
lipo: 2  TEKKEMEYVVKGLNVERYMGRWYEIASFPSRFQPKNGVDTRATYTLNPDGTIHLNETWSN 61

J5G20: 66  GKRSSIEGSAYKGRP 80  104 GNYWVMLLDQDYQ 116 118 ALIGEPSLTSRLVL 131
          GKR IEGSAYK P          G+YWV+ +D DYQ          ALIG+PS + L +L
lipo: 62  GKRGFIEGSAYKADP 76  100 GDYWVLYIDPDYQ 112 114 ALIGQPSRSYLWIL 127

J5G20: 131 LSRQPQLDEEIIYNRLLEKAQQQGYEVGRLRKTSQTAVPDETE-APKD-KGFWWIKALHG 181
          LSR Q++EE Y +L+EKA ++GY++ +L KT Q+ P E+ AP+D KG WW K+L G
lipo: 127 LSRTAQMEETKYLVEKAVEEGYDISKLHKTPQSDTPPESNTAPEDSKGVWWFKSLFG 185

```

Figura 12. Comparação entre a seqüência deduzida de aminoácidos do cDNA J5G-20 e a seqüência de aminoácidos da *outer membrane lipoprotein-like* de *A. thailana*. A proteína possui uma seqüência de 185 aminoácidos. As letras mostram os resíduos idênticos e as cruzes mostram os similares.

Tabela 12. Proteínas que apresentaram similaridade com o cDNA J5G-20, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Arabidopsis thaliana</i>	-outer membrane lipoprotein-like (AB024029)	5e ⁻⁵⁹	(58/75) 77%	(65/75) 86%
<i>Citrobacter freundii</i>	-outer membrane lipoprotein (pir I40710) - lipocalin precursor (U21727)	9e ⁻¹⁰	(29/69) 42%	(39/69) 56%
<i>Homo sapiens</i>	- apolipoprotein D precursor (M16696)	4e ⁻⁰⁷	(25/59) 42%	(38/59) 64%
<i>Escherichia coli</i>	- outer membrane lipoprotein (lipocalin) (AE005648)	2e ⁻⁰⁶	(23/49) 46%	(31/49) 62%

A proteína descrita como “outer membrane lipoprotein” ou apoproteína D faz parte da família das lipocalinas que são proteínas extracelulares que se ligam e transportam pequenos compostos orgânicos hidrofóbicos (Bishop *et al.*, 1995).

As lipocalinas humanas são amplamente utilizadas como marcadores de processos inflamatórios, câncer, problemas metabólicos de fígado e outros (Xu & Venge, 2000). Estas proteínas também foram descritas em bactérias. Elas são proteínas pequenas com estrutura terciária formada por planos β em forma de canudo, o que permite a ligação de moléculas hidrofóbicas como ácidos graxos e lipídeos (Bishop *et al.*, 1995).

No cotilédone do embrião maduro de *A. angustifolia* estas proteínas podem funcionar como transportadoras de lipídeos. A possível função da lipocalina de *A. angustifolia* será discutida no Capítulo 3.

3.4.6. cDNAs com expressão mais acentuada no eixo do embrião maduro

J5G-21:

O cDNA J5G-21 foi amplificado com os *primers* OPJ-05 e T₁₂VG e possui 668 pb. A análise da sequência deduzida de aminoácidos deste cDNA em bancos de dados revelou 87% de similaridade com 101 aminoácidos de uma *nucleic acid-binding protein* de *Zea mays* (Figura 13). As sequências que apresentaram similaridades mais relevantes são mostradas na Tabela 13.

J5G-21: 1 PWGTRVERPT-QVGS AHKVYVGNLPWQADDNSLLQLFSEHGKVL EAKVVYDRETGRSRGF 59
P G+RV+RP Q G + ++YVGNLPWQ DD+ L++LFSEHGKV++A+VVDRETGRSRGF
NBP : 203 PRGSRVDRPFRQSGPSLRIVVGNLPWQVDDSR LVELFSEHGKVVDARVVYDRETGRSRGF 262

J5G-21: 60 GFVTYSSENEVND AIAALDGADMDGRPLRVNVAEDRPRRQF 100
GFVT +S++E++DAIAALDG +DGR LRVNVAE+RPRR F
NBP : 263 GFVTMASQDELDDAIAALDGQSLDGRALRVNVAEERPRRGF 303

Figura 13. Comparação entre a sequência deduzida de aminoácidos do cDNA J5G-21 e a sequência de aminoácidos da *nucleic acid-binding protein* de *Zea mays*. A proteína possui uma sequência de 303 aminoácidos. As letras mostram os resíduos idênticos e as cruzes mostram os similares.

Tabela 13. Proteínas que apresentaram similaridade com o cDNA J5G-21, organismos nas quais elas foram descritas, *E-value* e a porcentagem de resíduos idênticos e similares.

Organismo	Descrição	E	Identidade	Similaridade
<i>Zea mays</i>	- <i>nucleic acid-binding protein</i> - <i>Maize</i>	$2.2e^{-33}$	(70/101) 59%	(88/101) 87%
<i>Hordeum vulgare</i>	- <i>cp31AHv protein</i> - <i>barley</i>	$5.9e^{-33}$	(59/101) 58%	(87/101) 85%
<i>Triticum aestivum</i>	- <i>Ps16 protein</i> - <i>wheat</i>	$2.0e^{-32}$	(58/101) 57%	(87/101) 85%
<i>Hordeum vulgare</i>	- <i>nucleic acid-binding protein</i> - <i>barley</i>	$4.8e^{-31}$	(55/99) 55%	(84/99) 84%

Segundo Cook & Walker (1992) a NBP de *Zea mays* é uma proteína de aproximadamente 30 kDa codificada pelo genoma nuclear e envolvida na regulação da expressão gênica no cloroplasto. A NBP apresentou afinidade *in vitro* por ácidos nucleicos. A sequência de aminoácidos da NBP apresentou similaridade com o domínio RNP-CS (*ribonucluprotein consensus sequence*). Todas estas proteínas possuem peso molecular aparente entre 29 e 31 kDa (Schuster & Gruissem, 1991). A expressão da NBP é tecido-específica e dependente de luz. (Cook & Walker, 1992). O domínio RPN-CS ou *rrm* (Pfam 00076) foi descrito para proteínas de várias angiospermas. Contudo, é a primeira vez que ele é encontrado em gimnospermas.

A expressão deste cDNA no estágio maduro do embrião, sugere que ele está envolvido na regulação da expressão de genes específicos da maturação do embrião (genes *lea*).

4. Referências Bibliográficas

- Bishop R.E., Penfold S.S., Frost L.S., Holtje J.V., Weiner J.H. (1995) Stationary phase expression of a novel *Escherichia coli* outer membrane lipoprotein and its relationship with mammalian apolipoprotein D. Implications for the origin of lipocalins. J. Biol. Chem. 270 (39): 23097.
- Cairney J., Nanfei X., Pullman G. S., Vincent T., Ciavatta V. T., Johns B. (1999) Natural and somatic embryo development in loblolly pine. Applied Biochem. and Biotech. 79: 5.
- Chatthai M., Misra S. (1998) Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziessi* [Mirb] Franco. Planta 206: 138.
- Cook W.B., Walker J.C. (1992) Identification of maize nucleic acid-binding protein (NBP) belonging to a family of nuclear-encoded chloroplast proteins. Nucleic Acids Res. 20 (2): 359.
- Dong J.Z., Dunstan D.I. (1996) Expression of abundant mRNAs during somatic embryogenesis of white spruce [*Picea glauca* (Moench) Voss]. Planta 199: 459.
- Dong J.Z., Dunstan D.I. (1997) Characterization of cDNA representing five abscisic acid-responsive genes associated with somatic embryogenesis in *Picea glauca*, and their responses to abscisic acid stereostructure. Planta 203: 448.
- Dong J.Z., Dunstan D.I. (1999) Cloning and characterization of six embryogenesis-associated cDNAs from somatic embryos of *Picea glauca* and their comparative expression during zygotic embryogenesis. Plant Mol. Biol. 39 (4): 859.
- Draper M. P., Salvatore C., Denis C. L. (1995) Identification of a mouse protein whose homolog in *Saccharomyces cerevisiae* is a component of the CCR4 transcriptional regulatory complex. Mol. Cell Biol. 15 (7): 3487.
- Gill R. T., Valdes J. J., Bentley W. E. (1999) Reverse transcription-PCR differential display analysis of *Escherichia coli* global gene regulation in response to heat shock. App. and Environ. Microbiology 65 (12): 5386.
- Girke T., Todd J., Ruuska S., White J., Benning C., Ohlrogge J. (2000) Microarray analysis of developing Arabidopsis seeds. Plant Physiol. 124: 1570.

- Goldberg R. B., Barker S. J., Perez-Grau L. (1989) Regulation of gene expression during plant embryogenesis. *Cell* 56: 149.
- 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.
- Hartzog G. A., Wada T., Handa H., Winston F. (1998) Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes Dev.* 12 (3): 357.
- Ivanov D., Kwak Y. T., Guo J., Gaynor R. B. (2000) Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol. Cell Biol.* 20 (9): 2970.
- Jarvis S. B., Taylor M. A., MacLeod M.R., Davies H. V. (1996) Cloning and characterization of the cDNA clones of three genes that are differentially expressed during dormancy breakage in the seeds of Douglas fir (*Pseudotsuga menziesii*). *J. Plant Physiol.* 147: 559.
- Jain S., Gupta P., Newton R. (1995) Somatic embryogenesis in woody plants. Kluwer Academic Publishers.
- Krassowski M. J., Owens J. N. (1993) The ultrastructural and histochemical post-fertilization megagametophyte and zygotic embryo development of interior spruce. *Can. J. Bot.* 71: 98.
- Leal I., Misra S. (1993a) Developmental gene expression in conifer embryogenesis and germination. III. Analysis of crystalloid protein mRNAs and desiccation protein mRNAs in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss). *Plant Sci.* 88: 25.
- Leal I., Misra S. (1993b) Molecular cloning and characterization of a legumin-like storage protein cDNA of Douglas fir seeds. *Plant Mol. Biol.* 21: 709.
- Marín I, Lloréns C. (2000) Ty3/Gypsy retrotransposons: description of new *Arabidopsis thaliana* elements and evolutionary perspectives derived from comparative genomic data. *Mol. Biol. Evol.* 17 (7): 1040.
- Misra S. (1994) Conifer zygotic embryogenesis, somatic embryogenesis and seed germination: Biochemical and molecular advances. *Seed Science Res.* 4: 357.

- Misra S., Green M. J. (1991) Developmental gene expression in conifer embryogenesis and germination. II. Crystalloid protein synthesis in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss. Plant Sci. 78: 61.
- Mullen R. T., King J. E., Gifford D. J. (1996) Changes in mRNA populations during loblolly pine (*Pinus taeda*) seed stratification, germination and post-germinative growth. Physiol. Plant. 97: 545.
- Nanfei X., Jhons B., Pullman G., 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: 377.
- Neutelings G., Domon J. M., Membré N., Bernier F., Meyer Y., Davia A., David H. (1998) Characterization of a germin-like protein gene expressed in somatic and zygotic embryos of pine (*Pinus caribaea* Morelet). Plant Mol. Biol. 38: 1190.
- Newton C. H., Flinn B. S., Sutton B. C. S. (1992) Vicilin-like seed storage proteins in the gymnosperm interior spruce (*Picea glauca/engelmannii*). Plant Mol. Biol. 20: 315.
- Preccott A., Martin C. (1987) Rapid method for quantitative assessment of levels of specific mRNAs. Plant Mol. Biol. Rep. 4: 219.
- Sabala I., Elfstrand M., Farbos I., Clapham D., Arnold S.V. (2000) Tissue – specific expression of Pa18, a putative lipid transfer protein gene, during embryo development in Norway spruce (*Picea abies*). Plant Mol. Biol. 42: 461.
- Schuster G., Gruissem W. (1991) Chloroplast mRNA 3' end processing requires a nuclear-encoded RNA-binding protein. EMBO J. 10 (6): 1493.
- Stachora A. A., Schafer R. E., Pohlmeier M., Maier G., Ponstingl H. (1997) Human Supt5h protein, a putative modulator of chromatin structure, is reversibly phosphorylated in mitosis. FEBS Lett. 409 (1): 74.
- Torii K. U., Misukawa N., Oosumi T., Matsuura Y., Yokoyama R., Whittier R. F., Komeda Y. (1996) The Arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. The Plant Cell 8: 735.
- Wada T., Takagi T., Yamaguchi Y., Ferdous A., Imai T., Hirose S., Sugimoto S., Yano K., Hartzog G. A., Winston F., Buratowski S., Handa H. (1998) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. 12 (3): 343.

- Walker J. C. (1994) Structure and function of the receptor-like protein kinases of higher plants. *Plant Mol. Biol.* 26: 1599.
- Wen Y., Shatkin A. J. (1999) Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev.* 13 (14): 1774.
- White J. A., Newman T., Girke T., Focus N., Martinez de Ilarduya O., Jaworski J. G., Ohlrogge J., Benning C. (2000) A new set of *Arabidopsis* ESTs from developing seeds: The metabolic pathway from carbohydrates to seed oil. *Plant Physiol.* 124: 1582.
- Wobus, U.; Weber H. (1999) Seed maturation: Genetic programmes and control signals *Curr. Opin. in Plant Biology* 2: 33.
- Xu N., John B., Pullman G., 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. Reporter* 15: 377.
- Xu S., Venge P. (2000) Lipocalins as biochemical markers of disease. *Biochim Biophys. Acta* 1482 (1-2): 298.
- Zimmerman J. (1993) Somatic embryogenesis: A model for early development in higher plants. *The Plant Cell* 5: 1411.

ANEXO 2

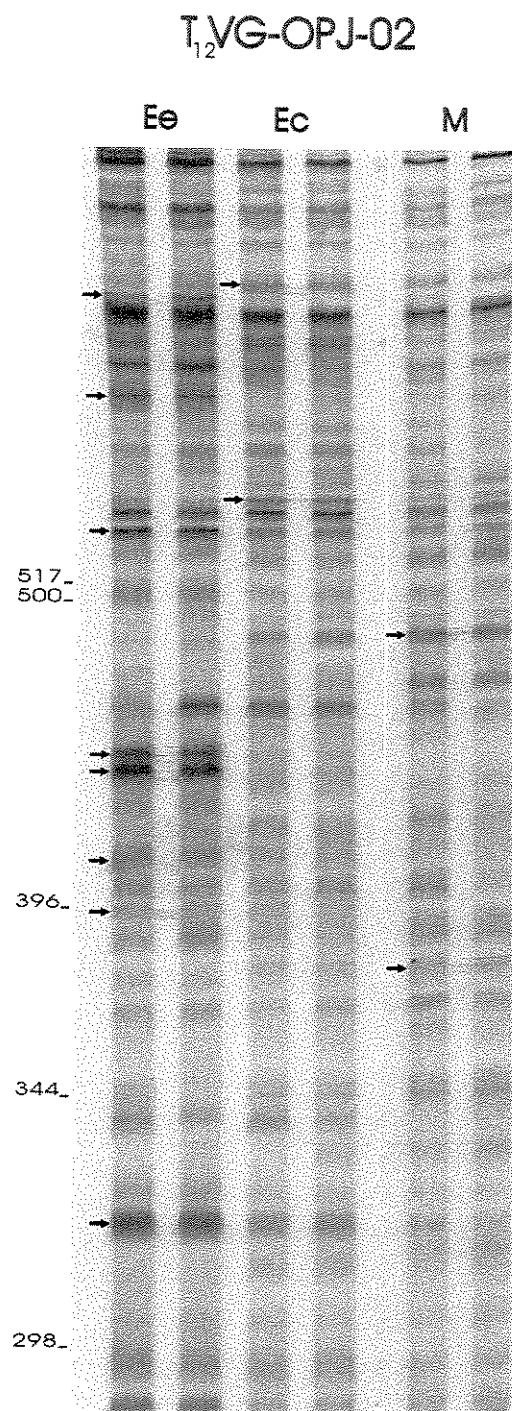


Figura 1. *Display* diferencial de mRNA utilizando-se como *primer* ancorador o T_{12} VG, e como *primer* arbitrário o OPJ-02 (Operon). No experimento foi utilizado RNA do megagametófito (M) e do embrião: região cotiledonar (Ec) e eixo (Ee). As bandas diferenciais estão indicadas na figura (→).

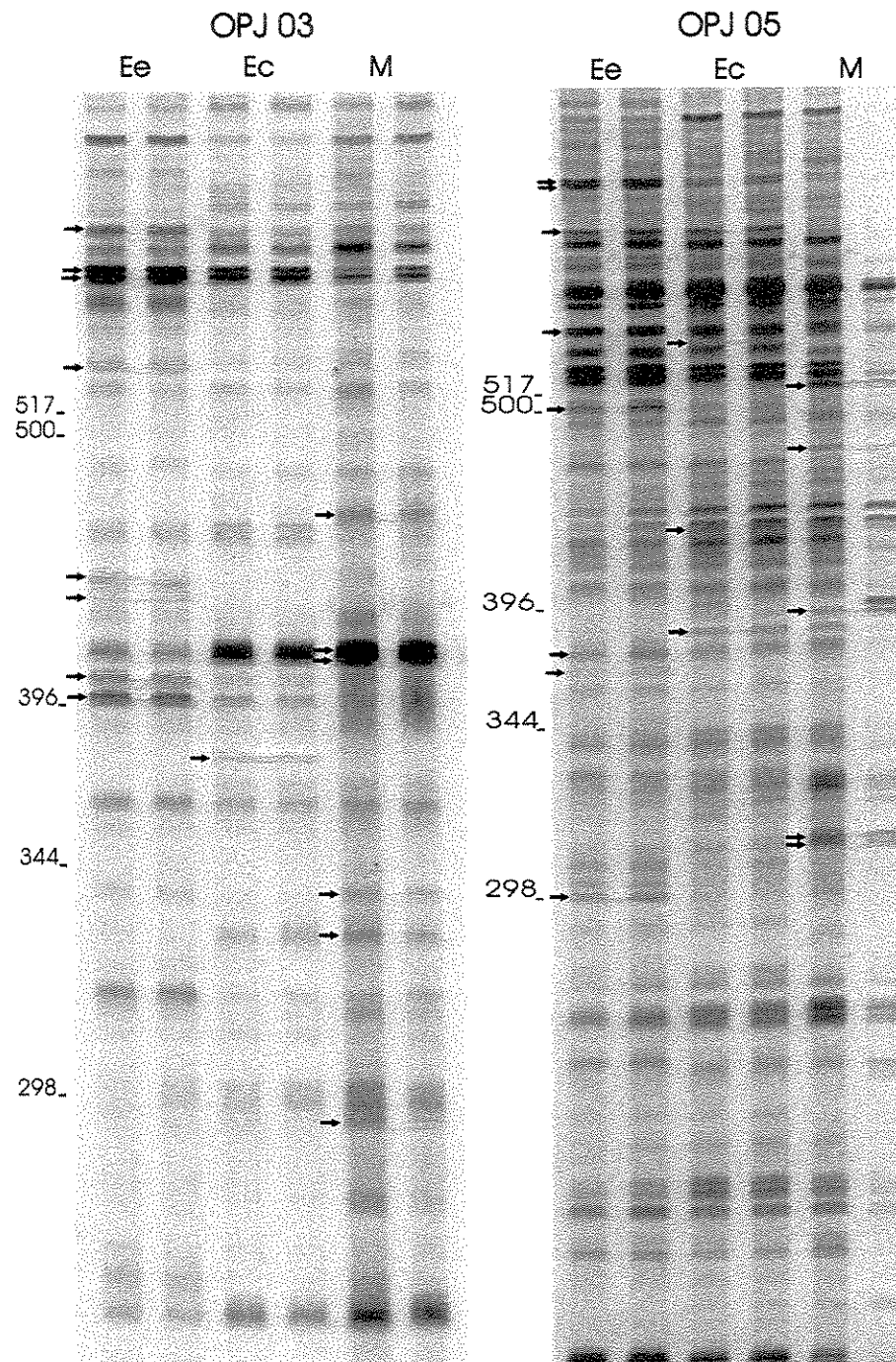


Figura 2. *Display* diferencial de mRNA utilizando-se como *primer* o ancorador T₁₂VG, e como *primers* arbitrários o OPJ-03 ou OPJ-05 (Operon). No experimento foi utilizado RNA do megagametófito (M) e do embrião: região cotiledonar (EC) e eixo (EE). As bandas diferenciais estão indicadas na figura (→).

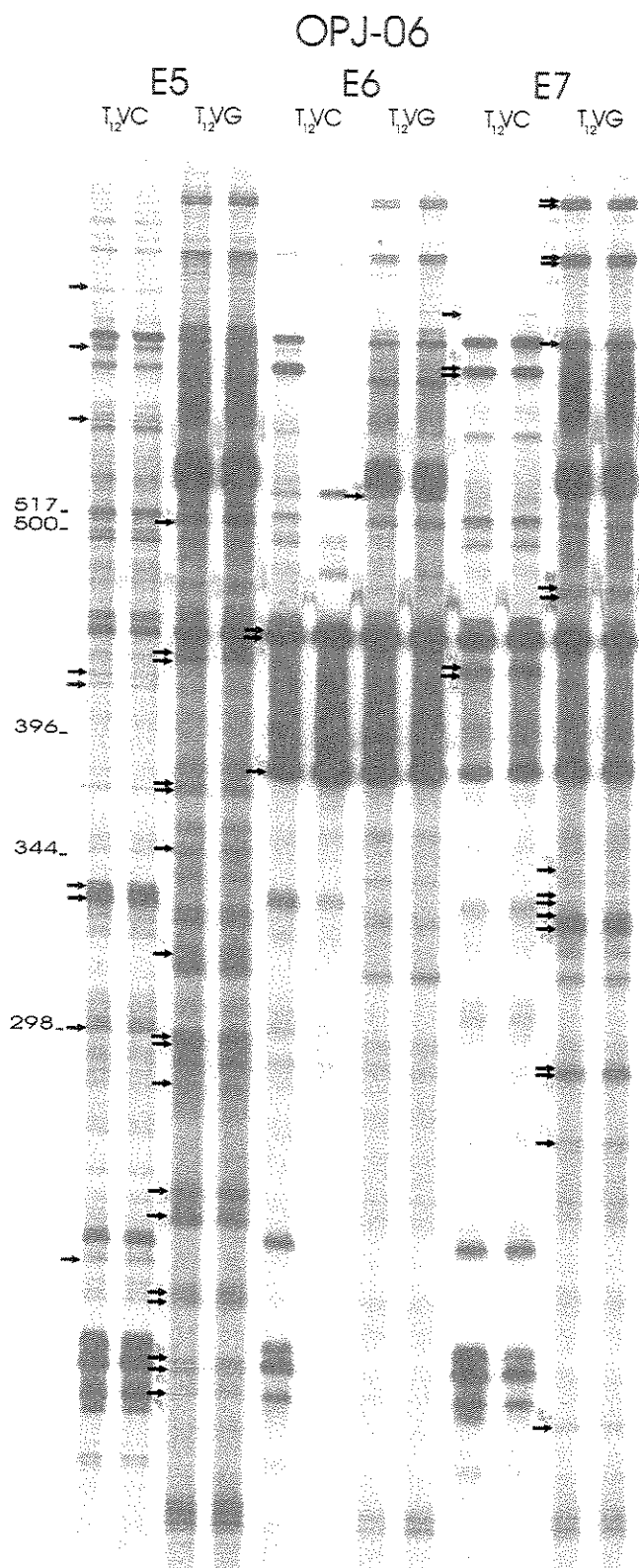


Figura 3. *Display* diferencial de mRNA utilizando-se como *primers* ancoradores $T_{12}VC$ ou $T_{12}VG$ e como *primer* arbitrário o OPJ-06 (Operon). No experimento foi utilizado RNA do embrião em diferentes estágios de desenvolvimento: estágio globular (E5), estágio cotiledonar (E6) e estágio maduro (E7). As bandas diferenciais estão indicadas na figura (→).

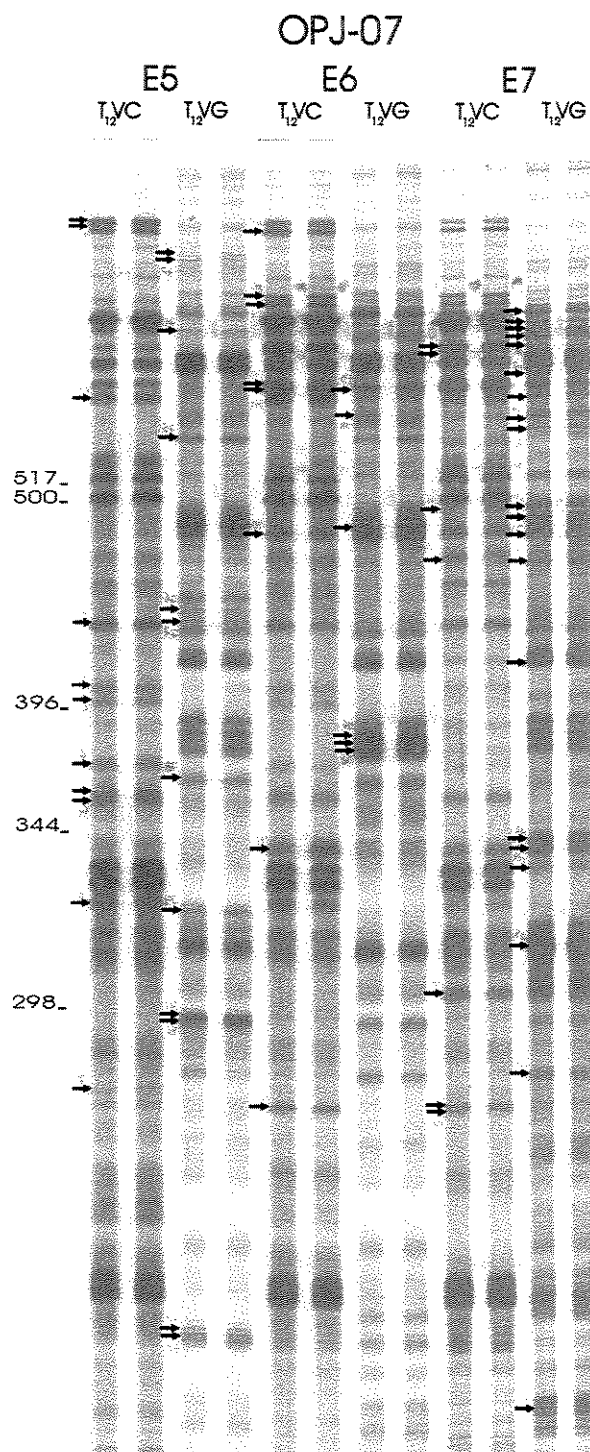


Figura 4. *Display* diferencial de mRNA utilizando-se como *primers* ancoradores T₁₂VC ou T₁₂VG e como *primer* arbitrário o OPJ-07 (Operon). No experimento foi utilizado RNA do embrião em diferentes estágios de desenvolvimento: estágio globular (E5), estágio cotiledonar (E6) e estágio maduro (E7). As bandas diferenciais estão indicadas na figura (→).

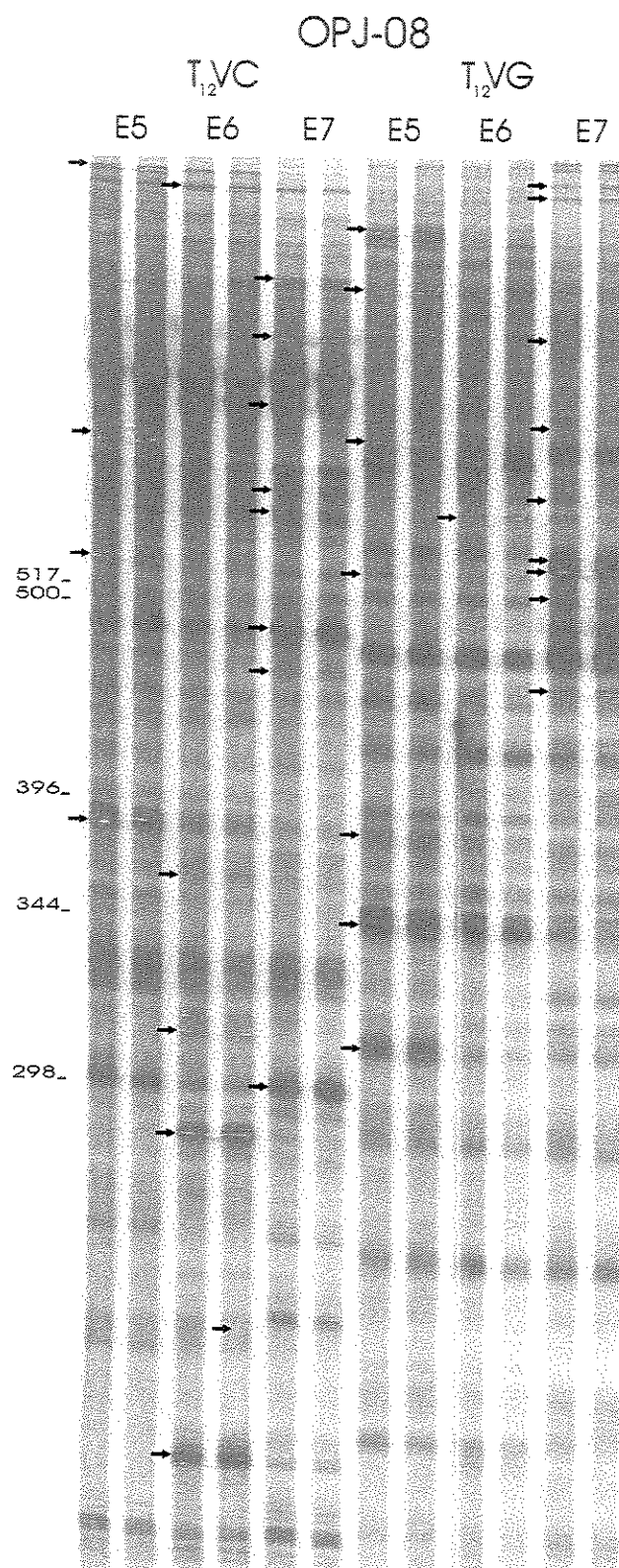


Figura 5. *Display* diferencial de mRNA utilizando-se como *primers* ancoradores T₁₂VC ou T₁₂VG e como *primer* arbitrário o OPJ-08 (Operon). No experimento foi utilizado RNA do embrião em diferentes estágios de desenvolvimento: estágio globular (E5), estágio cotiledonar (E6) e estágio maduro (E7). As bandas diferenciais estão indicadas na figura (→).

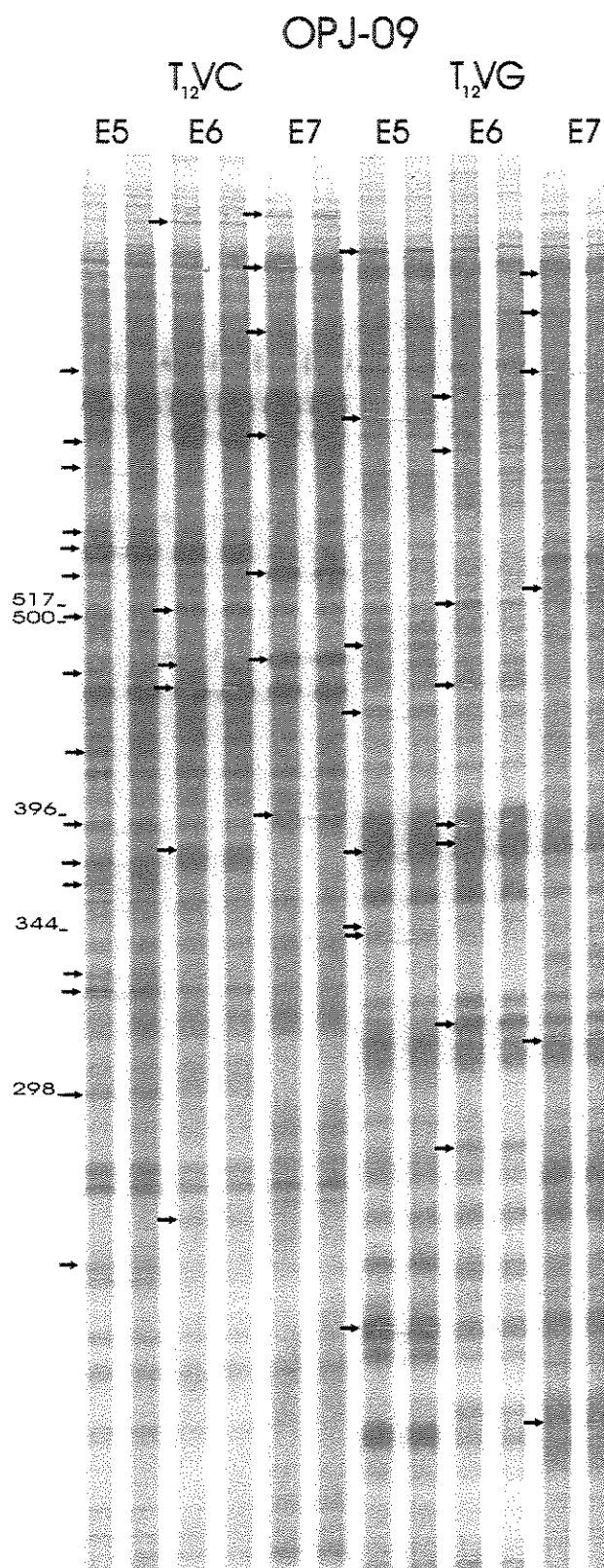


Figura 6. *Display* diferencial de mRNA utilizando-se como *primers* os ancoradores T₁₂VC e T₁₂VG, e como *primer* arbitrário o OPJ-09 (Operon). No experimento foi utilizado RNA do embrião em diferentes estágios de desenvolvimento: estágio globular (E5), estágio cotiledonar (E6) e estágio maduro (E7). As bandas diferenciais estão indicadas na figura (→).

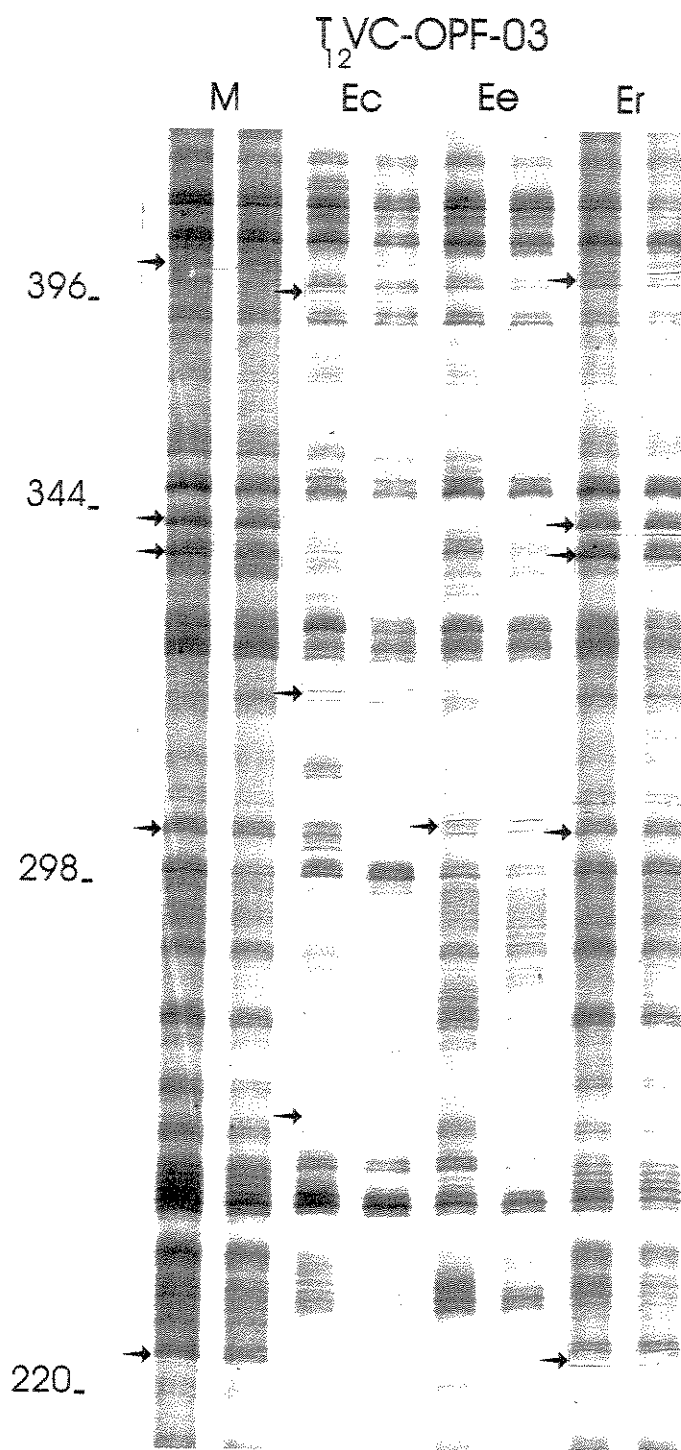
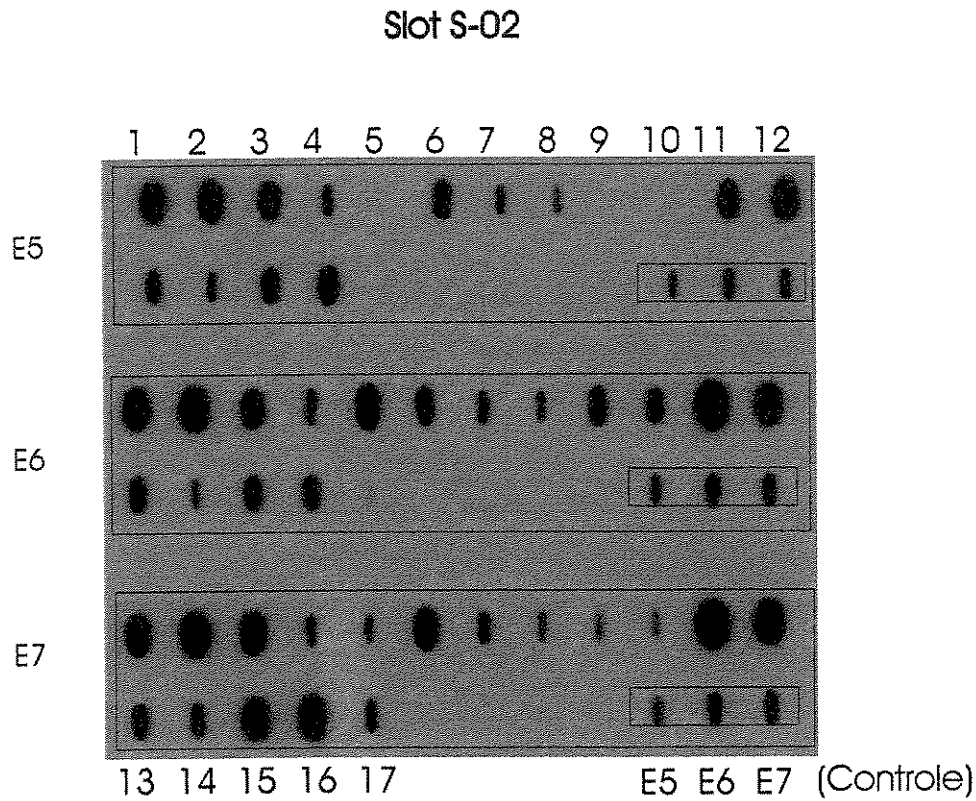


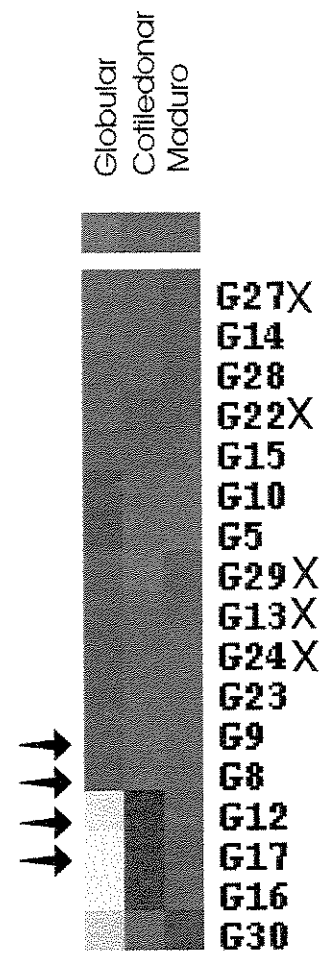
Figura 7. *Display* diferencial de mRNA utilizando-se como *primer* ancorador o $T_{12}VC$, e como *primer* arbitrário o OPF-03 (Operon). No experimento foi utilizado RNA do megagametófito (M) e do embrião: região cotiledonar (Ec), radicular (Er) e eixo (Ee). As bandas diferenciais estão indicadas na figura (→).

OPJ-08-T₁₂VC

A:



C:



B:

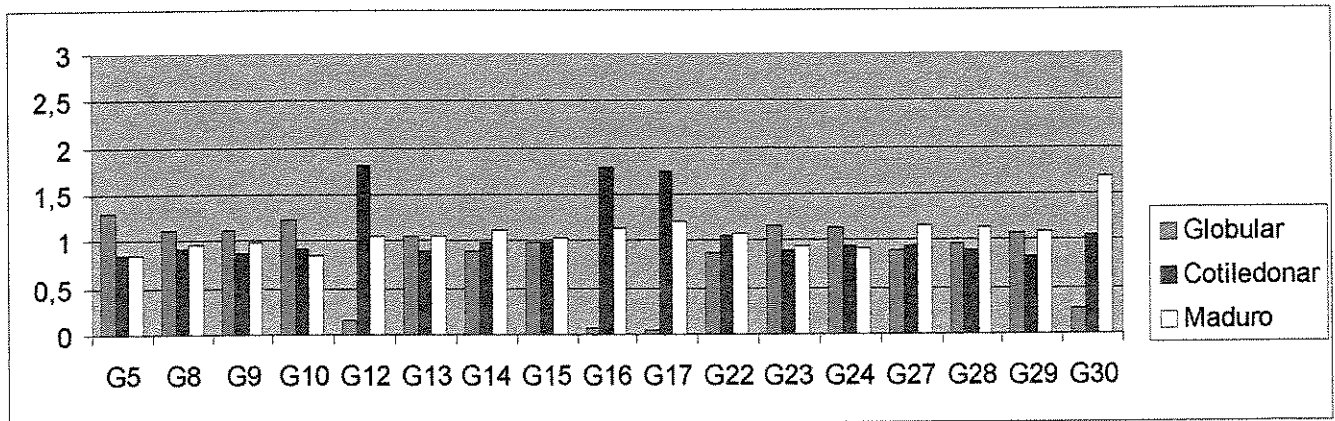
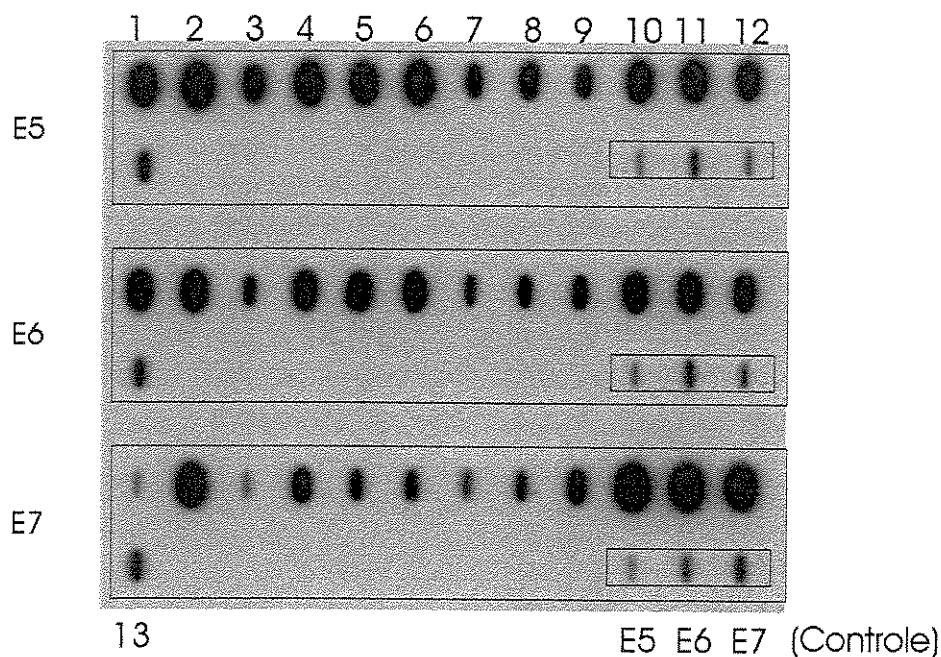


Figura 8. Confirmação da expressão diferencial, por *slot blot*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VC – OPJ08. A. Membrana hibridizada com sonda de cDNA de embrião em estágio globular (E5), embrião em estágio cotiledonar (E6) e estágio maduro (E7). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

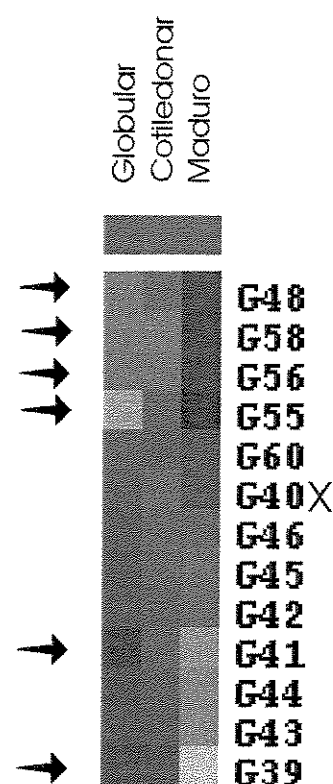
OPJ-08-T₁₂VG

A:

Slot S-03



C:



B:

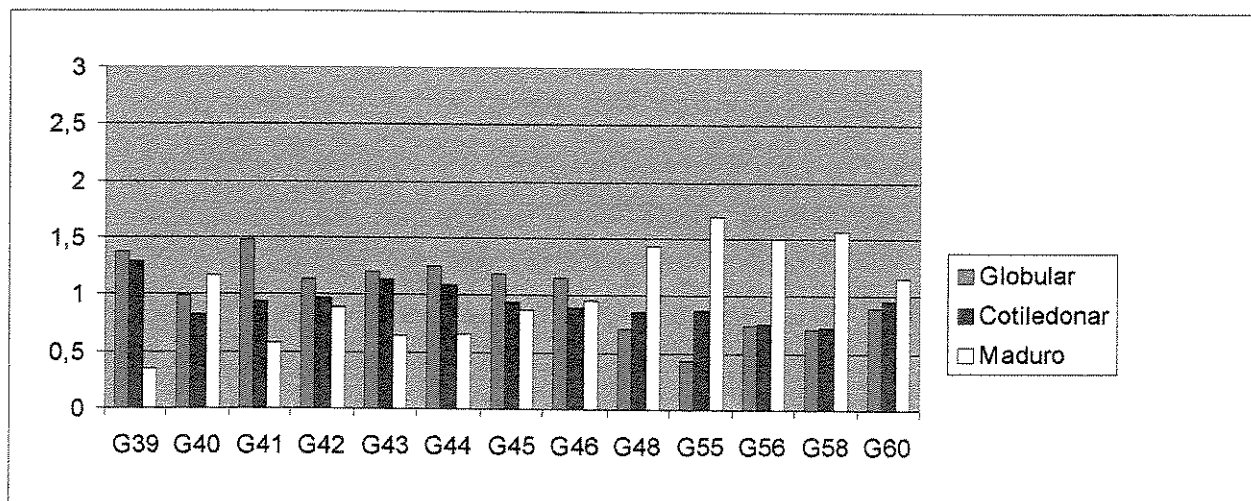
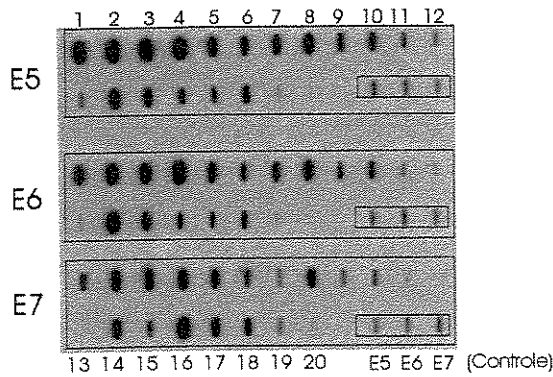


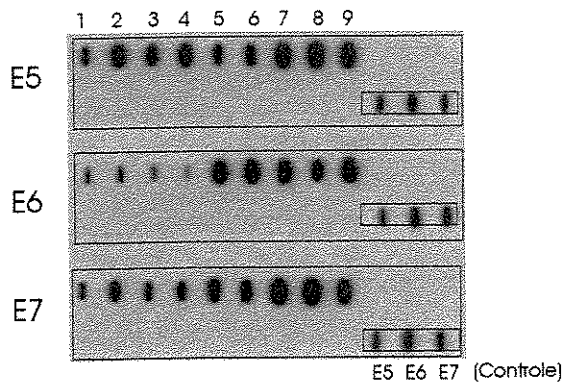
Figura 9. Confirmação da expressão diferencial, por *slot blot*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VG – OPJ08. A. Membrana hibridizada com sonda de cDNA de embrião em estágio globular (E5), estágio cotiledonar (E6) e estágio maduro (E7). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:

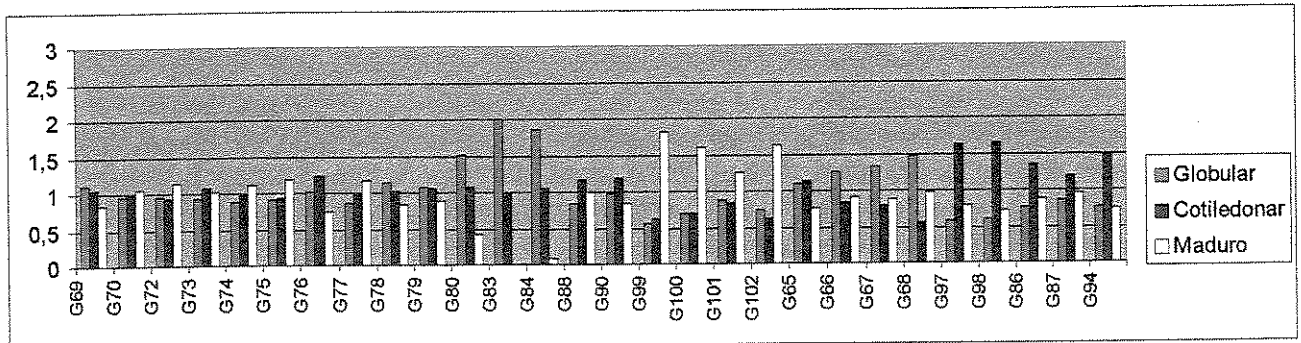
Slot S-04



Slot S-08



B:



C:

Globular
Cotyledonar
Maduro

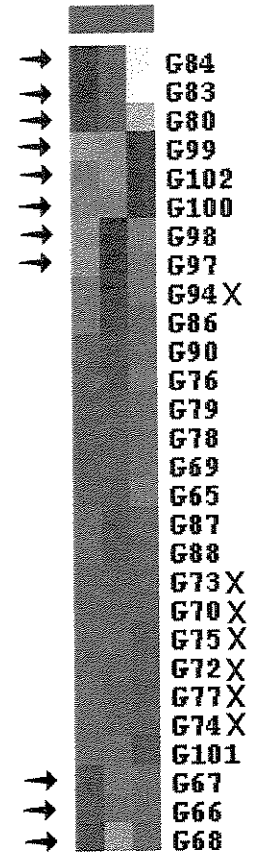
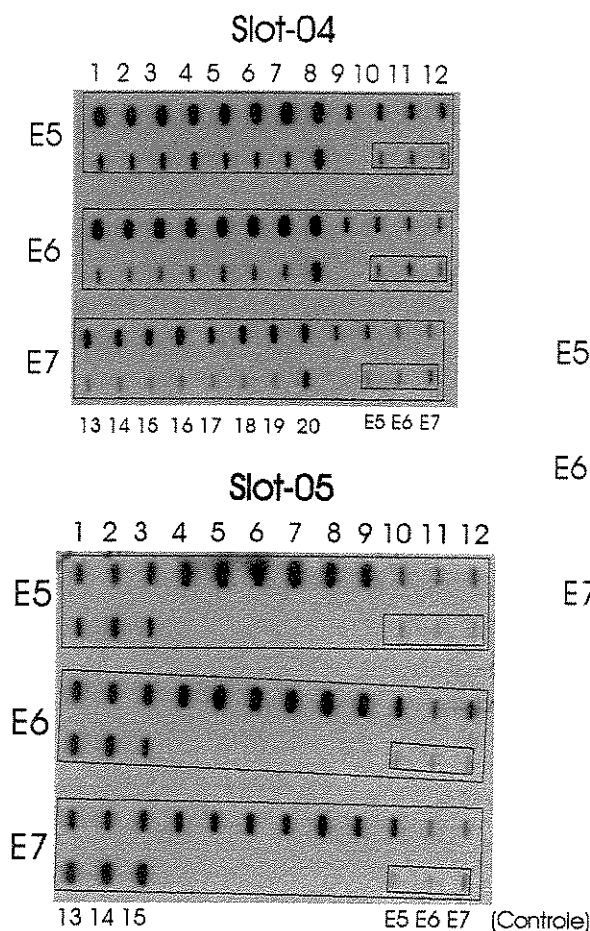
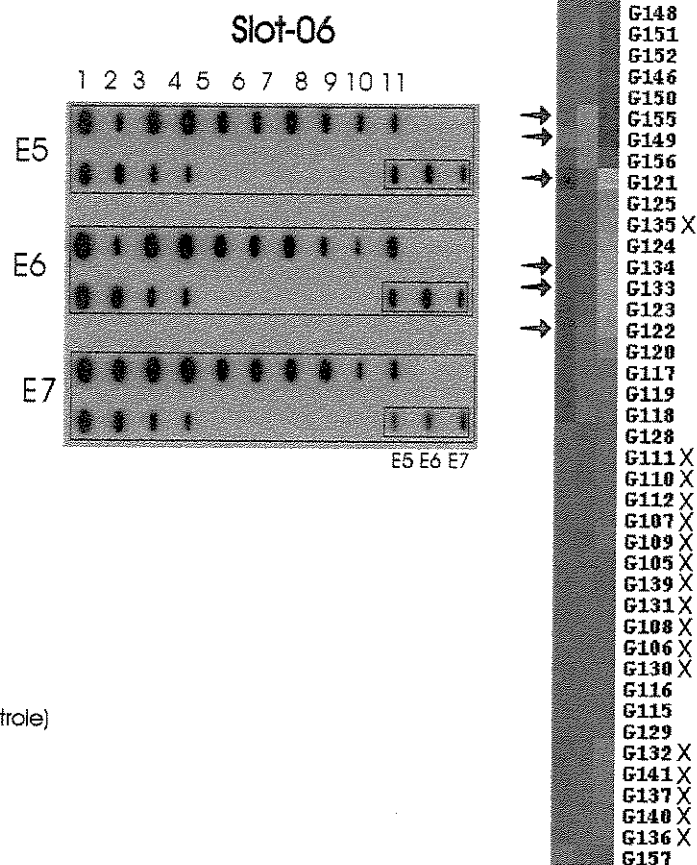


Figura 10. Confirmação da expressão diferencial, por *slot blot*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VC – OPJ09. A. Membrana hibridizada com sonda de cDNA de embrião em estágio globular (E5), estágio cotyledonar (E6) e estágio maduro (E7). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:



C:



B:

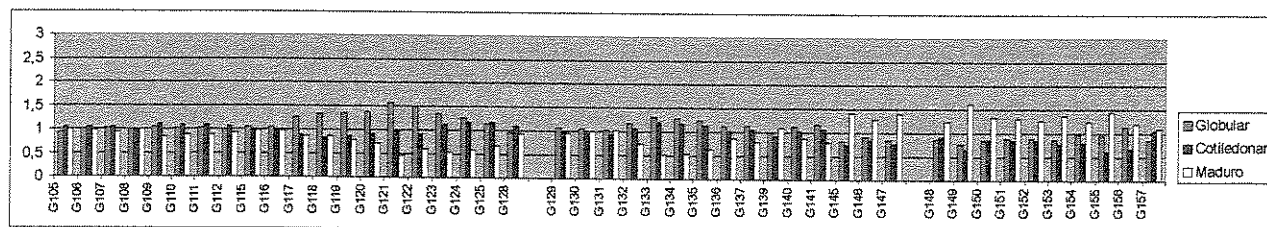
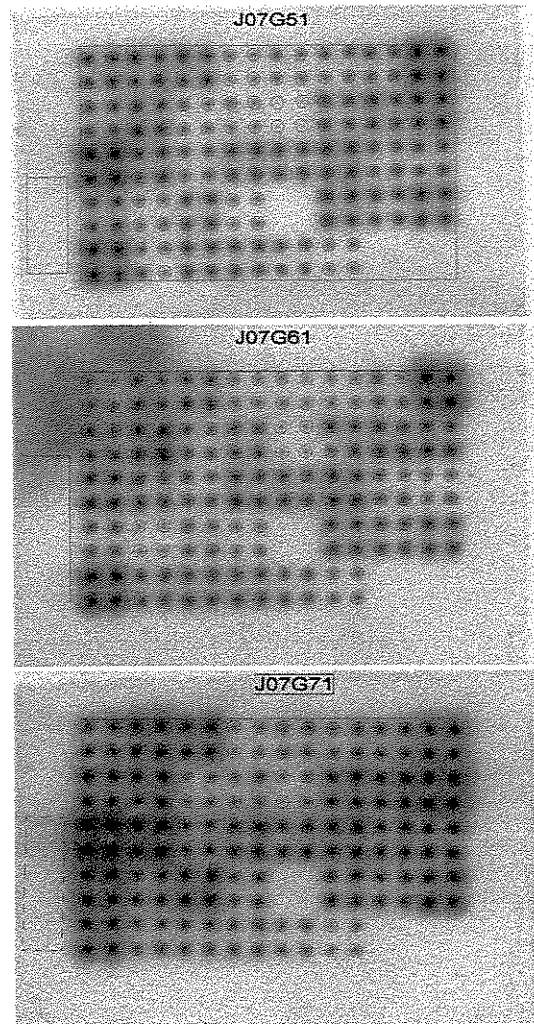
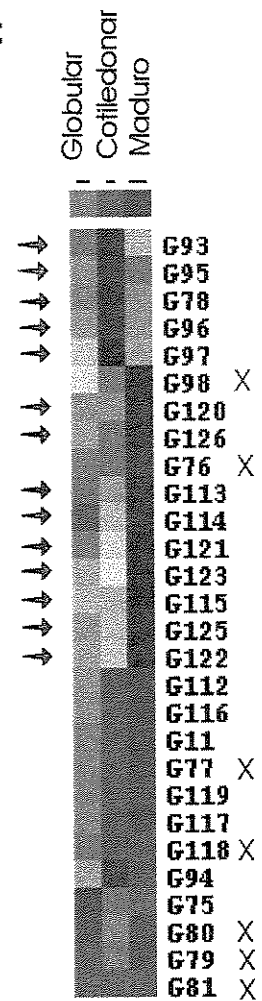


Figura 11. Confirmação da expressão diferencial, por *slot blot*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VG – OPJ09. A. Membrana hibridizada com sonda de cDNA de embrião em estágio globular (E5), estágio cotiledonar (E6) e estágio maduro (E7). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:



C:



B:

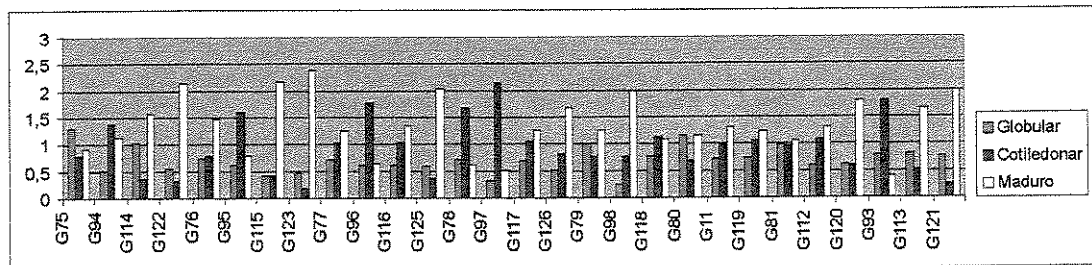
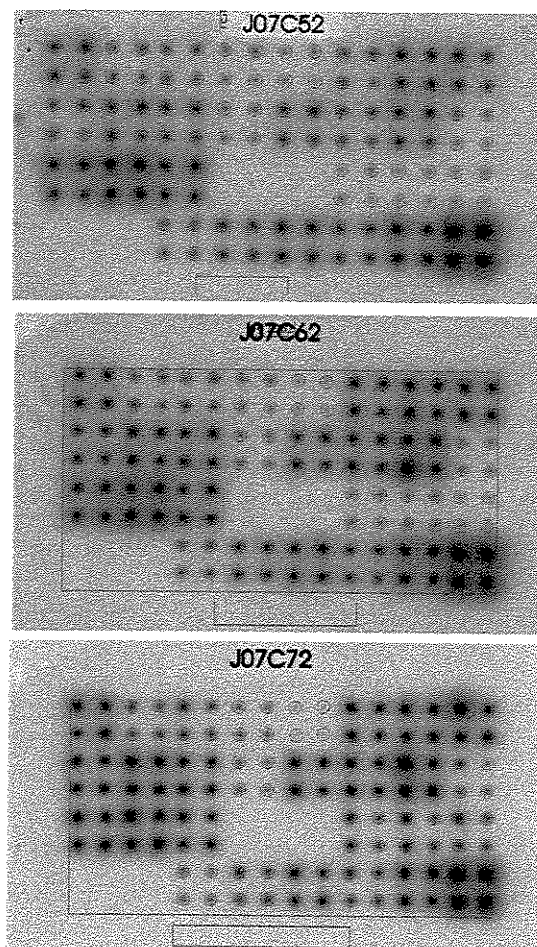
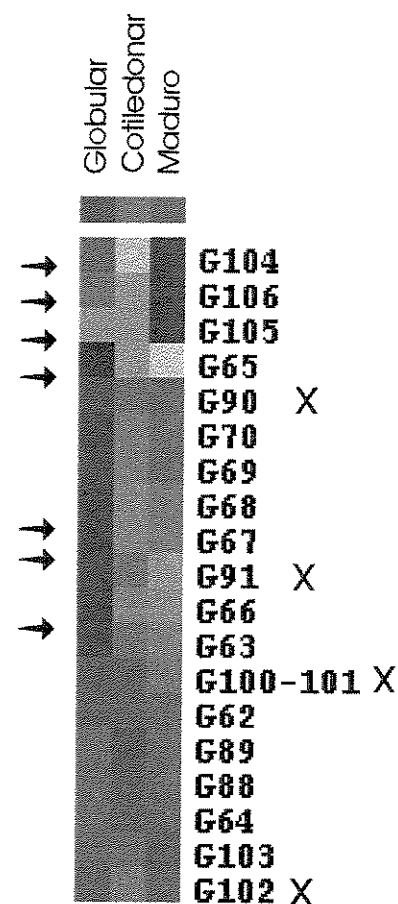


Figura 12. Confirmação da expressão diferencial, por *macroarray*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VG – OPJ07. A. Membrana hibridizada com sonda de cDNA de embrião em estágio globular (J07G51), estágio cotiledonar (J07G52) e estágio maduro (J07G53). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:



C:



B:

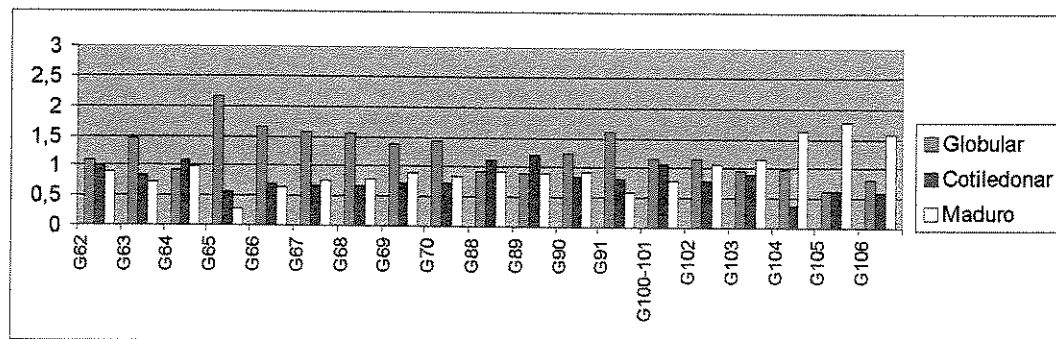
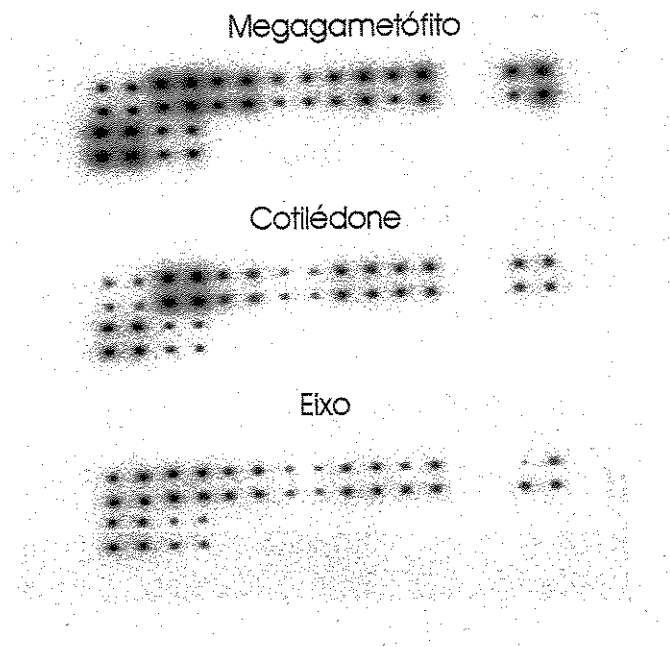
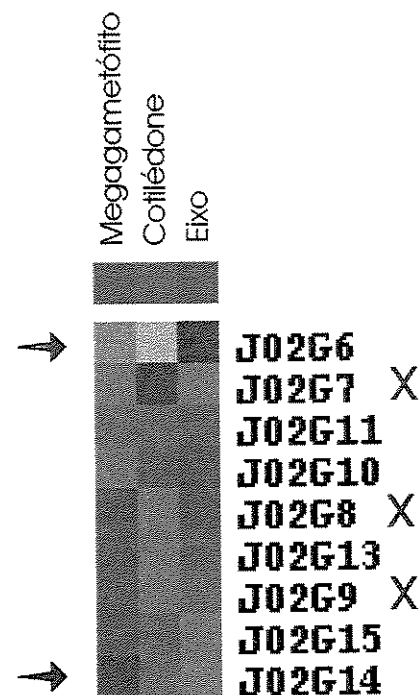


Figura 13. Confirmação da expressão diferencial, por *macroarray*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VC – OPJ07. A. Membrana hibridizada com sonda de cDNA de embrião em estágio globular (J07C52), estágio cotiledonar (J07G62) e estágio maduro (J07G72). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:



C:



B:

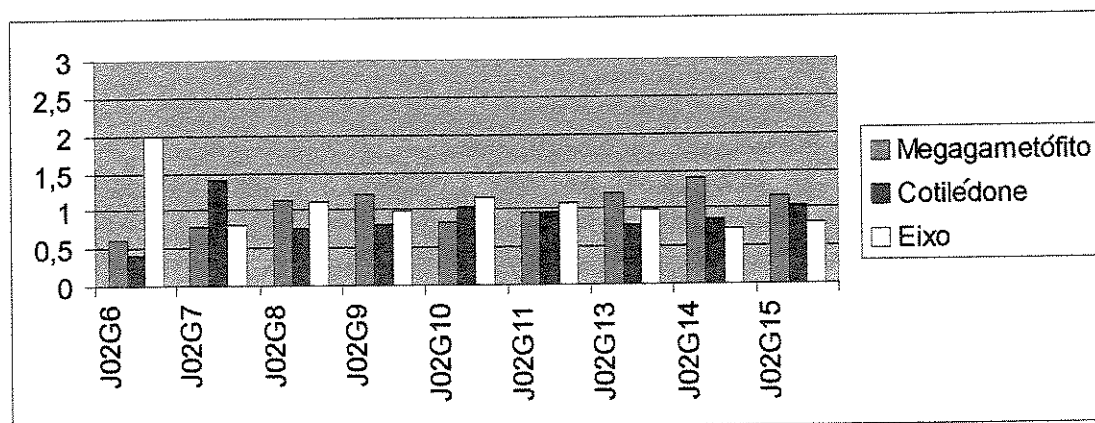
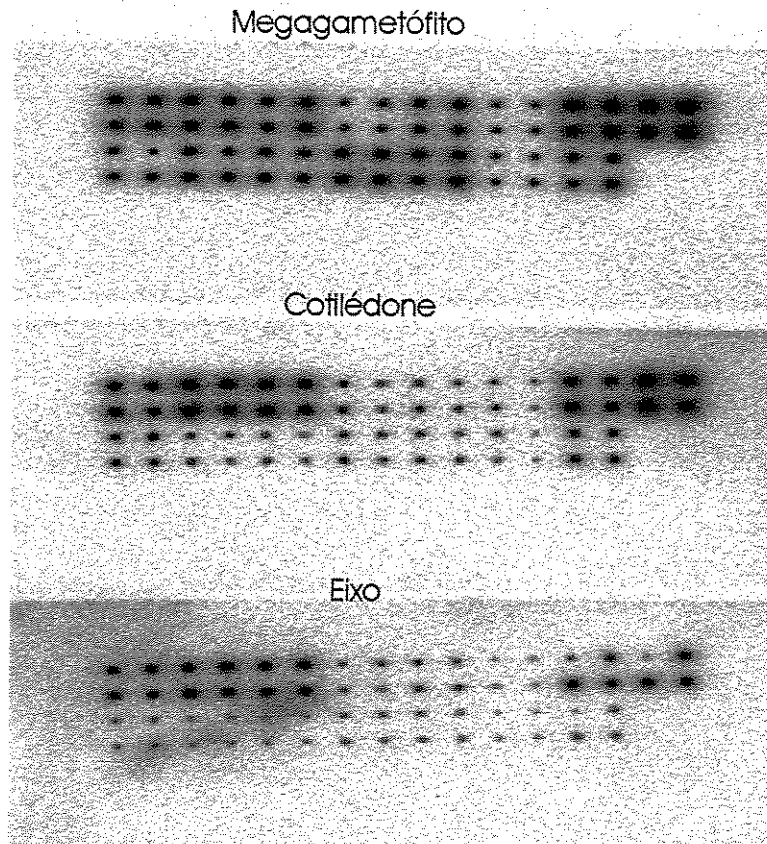
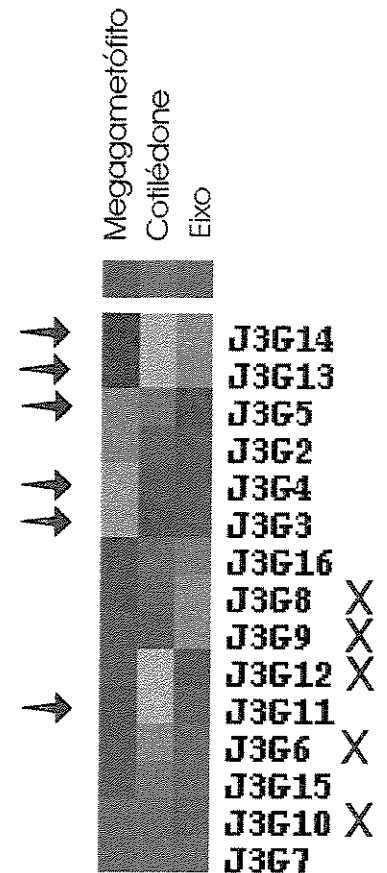


Figura 14. Confirmação da expressão diferencial, por *macroarray*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VG – OPJ02. A. Membrana hibridizada com sonda de cDNA do megagametófito (M) e do embrião zigótico: região cotiledonar (EC) e eixo (EE). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:



C:



B:

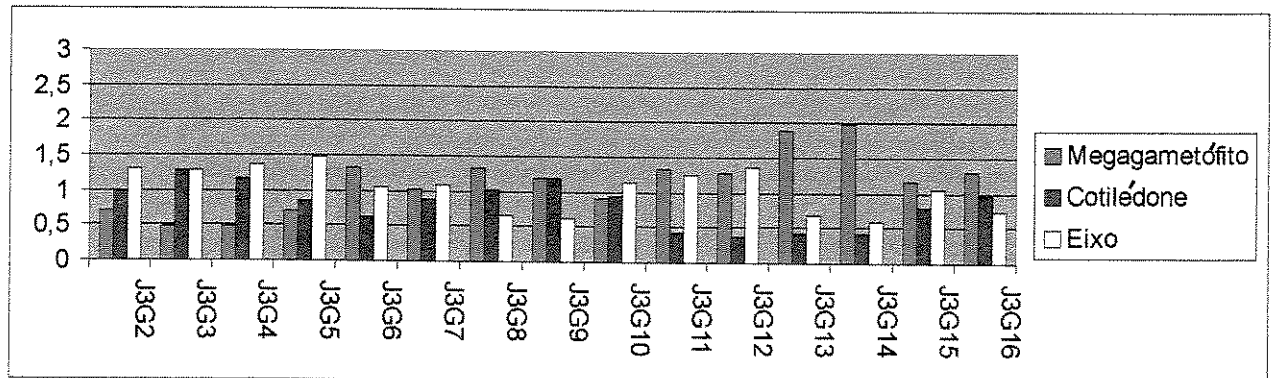
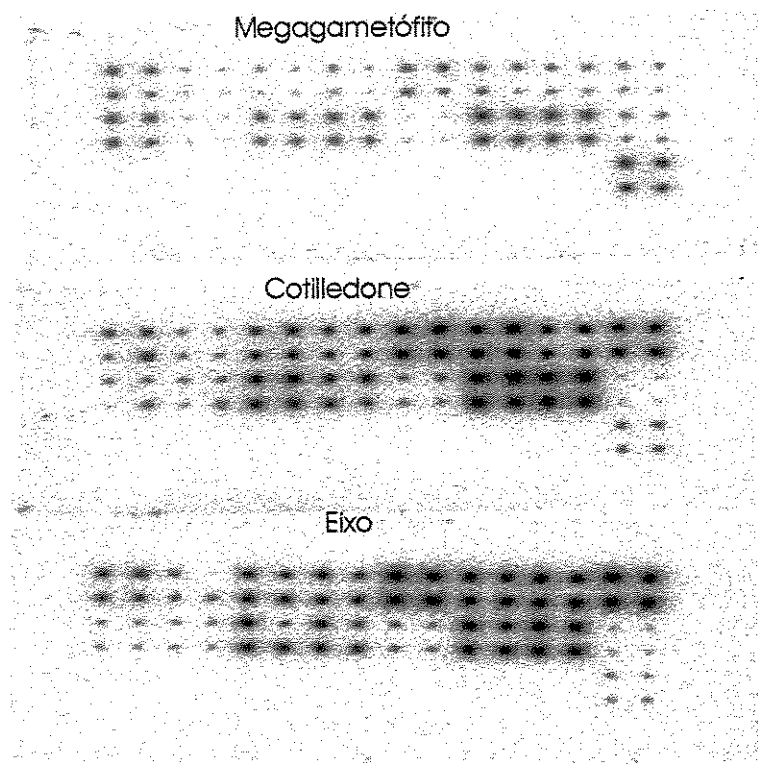
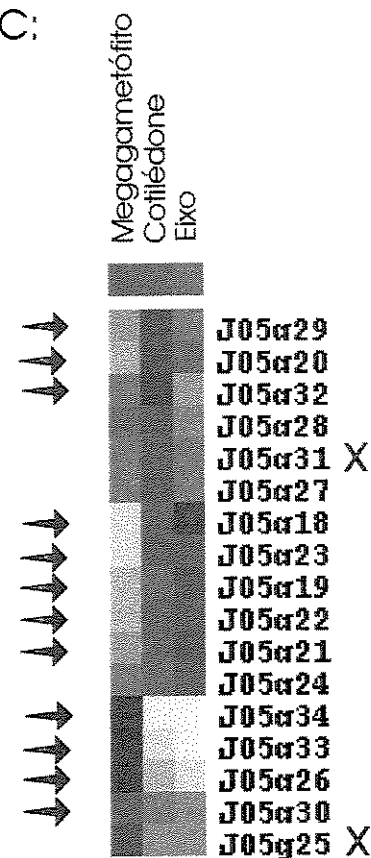


Figura 15. Confirmação da expressão diferencial, por *macroarray*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VG - OPJ03. A. Membrana hibridizada com sonda de cDNA do megagametófito (M) e do embrião zigótico: região cotiledonar (EC) e eixo (EE). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:



C:



B:

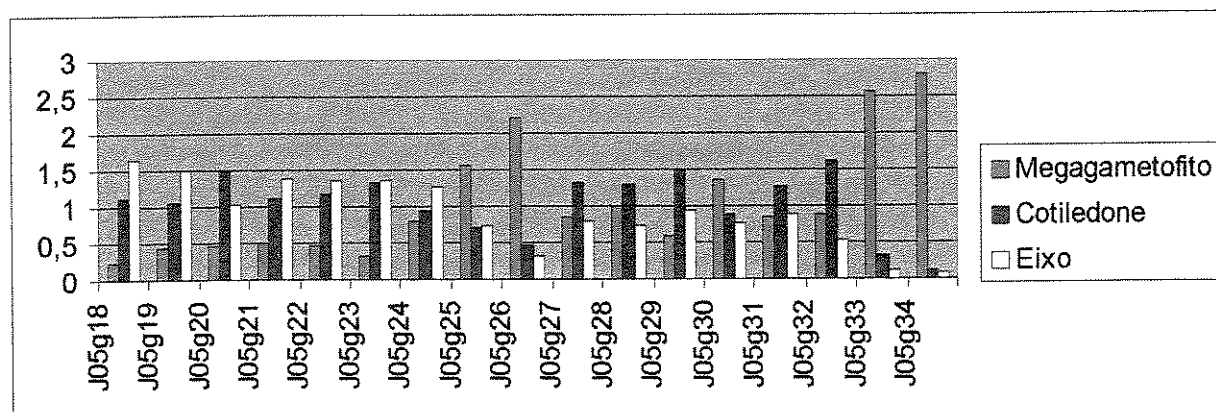
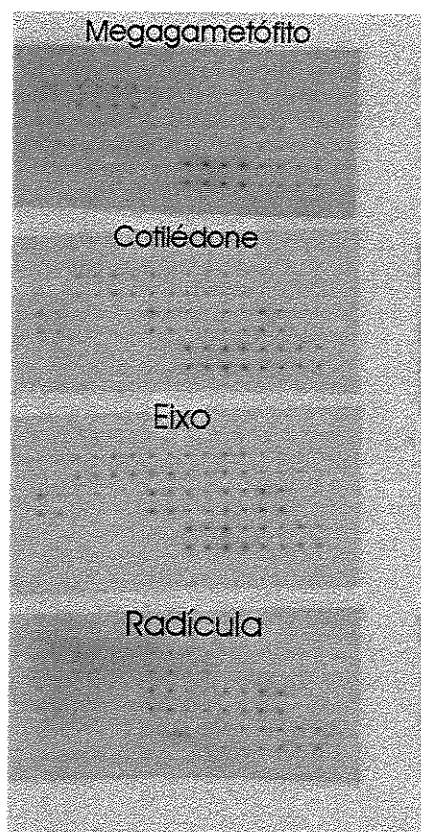
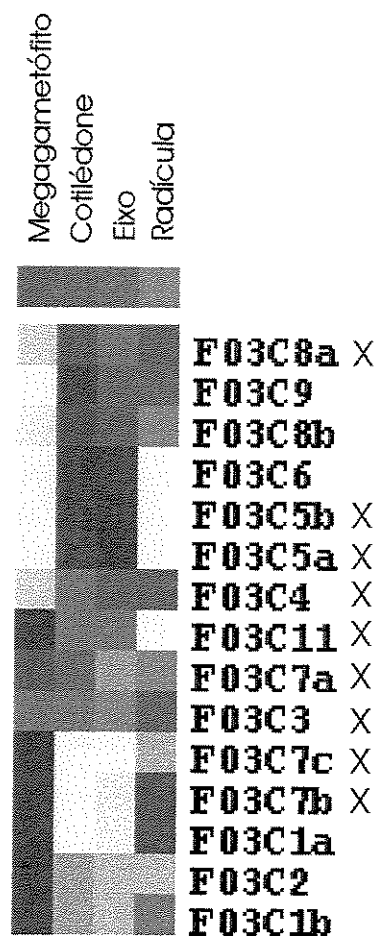


Figura 16. Confirmação da expressão diferencial, por *macroarray*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VG – OPJ05. A. Membrana hibridizada com sonda de cDNA do megagametófito (M) e do embrião zigótico: região cotiledonar (EC) e eixo (EE). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X), e as que apresentaram um padrão de expressão bem diferenciado foram marcadas com uma seta (→).

A:



C:



B:

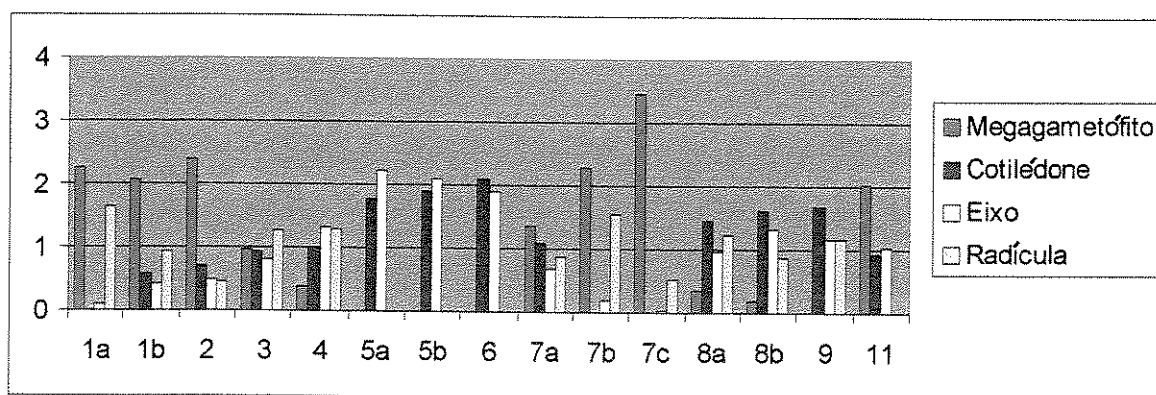


Figura 17. Confirmação da expressão diferencial, por *macroarray*, das bandas isoladas através de *display* diferencial de mRNA com os *primers* T₁₂VC – OPF03. A. Membrana hibridizada com sonda de cDNA do megagametófito (M) e do embrião zigótico: região cotiledonar (EC), eixo (EE) e região radicular (ER). B. Representação gráfica dos valores da expressão relativa das bandas. C. Agrupamento das bandas de acordo com a sua expressão. As tonalidades vermelhas representam expressão mais acentuada, enquanto as amareladas representam ausência de expressão. As bandas não confirmadas foram marcadas com um (X).

ANEXO 3

Sequências dos clones de cDNA de EG, EC, EM, Ec, Ee e M isolados através de display diferencial de mRNA.

>J07G-115

CCTCTCGACA TTGCTGGCAA ATAGGAAGAA AGGTCATTTT TTGAAAGGTG ATGCAGTTGT
AGTTGTACGG GGAGATTTAA AAAATTTGAT GGGTTGGGTG GAGAAAGTGG AGGATGACAA
CATCCATGTC AAACCAAAGG AAAAAGGCAT ACAGGGGACT TTGATCTGTA AGGAGAACGA
GCTTTGCAAG TTCTTCAAAC CCGGTGATCA TGTGAAGGTT GTGAGTGGCA CCCATCAAGG
TGCTACTGGC ATGGTGGTTA AAGTGGAAAA TCATGTTCTT ATAATAGTGT CGGACACAAC
AAAAGAAGAT ATTCGTGTAT TTGCGGATAA TGTTGTTGAA AGCTCTGAAA TAACTTCTGG
CGTTACTAAA CTTGGAGACT ATGAAC TACA TGATCTTGTG CAACTTGATC ATACTAGCTT
TGGCATTATT GTACG**TGTCG** **AGAGG**

>J07C-104

TTTTTTTTTTT **TTGCCAATAA** ATCATCCACA TAGTCTTCGA TGATCTTATG CATATAATCG
TGGAAGATGT AAGTCATGGC ACGTTGATAA GTAGCACTTG CATTTTTTAAG ACCAAAGGGC
ACAACCTTGT AACAAAAGGT CCCCCAAGGG GTTGTAATATG CTATCTTGTG TTGGTCTTCA
GGATTTATTG AAATTTGATT ATATCCTGAG AAACCATCCA TGAGTGAAAA CATTTTCATGT
CCAGCGTAGG **TGTCGAGAGG**

>J07G-126

CCTCTCGACA ATAGCCGACT TGCAAAGTAT ATGGGCTGAT CCAACCCATA TTGATCTTCT
TGTGCTAATG CACATCCAAT GGCAAATCTT GAGGCATCAA CATGTATGTA AAATTCTTTT
TCCCAATTTG GATAAACCAA TATGGGTGCT TTTACCAACT ATCCCTTCAG TCCTTCAAAT
GCCTCTTGGT GTTCCTCTGT CCATTTGAAT TGAACACCCT TTCGTGTGAG **TGTCGAGAGG**

>J07G-113

CCTCTCGACA GCCAAGGAGA AACTCCCAAT CATATTCATG GTGATGAAAC AATATAAACG
ATGGCATTAT ATGAATCTAA ATTTTCATGTC AACTCTGCAG ACGATTCTAT ACATTCCTGT
TTAACTAATG TATTTTTGCA TGACTATTGT TAGTGCATCT GTTAAAGTTG TTTTTTCAAC
CTTTACTAGT CATGCAAAT ATAGACTAGT GTCCTCTATG CAAGCTCCCT TCTCATAAGC
AACATGCTCA CATATAATCC ATTTGTCGAG AGGAGGTGAA GAGGATTAGC CATTTAGCAC
TTTAGCCTGA GGCCTTCTTG AGATGGGCAA TTTGGTTATG TCATCAATGT AATTCCACTT
AGTCAGCAAA GGTGCAACAT TTTACAGTTT TGCAGTATAT GAAATTGTTT TATGAAAGGA
AGTTACTGGT AGTGTGTAAC TTGGAGTGGA ATTTCTATTT GGAGGCTGAC ACTGCCAAAG
TTGCTCAAGA ATGGTGGCCA **AAAAAAAAA A**

>J07G-123

CCTCTCGACA CCTTATTACA TTTTACACCT ACTTTTTATG GTTAAGATTA GATTTAGAAT
TCCAACATAC CAACCAAAAA GTGGTAAAAA CCCATTTTTT CATTAAGAAA AAGTCTTAAG
GAACCATTTT AATCTACTTA TTTGGGATCT TAACTTTGAT CCATTCCTCTA TTTTGGTTAG
GATTCATTTT TTTATTCCAC AAATTAGCCA ACCATTTTAA AATACTTTCA TACTACAATA
CCAAAGGGTA TAATAATAAT ATTGAGGTTT TTACCATGTT CCACCAT**TGT** **CGAGAGG**

>J07G-114

CCTCTCGACA GG TAGCCAAA AATATGGTAA ACATCTTCTA AGGTAAGGGT CATTTCCCTA
GTTGGAAACT GAAAGGTTCT CTTATCAAAA TCTCACCTTT CAACTAGGGA AGATAGATAA
TACTTGTTAA TTTGCACTCA TTCTAAATGA GTGACACCAT CTAACCTCTT CTTTGAGTCA
GAACTAGTCA TCCACATCTA ACCTACTTCT AACTTTTTTAA AATATTGGCT TACACACTAT
TTGTGCTCGT CTAATCCTAA TCGACTCTAC CATGATCTTC TGACTTGCTT CAGGGTTACT
TGAAGCCATA AAAGTAGCTA AATGCTAGCA ACTAGTCAAC ATTCAAAAAT AATTGTACCA
GAATAGAGGG TTCGCTTCNA ATTACCAAAA TGGAGATGCA GTTGGAGTAA TGATGAATGA
TCTATAAGAT GCTAAAGGAA TCATTTTTGT GGGCCGTTGT GTGTAATTAT TGCAACACGG
GAAAAGATTC CCTATTTGTT **GTGTCGAGAG G**

>J07C-105

CCTCTCGACA TCCTTGACAA AACATGAGAT AAACACCAGG GATAGAAGCA CTAAACATTA
 TATAGGAATC TTTGGTCGTG CTCATTCTAT TTTCTCTTGT TCTAGACATC TTTTGTAGCT
 GGCATTACTT GGAATTGTTG CCAACAAAAT ATTTTCTCT GATTATTTAT CCCACTTCTT
 GCAATCAACC ATATCAGCAA CATGATAGGC AATGTCTGTC ACCACCTGTG **TCGAGAGG**

>J07G-96

CCTCTCGACA GCCAGAAGCA ACTAAATAT CAGAATTAGT ATATCCTAAA ATGATCACTA
 AGTCTTTAAC TCAACAAGCT CAATCTTGAC TAGAGGAATA ATTTTGTCAA ATTCTTCCCT
 TTACTTATTT GAGAGTGCTG CTTGAAACAT TTTAACAAGC TGA CTGGTCT TGA CTCTTGA
 AAACAAATAT CTATGCTTTG GCAAGTCTAT GAGCGCCTTG ATGTGGAATA TTTACAATCT
 GTGTTCTTCA TCATCCACA CAGGCACTCA CAAGGTCATT AGTGAGTGGA TCATTGATGA
 AGAAGCCACC TTCCTTATGG **CCAAAAAAAA AAAAA**

>J07G-97

CCTCTCGACA TAATATACAG AGCAAACACG TTCGATTTTA CCATGGCCCT GCACGCTTCA
 CATCAATAAA CATCTTAATG CACGAAAAAG GTACAGTGCA CTTACAATAC AAAAAGGATT
 ACGAACCCCA AGTTCATAAA TTTTAGCCAT TCCTTTTCTA CACAAGACAC ACAAACCTCG
 CGCTCATTCT GAATTCTAAC ATAGACCAGA AACGGCATT AAAAAATTCT CAGTACAATA
 AAAAGGTAAA ACGGNTAACA TTCATCATCA GATTTGACGA AAATCTCGAT CTCCGAGTCA
 AAACCTCGTC CCCCCC**CCAA AAAAAAAAAA**

>J07G-93

TTTTTTTTTT TTCGGAGTTA TTGCGGAGTG GCCAAGTAAC TCTCCAACT TAGAACTTAG
 GCTACCTCCT GAGCTGAGCG TTTTAAAACA CTGCCTATAG CAATTAGATA GATATTATGT
 TAGAGAAAAA ATATTTGATA CAGGTAGATG GAAGAGACCA AAAGAGGAAT GAGAAATGGGA
 TGAATTTTTT TTTTTTTACA TAGGAGAGTT GATTAGCATC AATGAAATTA GACTCGATAG
 ATAAACCTCT CCAGAACCAA AGCAGTCTCA TTCAACAAAC GAATGTGCTC TATAAAAAATC
 ACCAGAAACT CTATAGTTCA ATTCTTCAAC TTTTGAAACC TCTTTTGCCA TCTTGGCCAA
 CGTATCCCTT GCCCTCTTGA AAAAATTAAA CTAAAGAACA CAGAGAACGC TCCCAAATTT
 CAACAAATTT CACCTGATCG AGATACCAGA CTAGTTGTTG GGGACTGTTA TAAAACAACA
 CTAGCTGTAA AACTCATTTT GATAAATTTT TATACGTAAC TGCCTGGATG **TCGAGAGG**

>J08C-16

CATACCGTGG AGCATCGTAT AATGTAACGT ATATAAGCCT ACTTATGTAC GGGGTATTAT
 AGGTATTAGA TATTAGGGTG TGAGTAATCT AAATAAGAAG ACGAAATAAA TGTTTGAATG
 AGTATATNGT ATGTGAATAG CATCAAAATG AGATCAAGTG TTCTTCT**GCA AAAAAAAAAA**
A

>J06C-8

TCGTTCCGCA TTCACTCCGC CAATCAATCA AACTCTTTTA TTCTGCAGCA CTTGGAAGGT
 TTGTTGCTTC CAATCTCCGN TAATAGNCNT CATCTCAAGG CCTATTACCA GTGAAAAGGT
 ANCGAGTTTG ATCAAGGTCC CCTTATACAT ATGTAAAAAG CTTCTATTTT CCCTTTTCCT
 TTTTATTCA AATNTCTCTT TGCAAAAGAG AAATTTGATT NTGNATATAG ATGNACGAAC
 CTTGGATTTT TTCTAAGGGG TACAAATTCT TTCTATGTGC TCTTTG**GCaa AAAAAAAAAA**
A

>J06G-21

TCGTTCCGCA CCCATAAAAT CTTATGCCCA GCAGCAGCCT CCCTGTTTGG CAGCCAAATT
 TCCATGCTGG GACTCCGCAG ATCCCTGCAG CTCCTTTTGG GTCACCTGTA TCTAGCTACA
 GCACCCCCAA TGTGTCGTCT GGTCTCAAT ATTCCAATAG CCCTTATGGG TTTTCAGTTT
 CTGGTTCAAG CAATGGAAGC TTCAATCAGA CCCCCTGCTG GTCCCAACAA TATTGTTTAC
 AACCTCACT **GCCAAAAAAAA AAAAA**

>J09C-83

TGAGCCTCAC GATATGTAAT GTAATTTGGT GGATGTGTGA TGTGATGATG CAACAACATC
 CAGTTGCTGC TGCTGCTGCT GTTGCATTAT GCTTTTGCTT TTTCTGTCGC TACATAGTTG
 TAAACACATG GATTTAATAT ATGTTTTTAA TTCCACTGCA **AAAAAAAAAA A**

>J09C-66

TGAGCCTCAC CAGATGAGTA AGACCCTTGG GCTACTTGAT TGTGTGAGT ATTGTTGGTT
 GGTGCAACA AGGATTGCAA GAATGACATA GAAGGTGGTG ATTGAGGATA TTGCTAAAGA
 ATTTGTTGGG ATGAAATTGG GATTGCATGA GATGAACCTG GATAACAATG AGGAGTAGAA
 AATGATGGGT TAACAATAGA CTAACCCTGA TTGGAATGAG GAATTGGGAT AGCTACTTGG
 CTAGATGGGG TTACCCAAGA TAAGCTAAGG GTGAAGAAAT TTGGTTGAGA AGAAAGGGAA
 TTTTAAACCT CTACCGTTTG TTGCTGATCA AGATGAACAT CTGGTGAACT CATCAGCAAT
 TGATTTAAAT TGTTGCCATG CCTGTTGGAT GGCAATCTGA TGAGCAGTGG GGGTTTGT
 GGGTTGGAAG **AAGAAAAGGA**

>J05G-33

CTCCATGGGG CATGACATGA GGCAACGGAN AGTGTGATTC ACGAGGTGGA GTTGGGCCAT
 TAGGAGGCCA TCAATGTTCA CCATTTTGAT TGTGCCACGA CACATGTATA GATACGGGCT
 AGTTTAGTTG TAGGACAGCA TATAGATATT TTGCAGCCTG TGGGGCTGAA ATTTTGTATG
 TGACAACCAG AAATACTATG TGTGTATATT TTTGGATGAT GACACACATG CATTTGCTTC
 TTTATGATGT AATATTATGT ATTTATATTT TTTGATGATG ATAAACACAT **GCGAAAAAAA**
AAAAA

>J05G-20

CTCCATGGGG AAATCGGAGC AAAAGAGCTT GGAGGTGGTG AAGGGCTTGG ATCTGGAGAG
 GTACATGGGC CGCTGGTACG AGATCGCTTC CATGCCGTCA AGATTTTCAGC CCAAGAATGG
 CGTTAACACC AGAGCAACCT ACGCTCTCAA CGACGATGGG ACTGTCCACG TGCTCAACGA
 GACGTGGGTG GAGGGAAAGA GGAGCTCCAT TGAAGGCAGC GCATACAAAG GTCGACCCCA
 AGAGCGAAGA CGCCAAATTC AAAGTAAAT TCTTTGTTCC TCCATTTCTC CCTGTAATTC
 CTGTGTATGG CAACTACTGG GTAATGCTTC TTGATCAAGA CTATCAAGTG GGCTTTGATC
 GGAGAACCTT CTCTCACTTC CCTGAGGGTA CTTGAGTAGG CAACCTCAAT TGGATGAAGA
 GATTTATAAT CGCTTATTGG AGAAGGCGCA ACAGCAGGGG TATGAAGTGG GTCGGCTGAG
 GAAACTAGT CAGACCGCCG TTCCTGATGA GACTGAGGCC CCAAAGATA AGGGATTTTG
 GTGGATCAAG GCTCTCCATG GGGAAATAGG AGCGTTGGGA AGCAAATTTT TGC GTTGGGA
 AGCAAATTTT TGC GTTGGGA AGCAGATTCT TCTTTTGGG TCTGGTTGTA GTTAA**CAAAA**
AAAAAAA

>J05G-23

CTCCATGGGG ACAATCATCT CATTGGGAGA ATAAATTTT GACACAAGTT TTTGCACTAT
 GTTGTTTTTA CAAACAACAC ATTGCAAAAA GGGGTATTTG TAGACATAAA AAATCAGCAT
 AATAAATTAT GAATTAATAT ACCTAAATAT TCCTTTTTGC CACCTTATAA AATTATAATT
 AATCCATATA TATACCCAAA TTTTCTAATA TTTCTCCATA AATAACAAAC ACTAATAATA
 AACCAAATTT TGAATATATC TCCTATACCC ACAAATAA TAAATACTA AATAACTTCT
 AAATAATACT AAAATGGGGA GATAAATAAT TCTAGCATTT TTCAAATAAT TACAGATACG
 AAGACTATTT TGGGAAACCT AGAAGAAAGT CAAATCAAAT ATATTTATAG AGAGGGAAAC
 AAAGTTGTTG ATAAATGGT GAATATGGCA ATTAGATATG AGAATTTCAA GGTCTGGTCA
 TAAAGCCCCA **TGGAG**

>J05G-26

CTCCATGGGG GACTGAATAA GCTCGCTGAG TTACTGGATG TGAAGAGAAT TGGGGTGTGC
 CATCAAGCTG GTTCAGACAG CTTGCTCACA TCTTGACAT TCAGGAACT GAGAGAGGGT
 TTCTTCAATG GCTCCACTGA AAAGTATGCA GGTGTCTTGT ATGGGTGGG TTGTGAGGCT
 AGTGAAATG GTCATTGATA GCCAGATCAT AGCCTTTCAT CTCTTTGTAC ACTTNTTTT
 TTCTCTCGG GGAAAACCT GGAACCTAAG TTAATATTGC **CAAAAAAAA AAAAA**

>J05G-18

CTCCATGGGG ATGTGCATTG AGTTTTGGAG GGTTTATCAA AGATTTGTCT TGCCATCCAA
 GTGGCTCGAA AATGCATTTT GATGTCCTGG AAACATATCT TGTCATCTCG GACTGTCCTT
 GTTCTGGAAA CTTCATAGAA CATGGTTTAA AACTTTAATA TAGAAAATAT TCTCTCTATG
 AATGCCATCA AGAAAGAGAA GGTTGTGTGT TGTCATCCAT CCTTCAAGAC GATGGGTGGA
 AAAGTAAGAA GTATGCAGGA TTGAATAGAA GACCAAATA GTATGTGACA GCAATAAGTG
 GGAACATAG ATGGATGAAT GATCAAGTGT GAGACAGCTG AGCTCAAATT TAAGAAAGTT
 TTACTAGCTG GTTTAATGTG CTTCAGGGTT GAAGCAATAT AAGGGGAAGA GAATGTTATG
 ATCCTATCCA ATTATGGATT TTGTCCACGA GTATGATCAT AGAGTCAAGG GTAGGCCCAT
 TAAAGGTTAC AAAAGCTTTT TTAAGTTGTG ATTCCAATTT GAGTTTGGTA GAAATTTACA
 TTATTATATC TAGTCTTTTT AGACAAATCA ATTTAATCAC AAGGTATCTC TGTTTGTGTT
 TGTTTCCAAG TGAAAGTGGA TTTTTCTAT TTGAAGTTAG GTTTGCTCAT AGCTTGGGAG
 CCATTAGTAA AGCCATATGC ATTTTGTCAA AGTAAAGTAG GACACCGGCA AGGATGCCAC
 TTGCGTTATC GGTGATGTGA CCCGGGTTCA GT**CCCCATGG AG**

>J05G-19

CTCCATGGGG ATGTGCATTG AGTTTTGGAG GGTTTATCAA AGATTTGTCT TGCCATCCAA
 GTGGCTCGAA AATGCATTTT GATGTCCTGG AAACATATCT TGTCATCTCG GACTGTCCTT
 GTTCTGGAAA CTTCATAGAA CATGGTTTAA AACTTTAATA TAGAAAATAT TCTCTCTATG
 AATGCCATCA AGAAAGAGAA GGTTGTGTGT TGTCATCCAT CCTTCAAGAC GATGGGTGGA
 AAAGTAAGAA GTATGCAGGA TTGAATAGAA GACCAAATA GTATGTGACA GCAATAAGTG
 GGAACATAG ATGGATGAAT GATCCAAGTG TGAGCCAGCT GAGCTCAAAT TTAAGAAAGT
 TTTACTAGCT GGTTAATGT GCTTCAGGGT TGAAGCAATA TAAGGGGAAG AGAATGTTAT
 GATCCTATCC AATTATGGAT TTTGTCCACG AGTATGATCA TAGAGTCAAG GGTAGGCCCA
 TTAAAGGTTA CAAAAGCTTT TTTAAGTTGT GATTCCAATT TGAGTTTGGT AAAAATTTAC
 ATTATTATAT CTAGTCTTTT TAGACAAATC AATTTAATCA CAAGGTATCT CTGTTTGTGT
 TTGTTTCCAA GTGAAAGTG ATTTTTTCTA TTTGAAGTTA GGTTTGCTCA TAGCTTGGGA
 GCCATTAGTA AAGCCATATG CATTTTGTCA AAGTAAAGTA GGACACCGGC AAGGATGCCA
 CTTGCGTTAT CGGTGATGTG ACCCGGGTTC AGT**CCCCATG GAG**

>J05G-22

CTCCATGGGG ACAATCATCT CATTGGGAGA ATAAAATTTT GACACAAGTT TTTGCACTAT
 GTTGTTTTTTA CAAACAACAC ATTGCAAAAA GGGGTATTTG TAGACATAAA AAATCAGCAT
 AATAAATTAT GAATTAATAT ACCTAAATAT TCCTTTTTGC CACCTTATAA AATTATAATT
 AATCCATATA TATACCCAAA TTTTCTAATA TTTCTCCATA AATAACAAAC ACTAATAATA
 AACCAAATTT TGAATATATC TCCTATACCC ACAAAAATAA TAAAAACTA AATAATTTCT
 AAATAATACT AAAATGGGGA GATAAATAAT TCTAGCATTT TTCAAATAAT TACAGATACG
 AAGACTATTT TGGGAAACCT AGAAGAAAGT CAAATCAAAT ATATTTATAG AGAGGGAAAC
 AAAGTTGTTG ATAAAAGGGT GAATATGGCA ATTAGATATG AGAATTTCAA GGTCTGGTCA
 TAAAG**CCCCA TGGAG**

>F03C-1

CCTGATCAC CCCGCCCAAT AGGCATGATA TTGACCTCAA CTCACTGTTT AAGGTCATAT
 TCCACTAGTA CATCATATC ATTTATCAGG CTAAAAAAT CCAAATAAGA GGAAGAATGT
 GACATTACAG TCTCACTGCT AACCAAGATG ACACGTGTTA AGAGAATTAT GAAAGCTAAG
 AGGAAATTTT TCTTAACAGA AGATATTGTA ACTCCTGCAT TAGATTGCAT TTATGTCGCA
 ATTTATTTTG TACTTATTTT TTAAGTCAAT ACACACTGGG TTACTTCCCA GTGAAGTCAC
 TACTTTCAGT GAGACAGTAT TTTTAAATGG AAATATAACA AGATTATTCA CATTATTCT
 GTGGTTTGAA ATTAATGAAA ACCAAAGGCG AAGTATTTAT AT**GGTGATCA GG**

>F03C-2

CCTGATCAC CGAGAAGCAA AAATTACAGA GCCGCTGAGG TGGACAAATG TCGCCACAT
 CAACATCAAT TGGCTTGCAC AGTCCTCAAT GGGCAGAGAA GATTTGATA AGATCCTCCT
 ACAAAAAGAAA GAGGATGGCT TTTCATTGAC AACTACCATG GAGTCATAAA AAGTATTGCA
 GGAGCATCTT CAACATTCTG CTATTGGTAC AATGGATGAT CACCTTGTGG AATTTTCAGA
 AGAATTAAGA ACTGTTGCAA AGGCACTAAG ACAGGCAGCA GAGAGTAAAG CCAGGGCTCT
 AGAAGAGGCA GCTGAGT**GGA** **AAAAAAAAA** **A**

>F03C-6

CCTGATCAC CAATACACAT AAGAACATAG TGGTCAAATC ACACAACCAA CCTATGAGGT
 ACCAACTTTT AAATGGAAAA GACCCCTTAC AGGGCATCGT CATAATATCA TCTTCTTAAT
 CATTCTTAAT ACTATTGCTA CATAAATCCT TGACAACTT TGCTTGTGTTG ACATATAGTG
 CATACTAGCA TGCTTCGTTA AAAGTCAACT CTATTATGCC TTCCCAACA AACATAGTTT
 TTTCGTCTAA TCTATATAGG ACACAATGTG AGCCTTCAAC AACCATTATT CACTTCTTGT
 TCTCATGTCA TTCGCATCCC CTAAGTGTAG CTTCAATAGG AACTTTTACC ACTTCTAACA
 TTGTGCATCC TTGGGCCTTA TCCAATTTAA GCAC**GGTGAT** **CAGGA**

>F03C-8

CCTGATCACC ATAACCTCGC CCATGTGGCA AAAATGTGCGA AAATTATGGA TTCGGTCTGT
 TATCCCCTGC AAAGGGACAG TTCCCTAGTG TATTGTCCGT GTTATCTGAA CTCTGAAAAA
 TTGTACAGTA CCAGAAATGG GACAGTGTTA GAAATGTAA CTTCTCGATA TATGAATATC
 ATCCAAAAGG TATAAAGGAA AGTGAAGATG TTCATGGTCC AATTGAAGAC CAGTTCAACT
 CAGATGTGGG **AAAAAAAAA** **AA**

>J05G-21

CTCCATGGGG AACTAGGGTT GAAAGGCCAA CTCAAGTGGG TTCTGCCCAC AAAGTTTATG
 TTGGAAACTT ACCATGGCAA GCTGATGATA ATAGCCTGTT GCAATTGTTT AGTGAACATG
 GAAAGGTTTT GGAGGCAAAA GTTGTGTATG ATAGAGAGAC TGGGCGTTCA CGAGGCTTTG
 GATTTGTGAC ATACTCATCA GAGAATGAAG TGAATGATGC TATTGCTGCC CTTGATGGAG
 CTGACATGGA TGGGAGGCCT TTGAGAGTTA ATGTGGCTGA GGATAGGCCT CGACGTCAAT
 TTTGAAGCAG AGTAAAATTT GGTTGCTTCT CAATTTTCTC CATTGCTTCC ACTTTTGTAT
 AATGGATTGT TATCCATTTT CTTTTTCCTT CATCTTGAAA ATCCTTGTAT GCGTAACCAG
 AAGAAGTCTC TCAGTGTTC TACATACCGT TTTTGGGGTT GGTGGTTTAT TACACCATAT
 AATCAAATTA TTTAAGTGGT TTATTGNGGC **CAAAAAAAAAA** **AAA**

>J05G-29

CTCCATGGGG GACAGACTAT GTCAATTGCC AGTCAAGTTA TTAGCAATGG AATACATGGA
 GGACAGACAG AGTATTCGTA CGTACACTTC AGTTTTCTT CGAACATGCT GCCTGCTTTC
 AGTATTCTCG AGGGTTTTCC CAATTTAAAG GAACTCTACA TAGACACTGA TTGTTTTAG
 GACCGGTGTG TTGATATGAT CGGATGGTTA TTCAATTCTC CATTGAAGGA GGGGAGGATG
 CAATAGGAGT TGGCATTGTT GGTTCTTTGG AGAATAAGCG GATGGATCAA GAAAAACCGT
 GTGATAGGAC TAATATTGTA ATAGGCTCTA ATCTTCAAGT ATGATAAATT TTATACTGGG
CCAAAAAAAAA **AAAA**

>J05G-32

CTCCATGGGG CAGGCGGTTT ACATATTCAA GCTTACGGGA GAAATAACTT ATTGGTACGG
 TTTTGACACT CTCCCTGCTT CTGCCGTTTT TCAGTCCTCA ATGCTCCTAA ATCCTACATC
 CACTGCTCTG TGATTTTCTT GAAGTGCATA CGAGATTTAA CAAACCGAAT CATTGTTTGT
 GTCTTTGAGA TCCTATCGTC ACACGGGACA CGGATGTTGT CTCGCAACAT CGAGCTTAAA
 ACCGTTATCC ACAAAGACTC ACAATTAACA CCTTATGAAA ATCTAAGCGT GGTATTTGAG
 AACTCCCGCT GCAATTCCAA TGACATGATA CCGCTTATTA TGCTACAATA ATATAGGGAT
 TTGTCTGCG **AAAAAAAAA** **AA**

>J06G-53

TCGTTCCGCA GAGGTGAGGT GGTCGAACTG TGCTCATGGC GAAGATGAGG GTCGAAATCC
 ATACGTTTTT AAGGAGGAGG AGCAAGAGAG GAGACTCGCT ACAGATGCAG GCGAGATCCG
 AGCGGTTCCC CTCTTCAGAG AGTTCTCCAC CCTCGTAAGA GATCTCGAGA ACTATGAACT
 CAATTTCTTT CACATGAAGC CCGATGCGTT CATGCGTCAT CATTACAGCG GTGCAGATCA
 TTTATCGTTG TTCTACAAGG AAAGGCAAGA ATTCAGTGTG **TGCGGAACGA**

>J06G-57

TCGTTCCGCA GCCAGTAAAG AAAATATTAT TTTGTTTGCA AGATTGCTGT AAATTCAGCA
 TAATTTTGTT TTGAGATTCA TCATCCTCAG GAGGAATTGG ATAAAGACCG ATTCATCTTC
 AGATCGGAGC AGAGCAGACT TTCTACGAAA **AAAAAAAAAA**

>J08C-12

CATACCGTGG AGCATCGTAT AATGTAACGT ATATAAGCCT ACTTATGTAC GGGGTATTAT
 AGGTATTAGA TATTAGGGTG TGAGTAATCT AAATAAGAAG ACGAAATAAA TGTTCGAATG
 AGTATATCGT ATGTAAATAG CATCAAAATG AGATCAAATG TTGTTCTGCA AAGACGTCTC
 ATGTTCTTTA AAAATAGAAA TATTGTTCAA ATGTTGTTCT **GCAAAAAAAAAA AAAA**

>J06G-36

TCGTTCCGCA TCGTTCCGCA TCGTTCCGCA GAGGTAATGC ATCCACGAAT TATCTTTCTT
 CATGAATTCA TATTCCTTTT GTTCACACTG ACTTTCAGAG TTGATCAATG CAGGTGAGGT
 GGTCGAACTG TGCTCATGGC GAATGGTAGG TTCCCGTTGC TTTACTTGGT GCTTGTTCTT
 GTAGGTTTTG TGTGTGCAAT CCGTTGCAAT GAGGTTGATC ACTACAAGGT TCATCGACAT
 GGTGAAGAAG AGGGTCGAAA TCCATACGTT TTTAAGGAGG AGGAGCAAGA GAGGAGACTC
 GCTACAGATG CAGGCGAGAT CCGAGCGGTT CCCCTCTTCA GAGAGTTCTC CACCCTCGTA
 AGAGATCTCG AGAACTATGA ACTCAATTTT TTTACATGA AGCCCGATGC GTTCATGCGT
 CATCATTACA GCGGTGCAGA TCATTTATCG TTGTTCTACA AGGAAAGGCA AGAATTCAGT
 GTGT**TGCGGAA CGA**

>J09C-68

TGAGCCTCAC CAGAAGATAC TTTACAGTCT TTGGACTGGA TAACACGTTT AAAGATAGCG
 ATTGGAGCAG CCGAAGGTCT TTCATTTCTG CATCATGAGT GCTCACCTCC TCTTGTCAC
 AGGGATGTTT AAGCAAGTAG CATCCTTCTG GATGATAAAT TTGAAGTACG GCTTGGAAGC
 CTGAGTGAAG CTTGTTTACA AGAGGGAGAA GTTCACCAAA GTGTTATAGC TCGGCTACTC
 CGGCTGTCAC AAACATCAGA GCAAGGGATG TCAGGTAAAG AATGGAATTG CACTTCTCTC
 ATTTTTTTAT GATAATTAGG ATTTATTAAG GCCATAAGAT TGGTCTGCTG **CAAAAAAAAAA**
AAA

>J06G-22

TCGTTCCGCA TAGACATAAT AGCCCTTTTC AGCAATACAA ACAGGCGGCT GCCGCTGCTG
 TAGCAGATTC ACAGACCACA TTTGGTTTTT CTCTCGGATT CGATAGCGGC GCCGCTCGCC
 AAGCGCAGCA GCAGACCCG CAACGGAATA TTTGACCTTG ACCTTCTTAT TATTTGTGTT
 AAGCCTTCTG ATGACTGACG GCGGTTTTTG GTAACGAAAA GGAACCCCTT TTCAGCTTTT
 GTCTCC**CGAA AAAAAAAAAA**

>J07C-66

CCTCTCGACA TCACCTAGTC TACTTTATAT ATATTGTGTA TATATGTCTA CAATATTTAT
 GTGTATGTAT ATATGAAAGG AATTGTATGC AAACACCAAC AAACAACATG AGTACTTAAA
 CATGTAAAT ATATAAATG TATTTATACA TACTTAAATG TAGAACAATA TTAGCAACAT
 TTAGCTATTA AGTTTTCCAT CCTAGAATGA TAGTAGGAAC ACCTACTCAA ACATTTGTAG
 TCAGCATGTT TAACTACAA TATGCAAAT TGAACCAAGA TTATCAAATA AAACCTTGTTA
 AAATATTAGC **AGCAAAAAAAAA AAAAA**

>J08C-84

CCTCTCGACA TCCAGTGCAT TTACGTATAG AAATTTATCA AAATGAGTTT TACAGCTAGT
 GTTGTTTTAT AACAGTCCCC AACAACTAGT CTGGTATCTC GATCAGGTGA AATTTGTTGA
 AATTTGGGAG CGTTCTCTGT GTTCTTTAGA **TTTTTTTTTT TT**

>J09C-80

TGAGCCTCAC ATGTATAGGG ATATGGAGGA GGAACATCAT AATAAATTTG AATATTTTGC
 CTGAATGCAT AGTAGGGCCA TGGTGTGTTG CAATGATTTT GGATACGAAG ATCCAACCTA
 GACTTGCACC ACCAATGAGA TACAACATTC CTAACCTCCC ACCAACATTG ATGTTGTATA
 TACTGCTGAT ACAAGAGACT TTGATACAGA TAACTGTTAA AAGCAACTTC CCCTATTGGC
AAAAAAAAAA AA

>J07G-95

CCTCTCGACA GTAATTGCAG AGCATGTGAG TGCTAAAACT CAATGTTTCT TGGGTGCAGA
 TAAGTAAAAA GTTACATGGC ATCCAATCAA AAATTTATGC CCTCGAGCAT GAATCCAGGA
 TAATTTTCCA ACTTTACCAT TCTTTGAGCT ACATAAATTT TGTACAAAGT TTTCTTGTGT
 GTGTTATGCT GTTTCTGCCT CTAACAATCT CATTAGCAAT TCTGACTGTT TTAGACTACT
 GATGTACCAC TGACTTTAGC ATGATAGATG CCAAAGTGAT GGAATTGGCA TTGAATAAGA
 ATGATAGCTG ATGGAAAATT GTTCATCACC CTGTGGTCAT GTGAGGAAGA ACGAATGTAG
 AACTTTCCAC GAATTGTGCA CAGGATTCTT TGATAAATAT TAAGTTATTT CTTTGGCCA
AAAAAAAAAA A

>J09C-67

TGAGCCTCAC AAAGACCTGG AATCTTGGCT TAAACTATTT AAGGATTCTA ACCCCACAAT
 AGATGACAAG GGATCGGCAA CTATTGTTCT GCCCCCTGTC AAAGGCTTTC AAGACTACCT
 TTATCTTGAT CATGAACTGA GAATTACCAG GGGAAACAGA GGAACAGTCA ATGTTGTCAA
 ATGGCTGCCA CAGAAAGTTA TCCACGTTTG AGAAGAACTC ATTGCATTG TGATTAAATA
 ACGAAGAAAG TTGGTACCTT GCTGGAGAGT TGCGTTTGCT CTTTTGGGGA TGAGAAGCAG
 GTATTCTGCA GAACACACCT GTTTGCTCCA CTATGATTTA AGGACCCATA TTCACAGATA
 CAAATGTAGG TATTTGGCTG AAAAACTTTT ATTATCTTCT ATATTAAAGA ACCATATCCA
 TGTGTACAAA GTTATGAGTT TTAAATATTT GATCCTTTAA CTAAGGACCA AAGTTATGAG
 CCTCACTGAC AAGCTTTCAC AGATATGTGG GCACTGCCAA TATGTAGCAA AAAAAAAAAA
 AAAGGAAGAA AAGAGGAGAA GAAAAACTAA TGGAAGGGTT AAAATATGGC CTTGGAGGCG
 TTAGGAATAG TATGGG**GAAA** **AAAAAAAAAA**

CAPÍTULO 3

Isolamento da seqüência completa e caracterização dos cDNAs J6G-36, J5G-26 e J5G-20

1. Introdução

Visando uma melhor caracterização de cDNAs de interesse geral, foi realizado o isolamento da sequência completa daqueles que apresentaram homologia com a proteína de reserva (vicilina 7S), com o fator de transcrição AraCAF 1 e com a lipocalina, proteína envolvida nos processos de transporte de reservas.

O isolamento da sequência completa do clone homólogo às vicilinas 7S de *A. angustifolia* (J6G-36), tornou-se um objetivo importante na tentativa de esclarecer as características da proteína como o tamanho do polipeptídio codificado, e aspectos evolutivos da similaridade com outras sequências de vicilinas descritas em angiospermas e gimnospermas. Estes dados complementam os já obtidos nos capítulos 1 e 2.

O isolamento da sequência completa do clone homólogo ao fator de transcrição CAF 1 (J5G-26), permite a comparação desta sequência com outras já descritas para bactérias, leveduras, angiospermas e mamíferos, podendo ser observado o grau de similaridade entre estas proteínas e possibilitando estudos posteriores da função deste fator de transcrição no desenvolvimento da semente de *A. angustifolia*.

Por ultimo, a sequência completa do clone homólogo à lipocalina (J5G-20) foi caracterizada, possibilitando o estudo da similaridade com outras proteínas desta família.

2. Isolamento da sequência completa e caracterização do cDNA J6G-36 – vicilina 7S de *Araucaria angustifolia*

A sequência completa do cDNA J06G-36 foi isolada por RACE (*rapid amplification of cDNA ends*, Chenchik *et al.*, 1998). Como mencionado no Capítulo 2, o cDNA J06G-36 de 493 pb é expresso principalmente no estágio cotiledonar de desenvolvimento do embrião. Este cDNA apresentou similaridade com os genes de vicilina 7S descritos para angiospermas e gimnospermas (Capítulo 2, Tabela 8). A sequência deduzida de aminoácidos deste cDNA corresponde à parte N^o-terminal da proteína AE-21 (Capítulo 1, Figura 6). A AE-21 é uma das principais proteínas de reserva do embrião de *A. angustifolia* (Capítulo 1, Figura 5).

2.1. Isolamento da sequência completa do cDNA J6G-36 através de RACE 3'

Para o isolamento da sequência completa do cDNA J6G-36 foi utilizado o *kit Smart RACE cDNA Amplification* (Clontech) segundo especificações do fabricante. No isolamento do cDNA foram utilizados os *primers* vici C (5'-AGGAGGAGGAGC AAGAGAGGAGACT-3') e vici D (5'-AGCCCGATGCGTTCATGCGTCATC-3') desenhados com o auxílio do programa Gene Runner v. 3.0 (Figura 1A). Os cDNAs obtidos por RACE, Vici A (1400 pb) e Vici B (1200 pb), são mostrados na Figura 1B. O cDNA Vici B foi clonado em pGEM-T *Easy* (Promega) e 2 clones (5 e 6) foram digeridos com *EcoR* I. Os padrões de digestão dos 2 clones foram diferentes (Figura 1C). O cDNA Vici B e as bandas resultantes da digestão dos 2 clones, Vici B 800, Vici B 400 e Vici B 350 (Figura 1C), foram clonadas e seqüenciadas. As seqüências obtidas foram utilizadas na obtenção da sequência completa do gene de vicilina 7S de *A. angustifolia* com o auxílio dos programas Clustal W e *Contig Assembly Program* (C.A.P.). A estratégia para a obtenção da sequência completa do gene é mostrada na Figura 2.

2.2. O gene de vicilina 7S de *A. angustifolia* é semelhante aos descritos para angiospermas e gimnospermas

O cDNA completo da vicilina 7S de *A. angustifolia* possui 1678 pb (Figura 3). O sinal de poliadenilação do gene foi encontrado na posição 1571 (Figura 3). Foram seqüenciados dois transcritos, que apresentaram uma diferença de aproximadamente 103 pb na porção 3' do gene (Figura 2). O ATG inicial dos transcritos foi encontrado na posição 136. A transcrição do gene de vicilina 7S de *A. angustifolia* resulta em um polipeptídeo de 458 aminoácidos. Como mencionado no Capítulo 1, os primeiros 33 aminoácidos deste polipeptídios são clivados para formação da parte N'-terminal da proteína madura.

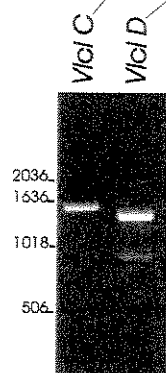
Como mostra a Figura 4, o precursor de 458 aminoácidos apresenta dois domínios bastante conservados. O primeiro vai da posição 39 até 232 e o segundo vai da posição 246 até 450. Entre estes dois domínios existe uma região pouco conservada (posição 231-247) onde a vicilina 7S de *A. angustifolia* poderia ser clivada. Os domínios conservados no

A:

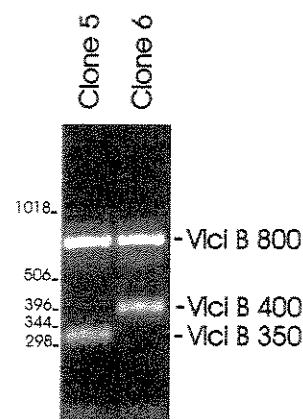
TCGTTCCGCATCGTTCCGCATCGTTCCGCAGAGGTAATGCATCCACGAATTATCTTTCTT
 CATGAATTCAATTCCTTTTGTTCACACTGACTTTCAGAGTTGATCAATGCAGGTGAGGT
 GGTCGAACTGTGCTCATGGCGAATGGTAGGTTCCCGTTGCTTTACTTTGGTGCTTGTTCTT
 GTAGGTTTTGTGTGTGCAATCCGTTGCAATGAGGTTGATCACTACAAGGTTTCATCGACAT
 GGTGAAGAAGAGGGTTCGAAATCCATACGTTTTTAAGGAGGAGGAGCAAGAGAGGAGACTC
 GCTACAGATGCAGGCGAGATCCGAGCGGTTCCCCTCTTCAGAGAGTTCTCCACCCTCGTA
 AGAGATCTCGAGAACTATGAACTCAATTTCTTTACATGAAGCCCGATGCGTTTCATGCGT
 CATCATTACAGCGGTGCAGATCATTTATCGTTGTTCTACAAGGAAAGGCAAGAATTTCAGT
 GTGTGCGGAA CGA

B:

3' RACE



C:



Clonagem

Figura 1. Isolamento por RACE da sequência completa do cDNA da vicilina 7S de *Araucaria angustifolia*. A: Sequência do clone J06G-36 e as sequências dos primers vici C e vici D desenhados. B: Resultado da reação de RACE. O cDNA vici B foi escolhido para clonagem. C: clones 5 e 6, do cDNA vici B, digeridos com *EcoR* I. O marcador de peso molecular 1 Kb (Gibco BRL) a esquerda em p.b.

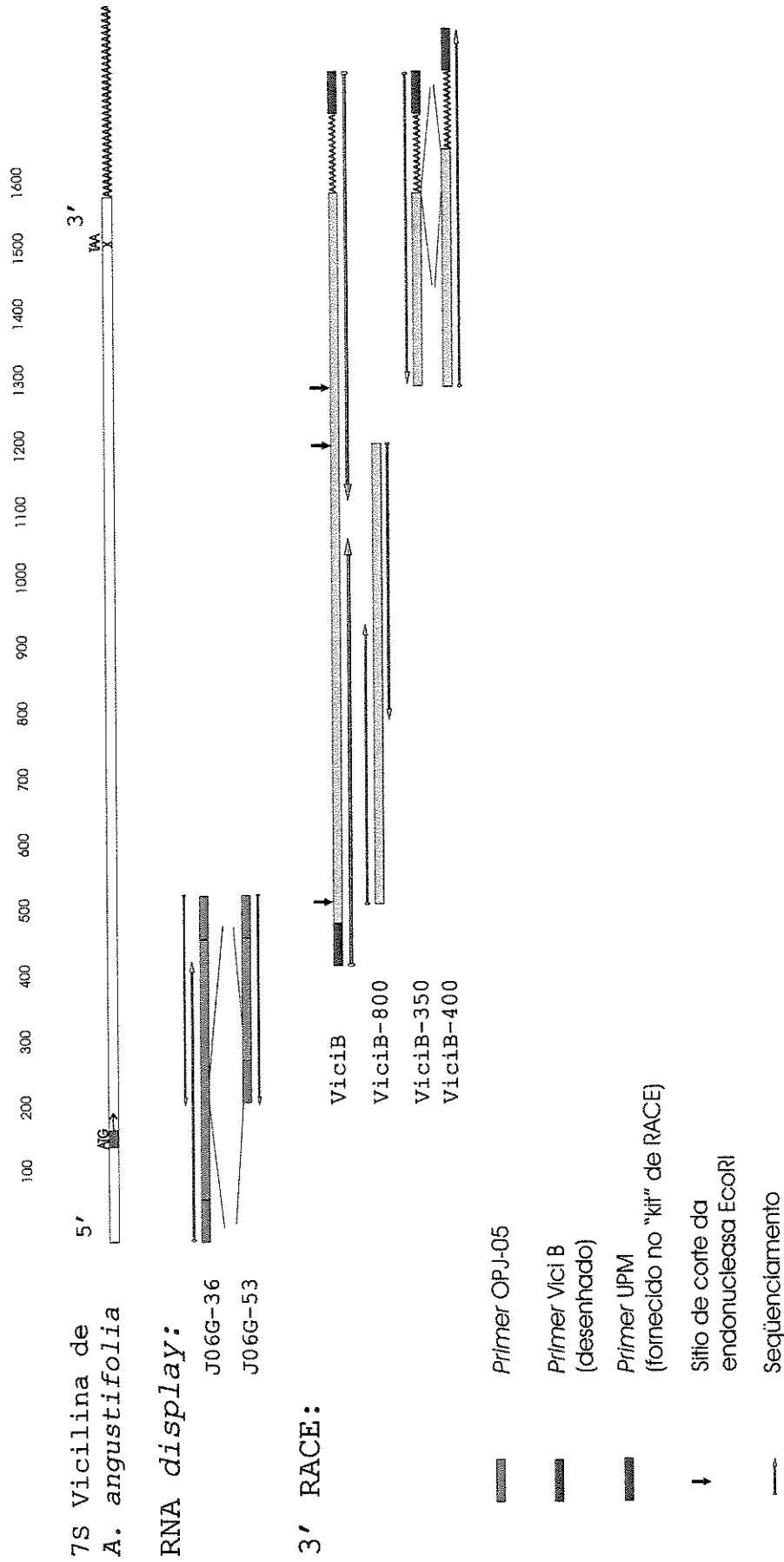


Figura 2. Esquema da montagem da sequência do gene de vicilina 7S. cDNAs J06G-36, e J06G-53 isolados nos experimentos de *display* diferencial. CDNA vici B obtido por RACE e os sub-clones vici B 800, vici B 400 e vici B 350.

alinhamento coincidem com os domínios N'- e C'- terminal propostos na estrutura tridimensional das vicilinas 7S (Laurence *et al.*, 1994). O polipeptídios de 458 aminoácidos poderia sofrer clivagem entre as posições 231-247 (Figura 4) dando assim origem as proteínas de 21 e 28 kDa observadas nos géis bidimensionais mostrados no Capítulo 1.

2.3. As proteínas de 20-21 kDa e 28 kDa da semente de *A. angustifolia* podem ser produtos da clivagem de um precursor comum

O peso molecular das vicilinas 7S descritas na literatura varia entre 36-50 kDa (Shewry & Casey, 1999). Este grupo de proteínas apresenta uma estrutura tridimensional conservada, com domínios N'- e C'-terminal de 200-220 aminoácidos (21-22 kDa) cada (Lawrence *et al.*, 1994; Shewry & Casey, 1999). As maiores vicilinas encontradas, principalmente em angiospermas, apresentam alongamentos nas partes N'- e C'-terminal principalmente. Em gimnospermas, as vicilinas encontradas são “pequenas” (Shewry & Casey, 1999).

Uma característica interessante descrita para as principais proteínas de reserva em *Pinus*, *Picea*, *Pseudotsuga*, e outras coníferas (leguminas 11S) é a clivagem proteolítica do precursor entre os domínios N'- e C'-terminal (Shewry *et al.*, 1995). Esta separação proteolítica dos domínios N' e C' não foi observada no grupo das 7S vicilinas de angiospermas (Shewry & Casey, 1999) embora a clivagem proteolítica de polipeptídios precursores já foram descritas (Yamada *et al.*, 1999; Shewry *et al.*, 1995). Em *Cucurbita*, o precursor (PV100) é clivado em múltiplos sítios Asn-Gln pela VPE (*vacuolar processing protein*), tendo como produto uma proteína de reserva do grupo das vicilinas 7S e vários peptídeos com ação inibidora de tripsina (Yamada *et al.*, 1999).

As vicilinas de 20-21 kDa e 28 kDa de *A. angustifolia* descritas no capítulo 1 podem ser produtos da clivagem do precursor de 44 kDa, se levarmos em consideração que mecanismos semelhantes de processamento de precursores de vicilinas 7S já foram descritos nas sementes maduras de *Cucurbita* (Yamada *et al.*, 1999). Evidências da clivagem do precursor em vicilinas de reserva foram encontradas na semente de *Pinus pinaster* (Allona *et al.*, 1994). Estes autores descreveram o acúmulo de dois polipeptídios de 22 e 27 kDa nestas sementes. O sequenciamento da parte N'-terminal destes

polipeptídios revelou similaridade às seqüências de vicilinas de angiospermas de gimnospermas (Allona *et al.*, 1994).

Neste trabalho foram encontradas vicilinas de 20-21 kDa (AE-21 e AM-20) em *A. angustifolia* (Capítulo I). Como foi demonstrado no seqüenciamento do N^o-terminal, estes polipeptídios representam o domínio N^o-terminal da vicilina. O tamanho destas proteínas coincide com o tamanho proposto para a estrutura clássica dos domínios N^o e C^o das vicilinas 7S (Lawrence *et al.*, 1994; Shewry & Casey, 1999). Assim sendo, é levantada a hipótese da clivagem do precursor viciAra seqüenciado (45-48 kDa estimado em computador) que daria origem as proteínas de 20-21 kDa (domínio N^o-terminal da vicilina). Esta hipótese não descarta a possibilidade do domínio C^o-terminal da vicilina ter peso molecular aparente de 28 kDa.

2.4. Expressão do gene de vicilina 7S no embrião e megagametófito durante o desenvolvimento e germinação da semente

A expressão do gene de vicilina 7S foi investigada durante o desenvolvimento e germinação da semente. Para isto, foi isolado RNA do embrião nos estágios: globular (EG), cotiledonar (EC), maduro (EM) e início da germinação (E45). Foi isolado também, RNA do megagametófito de sementes contendo embrião no estágio cotiledonar (MC) e maduro (MM); e de megagametófito no início da germinação da semente (M45) e radícula de plantas com 1 ano (R). Estes RNAs foram utilizados em experimentos de *Slot blot* e *Northern blot* (Sambrook *et al.*, 1989). Eles foram hibridizados com o fragmento vici B 800 (Figura 2) que representa 45% da parte conservada do gene de vicilina 7S. Como mostra a Figura 5, tanto no *Slot blot* como no *Northern blot* a expressão mais acentuada do gene ocorreu em EC e MC, confirmando a expressão diferencial observada no gel de *display*, diminuindo em EM e MM. Não foi observada expressão em E45, M45 e radícula. No *Northern blot*, foi observada uma banda de aproximadamente 1600 pb, em EC, EM, MC e MM (Figura 5A), que coincide com o tamanho do gene seqüenciado (Figura 3).

```

1 GGAGCCATTG AAGAAACCCT CTCTCAGTTG CCGTGACTTT CAAGCCGTTA ACTCTGTTTT
61 CGTTGCATAG ATTTTAAATA ACAAACCCTA GAGTGGGGTG GAAAATAGTG TGAAACATAG

+1      M S I L S Q S D S L R I R E V W A D
121 GGAAAATGTC AATCCTATCC CAGAGTGATT CTCTCAGGAT TAGGGAAGTA TGGGCCGACA

19 N L E E E I V L I R D I V D D Y P Y V A
181 ACTTGAGGGA GGAGATTGTC CTGATCCGGG ATATTGTGGA TGATTACCCT TATGTGGCCA

40 M D T E F P G I V V R P V G N F K N A S
241 TGGATACCGA GTTCCGGGC ATTGTGGTGA GACCTGTGGG CAACTTTAAG AATGCCTCGG

50 E Y H Y H A L K A N V D M L K L I Q L G
301 AGTACCATTA CCATGCCCTC AAGGCCAATG TGGACATGCT CAAGCTCATT CAGCTGGGCG

70 L T F S D E D G N L P T C G T D K Y C I
361 TCACTTTCTC TGATGAGGAT GGAACCTTC CAACCTGTGG GACAGATAAG TACTGTATCT

90 W Q F N F R E F N L E Q D L Y A N D S I
421 GGCAATTCAA TTTCAGGGAG TTAAATCTTG AGCAGGATCT GTATGCCAAC GATTCCATTG

110 E L L R Q S G I D F K K N N E K G I D S
481 AGCTGTTGAG GCAGAGTGGT ATTGATTTCA AGAAGAACAA TGAGAAGGGT ATCGATTCCG

130 Q V F G E L L M S S G I V L N D T V H W
541 AGGTCTTTGG GGAGCTGCTG ATGTCTTCTG GCATTGTGTT GAATGATACC GTTCACTGGG

150 V T F H S G Y D F G Y L L K L L T C Q N
601 TCACCTTCCA CAGCGGCTAC GACTTTGGGT ACCTGCTGAA GCTCCTAACA TGCCAGAACT

170 L P P T Q V G F F D L I N M Y F P T V Y
661 TGCTCCTAC CCAGGTGGGA TTTTTCGATC TGATCAACAT GTACTTTCCC ACTGTGTATG

190 D I K H L M K F C N S L H G G L N K L A
721 ATATCAAACA CCTCATGAAG TTCTGCAATA GCCTACATGG GGGACTGAAT AAGCTCGCTG

210 E L L D V K R I G V C H Q A G S D S L L
781 AGTTACTGGA TGTGAAGAGA ATTGGGGTGT GCCATCAAGC TGGTTCAGAC AGCTTGCTCA

230 T S C A F R K L R E G F F N G S T E K Y
841 CATCTTGTGC ATTCAGGAAA CTGAGAGAGG GTTCTTTCAA TGGCTCCACT GAAAAGTATG

250 A G V L Y G L G C E A S E N G H
901 CAGGTGTCTT GTATGGGTTG GGTGTGTGAG CTAGTGAAAA TGGTCATTGA TAGCCAGATC

961 ATAGCCTTTC ATCTCTTTGT ACACTTNTTT TCTTCTCTTC GGGGAAAACC CTGGAACCTA
1021 AGTTAATATT GCCAAAAAAA AAAAAAAA

```

Figura 7. Sequência nucleotídica completa do cDNA do fator de transcrição *araCAF 1* de *A. angustifolia*. A sequência nucleotídica e os aminoácidos deduzidos estão numerados a esquerda. Em itálico estão as sequências dos *primers* utilizados na amplificação dos cDNAs no *display* diferencial de mRNA e grifadas, as sequências dos *primers* utilizados na amplificação do experimento de RACE 5'.

1 TCGTTCGGCA TCGTTCGGCA TCGTTCGGCA GAGTAATGC ATCCAGGAT TATCTTTCTT
 61 CATGAATCA TATCTCTTTT GTTCACTG ACITTCAGAG TTGATCAATG CAGGTGAGT
 111 M A N G R F P L L Y L V L V L
 121 GGTGCAATG TGTGCAATG GATGCTAGG TTCCCGTTCG TTACTTTGGT GCTGTGTTCTT
 16 V G F V C A I R C N E V D H Y K V H R H
 181 GTAGGTTTGG TGTGTCAT CCGTTCGCAAT CCGTTCGCAAT GAGTTTGATC ACTACAAGGT TCATCGACAT
 36 G E E E G R N P Y V F K E E E Q E R R L
 241 GGTGAAGAAG AGGTTCGAAA TCCATACGTT TTAAAGGAGG AGGAGCAAGA GAGGAGACTC
 56 A T D A G E I R A V P L F R E F S T L V
 301 GGTACAGATG CAGGCGAGAT CCGAGCGGTT CCGCTCTTCA GAGAGTTCTC CACCTCGTA
 76 R D L E N Y E L N F F H M K P D A F M R
 361 AGAGATCTCG AGAATATGA ACTCAATTC TTTCACATGA AGCCCGATGC GTTCATGCGT
 96 H H Y S G A D H L S F V L Q G K A R I Q
 421 CATCAATACA GCGTTCGAGA TCAATTATCG TTGTGTTCTAC AAGAAAGGC AAGAATTCRG
 116 C V R K E K S M E E N L D R G N L L F I
 481 TGTGTGCGGA AAGAGAAATC GATGGAAGAG AATTTGGACA GAGGAAACTC CTTCTTCATT
 136 P R G S M L S I V N N D P R E E L I M V
 541 CCAGGGGCT CAATGTTGTC CATCGTTAAC AAGGATCCAC GCGAAGAGCT TATCATGGTC
 156 N L L Y N P N P Y R Q R H H E S F Y P V
 601 AATCTGTTAT ACAATCCTAA TCCTTACAGG CAGGACACC AGAGTCTGTT TTACCCAGTC
 176 D M L N A F R R E S L E A A F K V R S E
 661 GATATGTTAA ATGCTTTCAG ACGAGATCT CTGGAAGCCG CCTTTAAGGT TCGAAGTGAA
 196 D I E R M F S R Q D Q R V F R F L S R E
 721 GATATCGAAA GGATGTTTCTAG TAGGCAAGAT CAGCGGGTAT TTTCGATTTCTT AAGCAGAGAA
 216 E R E K I M G R D D S Q L S S F W P L K
 781 GAGAGGGAGA AGATTATGGG GCGAGATGAT AGTCAACTCT CATCAATTCG GCCATTGAAA
 236 T R K G E A E E H N K P F N L E K K D
 841 ACTAGGAAG GAGAAGCTGA GGAGAACAT AACAGCCAT TCAATCTGGA GAAGAAGAT
 256 A K Y S N K N G K Y M E V D S E D Y R P
 901 GCAAAATATT CTAACAGAAA TGGCAAGTAC ATGGAAGTGG ATAGTGAGGA CTATCGTCTT
 276 L K R Q E D R N S M G V G Y T R I E P G
 961 TTGAAGAGAC AAGAAGACAG AAATCCATG GCGTTCGTT ACACAGAGAT TGAACCTGGA
 296 K I T V P Y W H S H A F T I C V V V R G
 1021 AAGATAACTG TACCCTACTG GCACAGTCAT GCTTTCACGA TATGTGTAGT GGTGAGAGGA
 316 P G M L Q M H P R G G K Q Q T E A A K G
 1081 CCGGAATGC TCCAAATGCA TCCCGAGGT GGTAAAGCAGC AGACAGAGGC AGCGAAAGGG
 336 K G E E N Q N G N A R R R E E G E G E F
 1141 AAAGGGGAAG AGAACCCAGAA TGGGAACGCG AGACGGAGAG AAGAAGGAGA AGGAGAATTC
 356 R V S Y R R V E S E L R V G D V F V M P
 1201 AGAGTGTGAT ACAGACGAGT AGAGTCTGAA CTCAGAGTGG GGGAGTGT TGTAAAGCCA
 376 A G H A S V Q M A S S E R L E F L T F F
 1261 GCGGGCATG CATCGGTTCA AATGGCTCA TCCGAACGCC TTGAATTCCT TACCTTCTTC
 396 V N F D R D S G N F L A G N N S V L K Q
 1321 GTCAATTTTG ATCGCGACAG TGGCAACTTC CTCGAGGTA ATAATTCAGT GCTGAACAG
 416 L R E E Q L A A D F G V E R K E M Q R M
 1381 CTGCGGAAG AGCAGTTGGC TGCAGATTC GCGTAGAAA GGAAGANAAT GCAGAGAATG
 436 I G S Q D K A V F V D G P R G S R S L W
 1441 ATCGGATCNC AAGACAAAGC GGTCTTTGTA GATGTTCCGA GGGCTCCG ATCAGTGTGG
 456 S I V
 1501 AGCATCGTAT AATGTGATGT ATGTGACACT ACGTATGTAC GCGGTATTAG TAATCTAAAT
 1561 AAGAAGACCA AATAAATGTT CGGATGAGTA TATCTATGT GAATGATATC AAAATGAGTT
 1621 TTATCGATCA AATCTTGAC TGCNTTGACG CCAGAAAAAA AAAAAAAA AAAAAAAA

Figura 3. Sequência nucleotídica completa do cDNA da vicilina 7S de *A. angustifolia*. A sequência nucleotídica e os aminoácidos deduzidos estão numerados a esquerda. Em negrito, a sequência do N'-terminal de AE-21 (Capítulo 1). Em itálico as sequências dos *primers* utilizados na amplificação dos cDNAs no *display* diferencial de mRNA e na amplificação do RACE 3'. O sítio de poliadenilação encontra-se grifado.

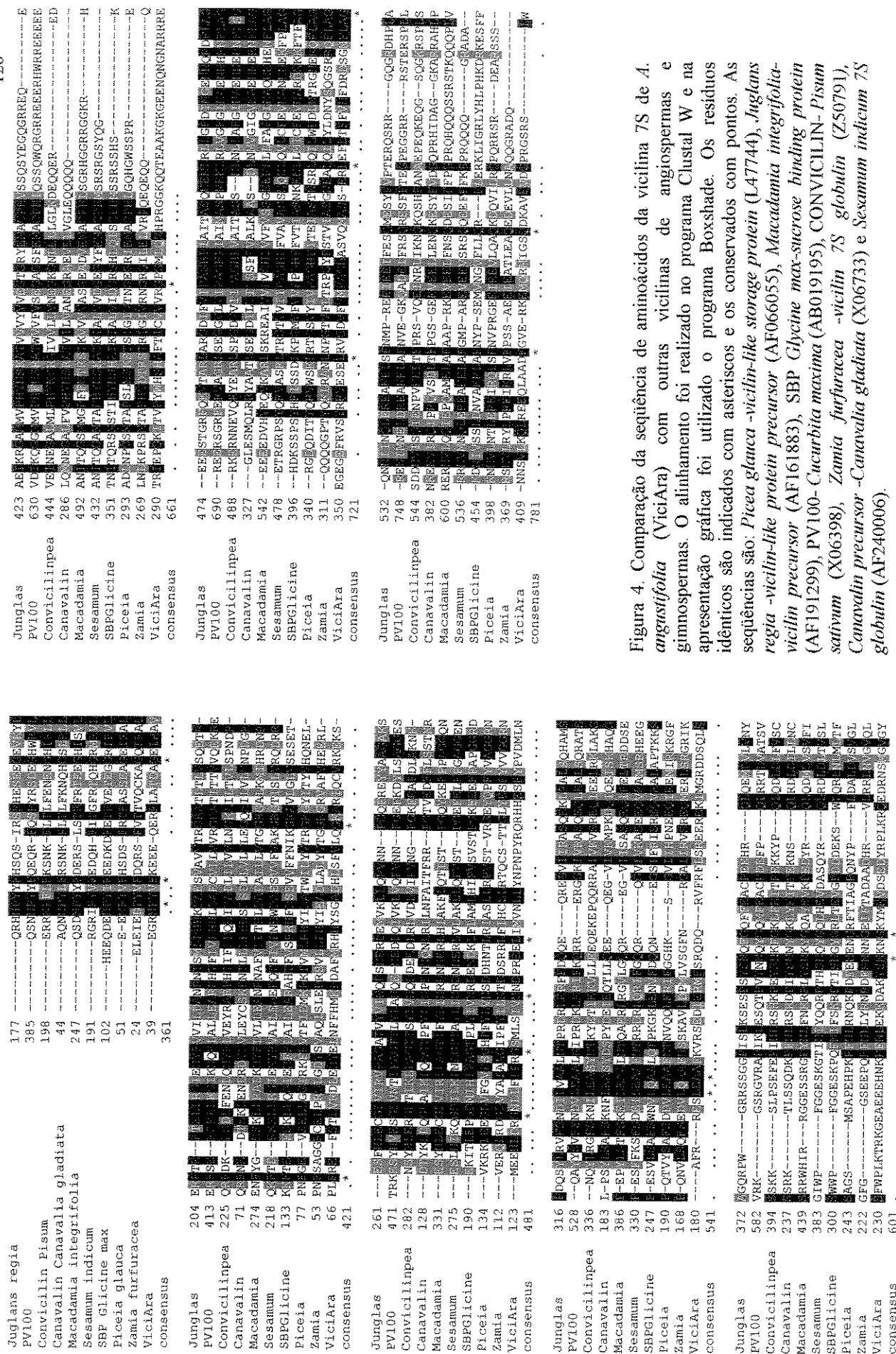
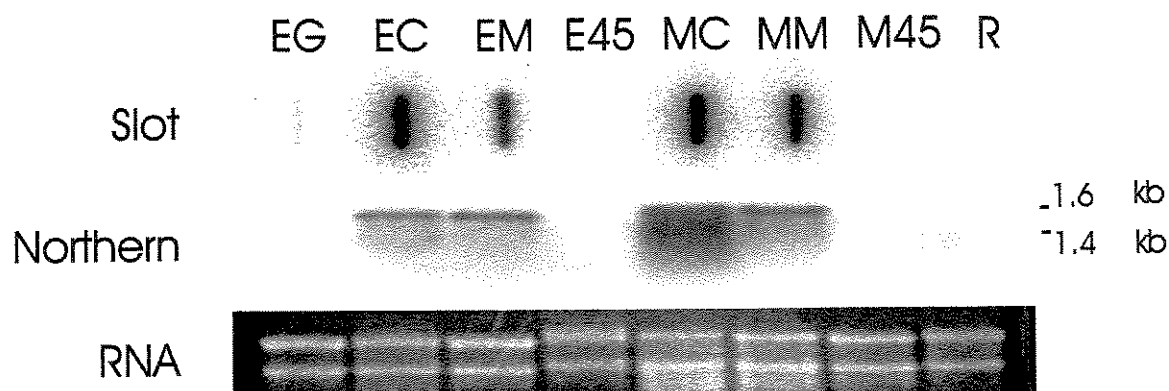


Figura 4. Comparação da sequência de aminoácidos da vicilina 7S de *A. angustifolia* (Viciara) com outras vicilinas de angiospermas e gimnospermas. O alinhamento foi realizado no programa Clustal W e na apresentação gráfica foi utilizado o programa Boxshade. Os resíduos idênticos são indicados com asteriscos e os conservados com pontos. As sequências são: *Picea glauca*-*vicilin-like storage protein* (L47744), *Juglans regia*-*vicilin-like protein precursor* (AF066055), *Macadamia integrifolia*-*vicilin precursor* (AF161883), *SBP Glycine max-sucrose binding protein* (AF191299), *PV100-Cucurbita maxima* (AB019195), *CONVICILIN-Pisum sativum* (X06398), *Zamia furfuracea*-*vicilin 7S globulin* (Z50791), *Canavallin precursor*-*Canavalia gladiata* (X06733) e *Sesamum indicum* 7S globulin (AF240006).

A:



B:

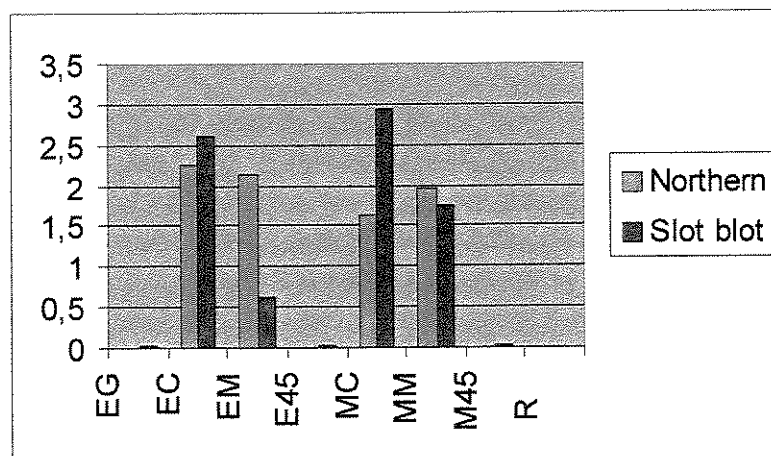


Figura 5. Expressão do gene da vicilina 7S de *A. angustifolia* nos diferentes estágios de desenvolvimento do embrião e megagametófito. A: *Slot blot* e *Northern blot* contendo RNA total do embrião globular (EG), embrião cotiledonar (EC), embrião maduro (EM), embrião no início da germinação (E45), megagametófito da semente contendo o embrião no estágio cotiledonar (MC), estágio maduro (MM), megagametófito no início da maturação (M45) e radícula (R). A expressão da vicilina 7S de *A. angustifolia* é mais acentuada nos estágios cotiledonar e maduro da semente, desaparecendo no início da germinação. B: Comparação dos resultados numéricos obtidos no *Northern blot* e no *Slot blot*.

3. Isolamento da sequência completa do gene *araCAF1* por RACE 5' a partir do cDNA J05G-26

O cDNA J05G-26 (296 pb), expresso principalmente no megagametófito da semente madura, apresentou similaridade com genes do fator associado CCR4 (CAF 1), que são fatores de transcrição descritos em angiospermas, leveduras e mamíferos (Capítulo 2, Tabela 10). O gene completo deste fator de transcrição foi isolado de *A. angustifolia* por RACE (Chenchik *et al.*, 1998) utilizando-se o kit *Smart RACE cDNA amplification* (Clontech) (Figura 6). Como *primer* específico, foi utilizado o CCR4 B (5'-GGAGCCATTGAAGAAACCCTCTCTCAG-3') desenhado com o auxílio do programa Gene Runner v. 3.0 (Figura 6A). Através de RACE, foram obtidos os cDNAs AraCAF1A (850 pb) e AraCAF1B (700 pb) (Figura 6B). Estes cDNAs foram clonados em pGEM-T *Easy* (Promega) e seqüenciados. As seqüências obtidas foram utilizadas na obtenção da seqüência completa com o auxílio dos programas Clustal W e *Contig Assembly Program* (C.A.P.). A estratégia para obtenção da seqüência completa do gene é mostrada na Figura 6C. O gene *araCAF1* seqüenciado possui 1048 pb e codifica para uma proteína de 274 aminoácidos (Figura 7). A seqüência deduzida de aminoácidos deste gene apresentou similaridade com os fatores de transcrição *mCAF 1* e *yCAF 1* (Figura 8) (Draper *et al.*, 1995). Estas proteínas estão associadas ao CCR4 (*carbon catabolite repressor protein*). Segundo Draper *et al.* (1995) elas estariam relacionadas com a repressão da expressão gênica nos processos relacionados ao metabolismo de carbono.

Na maturação da semente de *A. angustifolia*, acumulam-se grandes quantidades de amido, principalmente no megagametófito. Este processo de acúmulo inicia-se nos estágios iniciais do desenvolvimento embrionário (comunicação pessoal do doutorando Vanildo Silveira - BIOCEL, IB, USP; dados histológicos), mas intensifica-se no estágio cotiledonar do desenvolvimento (Capítulo 4). No megagametófito da semente madura de *A. angustifolia*, o amido possivelmente constitui a principal fonte de reservas. Este processo intensivo de síntese de amido é suprimido com a maturação da semente. Possivelmente nesta etapa atua AraCAF1, inibindo a ação de genes que atuam no metabolismo do carbono e tendo sua maior expressão no megagametófito maduro de *A. angustifolia*.

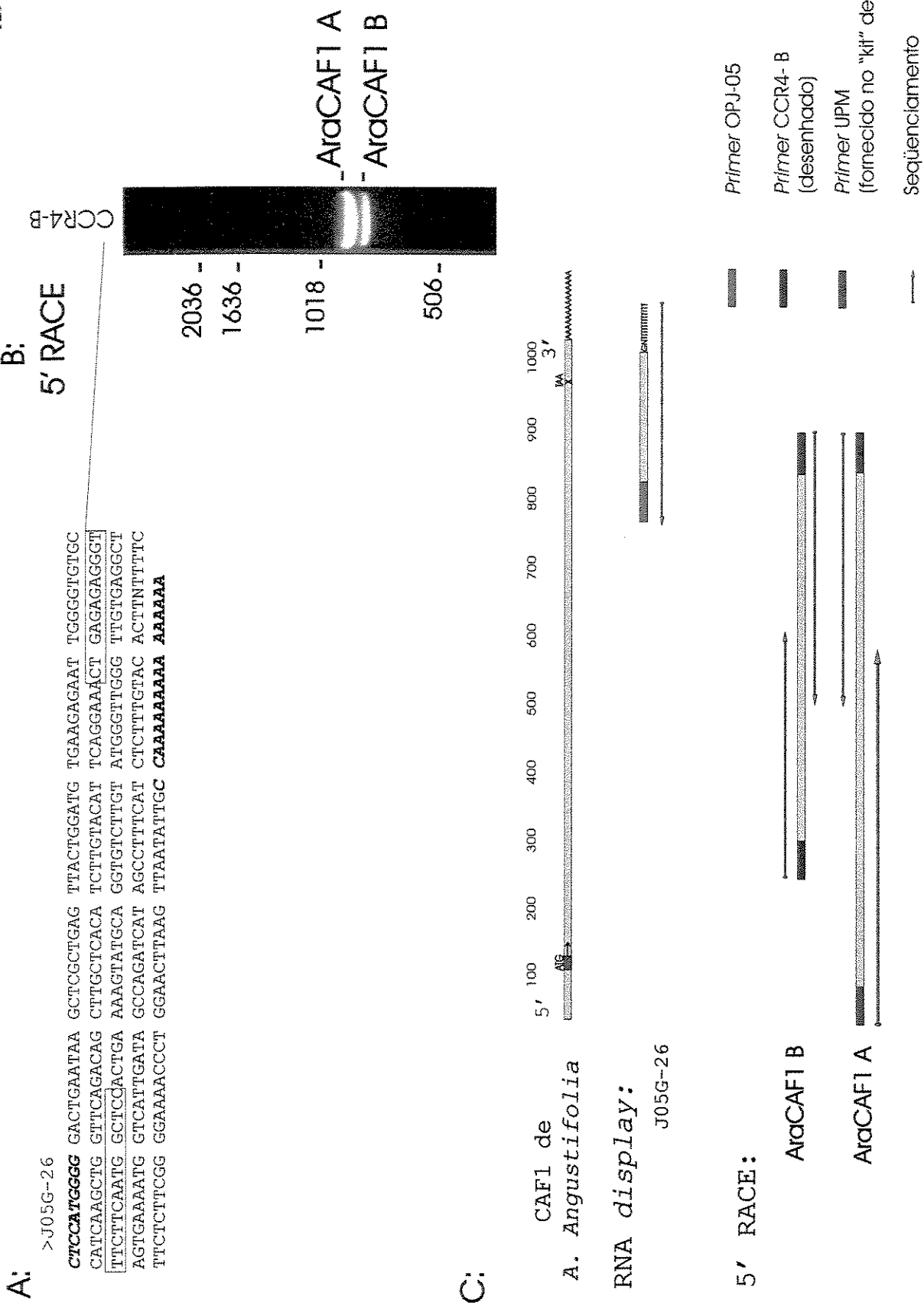


Figura 6. Estratégia do isolamento por RACE da sequência completa do fator de transcrição *araCAF1* de *A. angustifolia*. A: Primer específico CCR B desenhado com a ajuda do programa Gene Runner. B: Resultado da reação de RACE 5'. Os cDNAs *AraCAF1A* e *1B* foram escolhidos para clonagem em pGEM-T *Easy*. O marcador de peso molecular 1 Kb (Gibco BRL) a esquerda em p.b. C: Esquema da montagem dos diferentes cDNAs clonados e sequenciados. O cDNA J5G-26 foi isolado nos experimentos de *display* diferencial de mRNA.

A.thalianaCAF1	1	-----MAETLKEDSIMIREVWDYNLVEEFALIREIVDKHSYLAMDTEFPGVVLM
AraucariaCAF1	1	-----MSILQSDSLRIRREVWADNLEETALIRDIVDDYPKVAMDTFPGIVVR
H.SapiensBTG1	1	-----MKKIRQVIRKYNVAMDTFPGVVAR
MusCAF1	1	-----MPAALVDHSQFICEVWACHLDEEMKKIRQVIRKYNVAMDTFPGVVAR
C.elegansCAF1	1	MASSSSGGAGGAGGASGAPEVKIINVMSNVEEFFARIRGEVEDYPKVAMDTFPGVVAR
consensus	1 * * *
A.thalianaCAF1	50	FVATFRYNNDLNYRTERENVDLLKLIQVGLTFSDENCLNLTCCGTDKFCWQFNFRFNLG
AraucariaCAF1	50	PVGNFNASEYHYHALKANVDMKLKLIQLGLTFSDENCLNLTCCGTDKYCWQFNFRFNLG
H.SapiensBTG1	27	PIGEFRSNADYQYQLRCNVDLLKLIQLGLTFMNEOGGYPPG----TSTWQFNE-KENLT
MusCAF1	50	PIGEFRSNADYQYQLRCNVDLLKLIQLGLTFMNEOGGYPPG----TSTWQFNE-KENLT
C.elegansCAF1	61	PLGTERSKELNYQYQFCNVNMLKLIQVGFAMVNEKGELPPTG----DWQFNE-NESFA
consensus	61	*. * * * * *
A.thalianaCAF1	110	EDIVASESIELLRCGIDFKKNIEKGIQVVRFGELMSSGIVLNDASWVTFHGGYDFGY
AraucariaCAF1	110	QDLYANDSIELLRCSGIDFKKNNEKGIQVFGELMSSGIVLNTVWVTFHGGYDFGY
H.SapiensBTG1	82	EDMYAQDSIELLTSGIQFKKHEEGELETQYFAELLMTSGVVLCEGVNLSFHSGYDFGY
MusCAF1	105	EDMYAQDSIELLTSGIQFKKHEEGELETQYFAELLMTSGVVLCEGVNLSFHSGYDFGY
C.elegansCAF1	116	EDMESHSVEMLRCAGIDETLLQNNGIPTAVFGELTTSGLTDPRTHTLTSSGYDFGY
consensus	121	*. * * * * *
A.thalianaCAF1	170	LVKELTCKELPLKADDFKLLYVFPVYDIKHLMTF----CNGLEGGILNRFAELNGVER
AraucariaCAF1	170	LLKELTCQNLPTPTQVGFEDLINMYFPTVYDIKHLMTF----CNSLEGGILNKLAELLDVKE
H.SapiensBTG1	142	LLKELTNSNLPEEEDDFEILRLFFVYDYVYLMKS----CKNLKGGLQEVAEQLLEER
MusCAF1	165	LLKELTNSNLPEEEDDFEILRLFFVYDYVYLMKS----CKNLKGGLQEVAEQLLEER
C.elegansCAF1	176	LLASITLGLPLKEESTFEMCHTLEPTSEDIKILRTTPNCASAKLKGGLQEVAEQLDVKE
consensus	181	*. * * * * *
A.thalianaCAF1	226	VGCHQAGSDSLTLGSEFRKLRERFPGSTEKYT----GVLYGLGVEDGTTTTVAN----
AraucariaCAF1	226	IGVCHQAGSDSLTSCAFKLRERFENGSTEKYA----GVLYGLGCEASENGH-----
H.SapiensBTG1	198	IGPCHQAGSDSLTGMAFFKMRMFEEHIDDAK--YCGHLYGLGSGSSYVQNGTGNAYE
MusCAF1	221	IGPCHQAGSDSLTGMAFFKMRMFEEHIDDAK--YCGHLYGLGSGSSYVQNGTGNAYE
C.elegansCAF1	236	QGVCHQAGSDALLTAATFFKLRKQFEGDNWNQIAPLICGHMFGGLSSLSLFHSSGSTSRL
consensus	241	*. * * * *

Figura 8. Comparação da sequência de aminoácidos do fator de transcrição *araCAF 1* de *A. angustifolia* com sequências de angiospermas e gimnospermas. O alinhamento foi realizado no programa Clustal W e na apresentação gráfica foi utilizado o programa Boxshade. Os resíduos idênticos são indicados com asteriscos e os conservados com pontos. As sequências são: *A. thaliana* CAF1 (AC011713), *H. sapiens* BTG1 (AF053318), *Mus musculus* CAF1 (U21855) e *C. elegans* CAF 1 (U21854).

4. Isolamento da sequência completa do gene da lipocalina (cDNA J05G-20)

Com o seqüenciamento do cDNA J05G-20, foi obtida a sequência completa do gene de lipocalina. As lipocalinas são proteínas pequenas, extracelulares, com estrutura terciária formada por planos β em forma de canudo, o que permite a ligação e transporte de compostos orgânicos hidrofóbicos pequenos como ácidos graxos e lipídios (Bishop *et al.*, 1995). O cDNA seqüenciado apresenta 705 pb e codifica uma proteína de 225 aminoácidos (Figura 9). A lipocalina de *A. angustifolia* apresentou similaridade com lipocalinas descritas para angiospermas, bactérias e mamíferos (Figura 10).

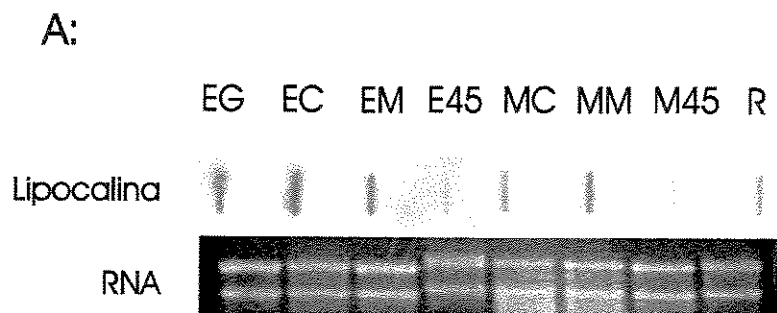
Embora a similaridade encontrada com outras seqüências não seja alta, esta família de proteínas é definida pela sua estrutura terciária conservada (Bicho *et al.*, 1995). Esta lipocalina pode estar envolvida com o transporte e acúmulo de reservas nos cotilédones do embrião (na Figura 4 do Capítulo 2, J05G-20 apresenta expressão acentuada no cotilédone do embrião maduro) e no megagametófito de *A. angustifolia*. Como mostra a Figura 11, a expressão mais acentuada do cDNA de lipocalina ocorre no embrião cotiledonar que é uma fase de deposição de reserva.

```

      M G K S E Q K S L E V V K G L D L E R
1  CTCCATGGGG AAATCGGAGC AAAAGAGCTT GGAGGTGGTG AAGGGCTTGG ATCTGGAGAG
      Y M G R W Y E I A S M P S R F Q P K N G
20 61 GTACATGGGC CGCTGGTACG AGATCGCTTC CATGCCGTCA AGATTTCAGC CCAAGAATGG
      V N T R A T Y A L N D D G T V H V L N E
40 121 CGTTAACACC AGAGCAACCT ACGCTCTCAA CGACGATGGG ACTGTCCACG TGCTCAACGA
      T W V E G K R S S I E G S A Y K G R P Q
60 181 GACGTGGGTG GAGGGAAGA GGAGCTCCAT TGAAGGCAGC GCATACAAAG GTCGACCCCA
      E R R P Y K V D P K S E D A K F K V K F
80 241 AGAGCGAAGA CCATACAAGG TCGACCCCAA GAGCGAAGAC GCCAAATTCA AAGTAAATTT
      F V P P F L P V I P V Y G N Y W V M L L
100 301 CTTTGTTCCT CCATTTCTCC CTGTAATTCC TGTGTATGGC AACTACTGGG TAATGCTTCT
      D Q D Y Q W A L I G E P S L T S L W V L
120 361 TGATCAAGAC TATCAGTGGG CTTTGATCGG AGAACCTTCT CTCACTTCCC TGTGGGTACT
      S R Q P Q L D E E I Y N R L L E K A Q Q
140 421 GAGTAGGCAA CCTCAATTGG ATGAAGAGAT TTATAATCGC TTATTAGAGA AGGCGCAACA
      Q G Y D V G R L R K T S Q T A V P D E T
160 481 GCAGGGGTAT GATGTGGGTC GGCTGAGGAA AACTAGTCAG ACCGCCGTTC CTGATGAGAC
      E A P K D K G F W W I K A L H G E I G A
180 541 TGAGGCCCCA AAAGATAAGG GATTTTGGTG GATCAAGGCT CTCCATGGGG AAATAGGAGC
      L G S K F L R W E A N F C V G K Q I L L
200 601 GTTGGGAAGC AAATTTTTCG GTTGGGAAGC AAATTTTTCG GTTGGGAAGC AGATTCTTCT
      F G S G C S
220 661 TTTTGGGTCT GGTGTAGTT AACATAATTT GCCAAAAAAA AAAAA

```

Figura 9. Sequência nucleotídica completa do cDNA da lipocalina de *A. angustifolia*. A sequência nucleotídica e os aminoácidos deduzidos estão numerados a esquerda. Em negrito estão as sequências dos *primers* utilizados na amplificação do cDNA no *display* diferencial de mRNA.



B:

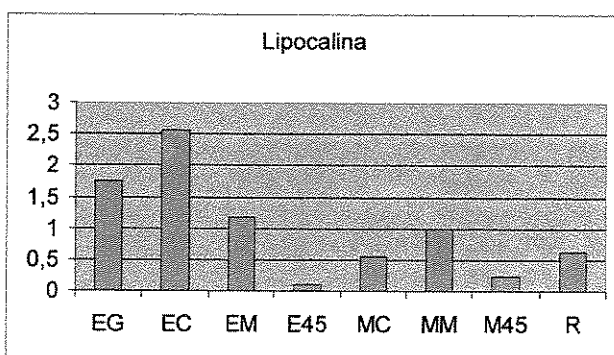


Figura 11. Expressão do gene de lipocalina nos diferentes estágios de desenvolvimento do embrião e megagametófito de *A. angustifolia*. A: *Slot blot* com RNA total do embrião globular (EG), embrião cotiledonar (EC), embrião maduro (EM), embrião no início da germinação (E45), megagametófito no estágio cotiledonar (MC), megagametófito maduro (MM), megagametófito no início da germinação (M45) e radícula (R). B: Apresentação gráfica da quantificação da expressão do gene de lipocalina.

5. Referências Bibliográficas

- Allona I., Casado R., Aragoncillo C. (1994) Biochemical genetics of a 7S globulin-like protein from *Pinus pinaster* seed. *Theor. Appl. Genet.* 88: 454.
- Bishop R. E., Penfold S. S., Frost L. S., Holtje J. V., Weiner J. H. (1995) Stationary phase expression of a novel *Escherichia coli* outer membrane lipoprotein and its relationship with mammalian apolipoprotein D. Implications for the origin of lipocalins. *J Biol. Chem.* 270 (39): 23097.
- Chenchik A., Moqadam F., Siebert P. (1998) A new method for full-length cDNA cloning by PCR. In *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis*. Krieg, P.A. (Ed.) Wiley-Liss, inc.: 273.
- Draper M. P., Salvatore C., Denis C. L. (1995) Identification of a mouse protein whose homolog in *Saccharomyces cerevisiae* is a component of the CCR4 transcriptional regulatory complex. *Mol. Cell Biol.* 15 (7): 3487.
- Lawrence M. C., Izard T., Beuchat M., Blagrove R. J., Colman P. M. (1994) Structure of Phaseolin at 2.2 Å resolution: Implications for common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* 238: 748.
- Sambrook J., Fritsch E. F., Maniatis T. (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Lab. Press. 2nd ed.
- Shewry R., Casey R. (1999) *Seed proteins*. Kluwer Academic Publ.
- Shewry P. R., Napier J. A., Tatham A. S. (1995) Seed storage proteins: Structures and biosynthesis. *The Plant Cell* 7: 945.
- Yamada K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I. (1999) Multiple functional proteins are produced by cleaving Asn-Gln bonds of single precursor by vacuolar processing enzyme. *J. Biol. Chem.* 274 (4): 25663.

CAPÍTULO 4

Considerações finais sobre o desenvolvimento embrionário de *Araucaria angustifolia*

Depois de 4 anos de trabalho, o estudo do desenvolvimento embrionário da *A. angustifolia* revelou-se prolífero em “surpresas”. Os resultados obtidos no estudo das proteínas de reserva e na expressão gênica do embrião zigótico revelam que a *A. angustifolia* apresenta um modelo de desenvolvimento muito diferente ao tido como “padrão” nas coníferas. Para uma melhor compreensão dos dados obtidos neste trabalho, e no intuito de apresentar um modelo básico para esta conífera, é proposto um esquema geral de desenvolvimento para *A. angustifolia*. Este modelo serve de base para futuros estudos dos mecanismos moleculares de desenvolvimento nesta conífera.

1. Desenvolvimento do embrião e acúmulo de reservas na semente de *A. angustifolia*

Embora pouco estudado, o desenvolvimento embrionário em coníferas é um processo com características comuns em várias espécies diferentes, definidas em trabalhos realizados em *Pseudotsuga menziesii* (Owens et al., 1993), *Picea glauca* (Misra & Green, 1990, 1991), *Picea glauca/engelmannii* (Krassowski & Owens, 1993), *Picea abies* (Hakman, 1993), e *Pinus taeda* (Cairney et al., 1999; King & Gifford, 1997). As espécies *Pinus sp.* e *Picea sp.* são as mais estudadas, e consideradas como “modelo”. Nelas, as sementes maduras são pequenas e o desenvolvimento do pró-embrião à semente madura é completado em 70-80 dias (junho – agosto) (Hakman, 1993; Misra, 1994). No processo de maturação da semente, são acumulados principalmente lipídeos, proteínas e amido, tanto no embrião como no megagametófito.

Em *Picea abies*, o acúmulo de reserva é um processo gradativo. Nas etapas iniciais do desenvolvimento acumula-se amido, principalmente no megagametófito. Posteriormente ocorre o acúmulo de lipídeos principalmente nos cotilédones do embrião e por último, ocorre o acúmulo de proteínas no megagametófito (Krasowski & Owens, 1993). Com pequenas variações, este “padrão” de acúmulo de reservas é mantido em *Pseudotsuga* (Owens et al., 1993), *Pinus* (King & Gifford, 1997) e outras coníferas (Misra, 1994).

Na *A. angustifolia*, o desenvolvimento do pró-embrião à semente madura demora 6 meses (dezembro – maio), a semente madura tem peso fresco aproximado de 7 gramas (Figura 1) e acumula amido como principal fonte de reservas (comunicação pessoal do doutorando Vanildo Silveira - BIOCEL, IB, USP; dados histológicos). O “encurtamento”

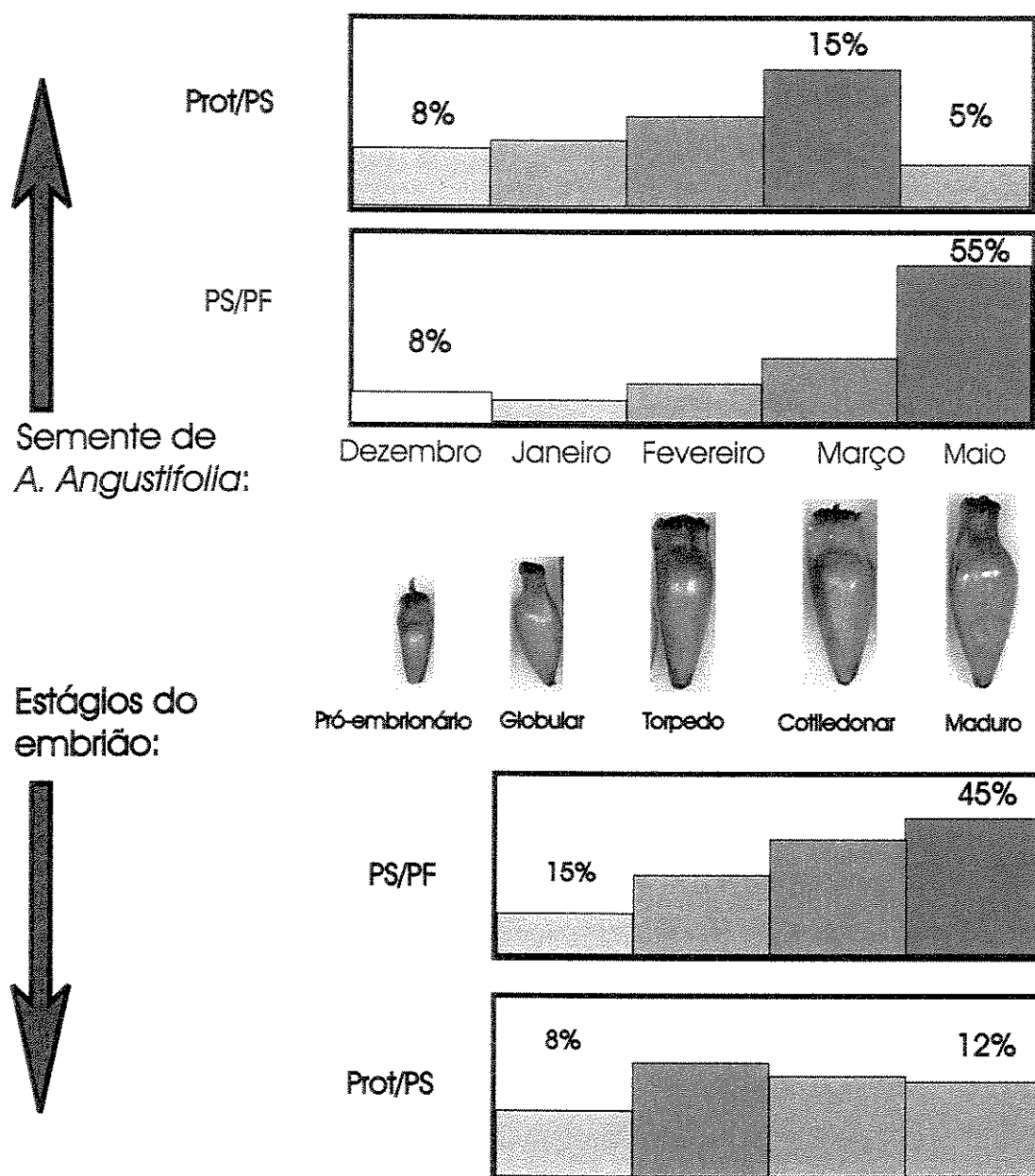


Figura 1. Representação esquemática dos principais processos que ocorrem durante o desenvolvimento e maturação da semente de *A. angustifolia*. Estão representadas as variações do peso seco com relação ao peso fresco (PS/PF) e da quantidade de proteínas com relação ao peso seco (Prot/PS) no desenvolvimento do embrião e da semente. O parâmetro PS/PF é proporcional à desidratação e acúmulo de reservas. No desenvolvimento da semente, o crescimento acelerado nos meses de janeiro e fevereiro não é acompanhado de acúmulo de reservas. O acúmulo de proteínas, lipídios e amido ocorre conjuntamente com a desidratação da semente nas etapas finais da maturação (Abril-Maio). No embrião, este processo ocorre gradativamente desde o estágio globular, até o maduro. O acúmulo de proteínas na semente ocorre gradualmente, mas a quantidade de proteína com relação ao peso fresco da semente cai rapidamente com a maturação devido ao acúmulo de amido no megagametófito da semente madura. No embrião, o acúmulo de proteínas é um processo gradual que acompanha o crescimento dos cotilédones. Adaptado de Fernandez *et al.*, 1997.

do ciclo de desenvolvimento da semente em *Pinus* e *Picea*, com relação à *A. angustifolia*, é considerado uma adaptação evolutiva das coníferas superiores às condições climáticas do norte da Europa e América (Haines & Prakash, 1980). Já o acúmulo de amido como principal fonte de reservas na *A. angustifolia*, é uma mudança metabólica importante, que requer expressão gênica específica. Esta diferença metabólica de acúmulo de reservas deve ser cuidadosamente estudada no futuro.

Embora a principal fonte de reserva da semente madura de *A. angustifolia* seja amido, as proteínas compõem uma importante fonte de reserva, representando respectivamente 12 e 5% do peso seco do embrião e megagametófito da semente madura (Figura 1). Esta quantidade é menor do que os 12 e 10% encontrados respectivamente no megagametófito e embrião de *Pseudotsuga* (Owens *et al.*, 1993), se consideramos que o megagametófito é a principal fonte de reservas na semente das coníferas. Embora não tenham sido feitas dosagem da quantidade de lipídeos na semente de *A. angustifolia*, foram encontradas evidências indiretas do acúmulo destes compostos nos cotilédones do embrião e no megagametófito maduro (dados não mostrados).

Como mostrado na figura 1, a semente de *A. angustifolia* aumenta de tamanho bruscamente no mês de janeiro, coincidindo com o rápido desenvolvimento do embrião nos estágios globular e torpeda. Isto tem a ver possivelmente com o processo de alongamento e vacuolização das células do megagametófito (Fernandez *et al.*, 1997). Este processo é parecido ao descrito em *Pseudotsuga* (Owens *et al.*, 1993) e *Picea* (Krassowski & Owens, 1993). Em *A. angustifolia*, como em outras coníferas, as reservas de lipídeos e proteínas acumulam-se rapidamente no megagametófito só depois do embrião atingir o estágio cotiledonar. Diferentemente de *Picea* e *Pseudotsuga*, a *A. angustifolia* acumula grande quantidade de amido neste estágio do desenvolvimento.

2. As principais proteínas de reserva de *A. angustifolia* são vicilinas 7S

As principais proteínas de reserva descritas em coníferas são leguminas 11S (Misra, 1994; Shewry & Casey, 1999). Estas proteínas acumulam-se em corpos protéicos, principalmente no megagametófito nos estágios cotiledonar e maduro do desenvolvimento. Outras proteínas de reserva foram descritas em coníferas como vicilinas 7S em *Pinus*

pinaster (Allona *et al.*, 1994) e *Picea* (Newton *et al.*, 1992), e albuminas 2S em *Pseudotsuga* (Chattai & Misra, 1998) mas em menor quantidade.

Na *A. angustifolia*, como foi descrito nos capítulos 1 e 3, as principais proteínas de reserva são vicilinas 7S de 20-21 e 28 kDa. Estes polipeptídeos representam mais de 50% das proteínas presentes na semente madura de *A. angustifolia*. Ao que parece, estes polipeptídeos são produto da clivagem de um precursor de 458 aminoácidos que teve o seu cDNA isolado e seqüenciado no capítulo 3. O transcrito seqüenciado no capítulo 3 apresenta 28-31% de similaridade com outras vicilinas 7S descritas para gimnospermas e angiospermas. Estas proteínas acumulam-se desde o estágio globular no embrião, mas no megagametófito, só quando o embrião atinge o estágio cotiledonar de desenvolvimento.

Neste trabalho foram descritas proteínas diméricas unidas por pontes bissulfídicas, de monômeros com peso molecular aproximado de 12 kDa. Elas foram isoladas na fração das albuminas e representaram menos de 5% das proteínas da semente madura. Como foi mostrado no capítulo 1, estas proteínas aparecem no embrião e no megagametófito desde os primeiros estágios de desenvolvimento.

3. Expressão gênica durante o desenvolvimento do embrião zigótico de *A. angustifolia*

Como discutido no capítulo 3, o isolamento, clonagem, seqüenciamento e caracterização de genes diferencialmente expressos durante o desenvolvimento do embrião zigótico de *A. angustifolia*, teve como resultado uma grande quantidade de seqüências sem similares nas bases de dados. Este resultado era esperado, já que no desenvolvimento embrionário, são ativadas uma grande quantidade de genes altamente específicos para este processo. A expressão gênica durante o desenvolvimento da semente foi estudada em *Arabidopsis* (Girke *et al.*, 2000) e *Pinus taeda* (Cairney *et al.*, 1999), mas as seqüências geradas em estudos deste tipo são poucas. A falta de representatividade nas bases de dados de seqüências específicas do desenvolvimento embrionário, junto à pouca quantidade de seqüências de gimnospermas, impossibilita a atribuição de identidade à maior parte dos cDNAs seqüenciados.

As seqüências identificadas estão relacionadas a processos de acúmulo de reservas nos cotilédones do embrião e no megagametófito nas etapas finais do desenvolvimento

(vicilinas 7S), ao transporte e acúmulo de lipídeos no embrião maduro (lipocalina), à inibição da transcrição no embrião maduro (Spt 5), ao processo de alongamento celular no embrião globular (receptor LLR) e à repressão da expressão de genes envolvidos na síntese de amido no megagametófito maduro (AraCAF 1). Não está muito claro o papel da RBP (*Ribonucleic binding protein*) no embrião zigótico maduro. Este grupo de proteínas está associado à regulação da expressão gênica em cloroplastos (Schuster & Gruissem, 1991). Uma hipótese interessante seria o envolvimento desta proteína na ontogênese de plastídios. A expressão destes genes, e a possível função deles no desenvolvimento da *A. angustifolia*, encontra-se resumida na figura 2.

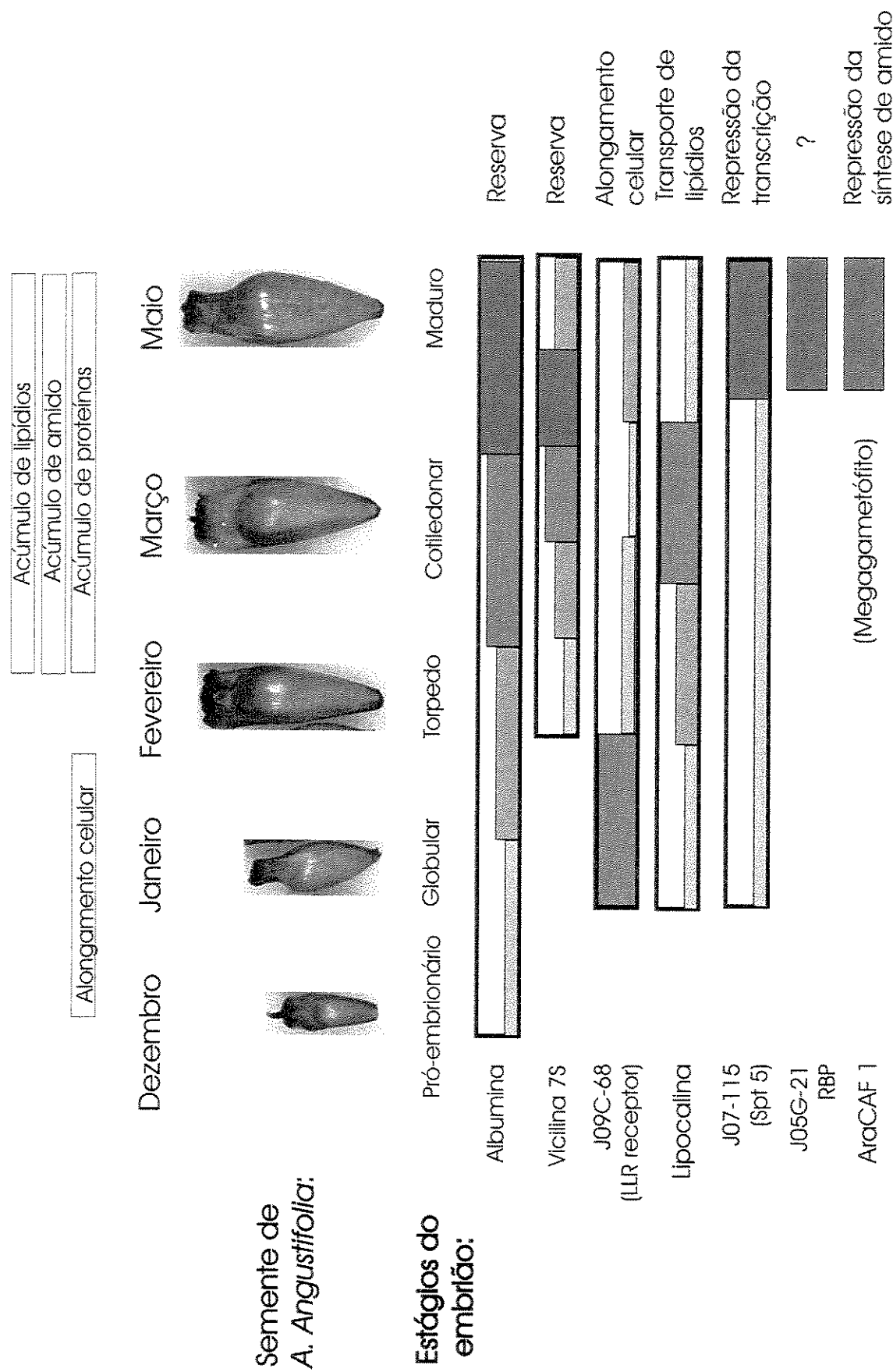


Figura 2. Representação esquemática da expressão das proteínas e genes caracterizados neste trabalho, e a possível função no desenvolvimento e maturação da semente de *A. angustifolia*.

4. Referências Bibliográficas

- Allona I., Casado R., Aragoncillo C. (1994) Biochemical genetics of a 7S globulin-like protein from *Pinus pinaster* seed. *Theor. Appl. Genet.* 88: 454.
- Cairney J., Nanfei X., Pullman G. S., Vincent T., Ciavatta V. T., Johns B. (1999) Natural and somatic embryo development in loblolly pine. *Applied Biochem. and Biotech.* 79: 5.
- Chatthai M., Misra S. (1998) Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziessi* [Mirb] Franco. *Planta* 206: 138.
- Fernandez J.H., Silveira V., Guerra M.P., (1997) Poliembriogénese zigótica e somática em *Araucaria angustifolia*. II Encontro Brasileiro de Biotecnologia Vegetal (REDBIO).
- Girke T., Todd J., Ruuska S., White J., Benning C., Ohlrogge J. (2000) Microarray analysis of developing *Arabidopsis* seeds. *Plant Physiol.* 124: 1570.
- Haines R. J., Prakash N. (1980) Proembryo development and suspensor elongation in *Araucaria Juss.* *Aust. J. Bot.* 28: 511.
- 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.
- King J. E., Gifford D. J. (1997) Amino acid utilization in seed of Loblolly Pine during germination and early seedling growth. *Plant Physiol.* 113: 1125.
- Krassowski M. J., Owens J. N. (1993) The ultrastructural and histochemical post-fertilization megagametophyte and zygotic embryo development of interior spruce. *Can. J. Bot.* 71: 98.
- Misra S. (1994) Conifer zygotic embryogenesis, somatic embryogenesis and seed germination: Biochemical and molecular advances. *Seed Science Res.* 4: 357.
- Misra S., Green M. J. (1990) Developmental gene expression in conifer embryogenesis and germination. I. Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* [Moench] Voss. *Plant Sci.* 68: 163.
- Misra S., Green M.J. (1991) Developmental gene expression in conifer embryogenesis and germination. II. Crystalloid protein synthesis in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss. *Plant Sci.* 78: 61.

- Newton C. H., Flinn B. S., Sutton B. C. S. (1992) Vicilin-like seed storage proteins in the gymnosperm interior spruce (*Picea glauca/engelmannii*). *Plant Mol. Biol.* 20: 315.
- Owens J. N., Morris S. J., Misra S. (1993) The ultrastructural, histochemical, and biochemical development of the post-fertilization megagametophyte and the zygotic embryo of *Pseudotsuga menziesii*. *Can. J. For. Res.* 23: 816.
- Schuster G., Gruissem W. (1991) Chloroplast mRNA 3' end processing requires a nuclear-encoded RNA-binding protein. *EMBO J.* 10 (6): 1493.
- Shewry R., Casey R. (1999) Seed proteins. Kluwer Academic Publ.

CONCLUSÕES

As principais proteínas de reserva de *A. angustifolia* pertencem ao grupo das vicilinas 7S.

As vicilinas 7S de *A. angustifolia* são sintetizadas na forma de um precursor de aproximadamente 458 aminoácidos.

As vicilinas 7S de *A. angustifolia* são expressas principalmente nos estágios cotiledonar e maduro do desenvolvimento do embrião e do megagametófito da semente contendo embrião no estágio cotiledonar e maduro.

Foram isolados 315 cDNAs através de *display* diferencial de mRNA: 99 do embrião globular (EG), 26 do cotiledonar (EC), 121 do maduro (EM), 27 do eixo do embrião maduro (Ee), 11 do cotilédone do embrião maduro (Ec), 10 da região radicular do embrião maduro (Er) e 21 do megagametófito maduro (M).

Dos 315 cDNAs isolados nos experimentos de *display* diferencial de mRNA, 189 tiveram a expressão diferencial confirmada através de experimentos de *slot blot* e *macroarray*.

Setenta e sete cDNAs apresentaram CER (coeficiente de expressão relativa) acima de 2 (23 do EG, 11 do EC, 24 do EM, 7 do M, 3 do Ec e 9 do Ee). Entre estes cDNAs, 39 foram escolhidos para clonagem e seqüenciamento.

Dos 39 cDNAs clonados e seqüenciados, 10 apresentaram similaridade com seqüências disponíveis em bancos de dados, sendo 2 de EG (retroelemento gypsy/Tye3 e *LRR receptor protein kinase*), 1 de EC (Vicilina 7S), 4 do EM (fator de transcrição SPT5, Vicilina 7S e retroelementos gag/pol e Osvaldo), 1 de M (fator de transcrição araCAF1), 1 de Ec (lipocalina) e 1 de Ee (*Ribonucleic acid-binding protein*).

A proteína CAF1 (*Catabolic associative factor*) de *A. angustifolia* (*araCAF1*) apresenta 274 aminoácidos, é codificada por um transcrito de 1048 pb altamente similar aos fatores de transcrição CAF já descritos em angiospermas e mamíferos. Esta proteína é expressa diferencialmente durante o desenvolvimento embrionário.

A lipocalina de *A. angustifolia* apresenta 225 aminoácidos e é expressa principalmente no estágio cotiledonar do desenvolvimento do embrião.