

Universidade Federal do Rio Grande do Sul

Instituto de Biociências

Programa de Pós-Graduação em Genética e Biologia Molecular

**IDENTIFICAÇÃO E ANÁLISE DO PADRÃO DE EXPRESSÃO DE GENES
CODIFICANTES DE ENZIMAS ENVOLVIDAS NA BIOSÍNTESE DE
TRIACILGLICERÍDIOS EM MAMONA (*Ricinus communis* L.)**

Alexandro Cagliari

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul como requisito para a obtenção do Título de Mestre em Genética e Biologia Molecular.

Orientadora: Dra. Márcia Maria Pinheiro Margis

Co-orientador: Dr. Rogério Margis

Porto Alegre, novembro de 2009.

Universidade Federal do Rio Grande do Sul

**IDENTIFICAÇÃO E ANÁLISE DO PADRÃO DE EXPRESSÃO DE GENES
CODIFICANTES DE ENZIMAS ENVOLVIDAS NA BIOSÍNTESE DE
TRIACILGLICERÍDIOS EM MAMONA (*Ricinus communis* L.)**



Alexandro Cagliari

Dissertação de Mestrado

Porto Alegre

2009

Este trabalho foi realizado no Núcleo de Genômica Funcional de Plantas do Departamento de Genética da Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brasil. O trabalho faz parte do “Projeto Estruturante de Agroenergia do Estado do Rio Grande do Sul” sendo subvencionado pela FINEP (Financiadora de Estudos e Projetos) e pela FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul). O mestrando obteve bolsa de estudos do Centro Nacional de Desenvolvimento Científico e Tecnológico (CNPQ).

Dedico esta dissertação a minha mãe, Iraci, e a minha avó, Adele, que me deram todo o suporte afetivo e emocional necessários ao longo desta caminhada.

AGRADECIMENTOS

Aos meus orientadores, Prof^ª. Dr^ª. Márcia Margis e Prof. Dr. Rogério Margis, por terem aceitado o desafio dessa orientação, pelo voto de confiança, paciência, estímulo e amizade. Pelo exemplo como profissionais e como seres humanos.

Às professoras Maria Helena Bodaneze Zanettini, Luciane Passaglia, Eliane Kaltchuk dos Santos e Fernanda Bered pelo apoio científico, convivência e contribuições para o bom funcionamento do laboratório.

À Silvia pelas dicas e conselhos de conduta laboratorial.

Ao secretário do PPGBM Elmo J. Antunes Cardoso e à Ellen pelo suporte burocrático e total disponibilidade.

Às professoras da UnC, Dr^ª Vânia Techio, Dr^ª Elisete Barp e Msc. Celí Araldi pelo incentivo em meus primeiros passos na biologia.

Ao CNPq pelo apoio financeiro.

Ao amigo Guilherme (Loss) pela parceria nas pesquisas com nossa querida mamona.

Aos amigos e co-orientadores esporádicos, Rafael O. (Japonês), Rafael O. (Rafinha) e Lauro (Pingola), pelas dicas, auxílio e ensinamentos (nem todos proveitosos, é claro).

Às amigas Beatriz (Bea) e Marta pelos ensinamentos quanto ao cultivo de embriões, de forma a quase me convencer que a técnica funciona.

Aos amigos de laboratório Roberto (Beto), Diego (Padre), Daniel, João, Ricardo, Rafael R. (Rauber), Marcelo (Pedron) pela amizade e companheirismo.

Às amigas “bromelinas” Camila, Márcia, Jamilla, Manuela e Mirian pelos abacaxis fornecidos.

Às amigas Carol, Fernanda, Andréia, Gisele, Marina por representarem a parte intelectual do grupo de genômica de plantas, compensando assim a parte masculina, limitada por sua essência rude e primitiva.

Aos grandes amigos e irmãos do Daikatana hakushin-gumi himitsu shakai, Gustavo, Emili, Ivanei, Adenilso (Verde), Yuka, Marcelo, Marcela, Marcos (Feijão), Jaime, Eleandro (Gordo), Adilso, Carlos (Tatu), Cleudemar (Tatuzinho), Aline, Giovanni.

Namaste!

Aos amigos de infância Jhonatan (Jhoni), Dailobi (Lobe), Adriano, Marcelo (Bola), Paulo Júnior, Cledson (Gringo), Cleomar (Maraga), Vanderlei (Boiadeiro) pela parceria, farras e histórias acumuladas em anos de convivência.

A todos, que de alguma forma contribuíram e conviveram comigo, meus sinceros agradecimentos pela contribuição, convivência e carinho.

Muito Obrigado!!!

SUMÁRIO

Resumo.....	7
Abstract.....	8
1 INTRODUÇÃO.....	9
1.1 Aspectos históricos.....	9
1.2 Combustíveis alternativos.....	14
1.3 Óleos vegetais.....	17
1.3.1 Ácidos graxos não usuais.....	18
1.3.2 Ácido ricinoléico.....	21
1.4 Biossíntese de óleos vegetais.....	22
1.5 Aplicações industriais dos óleos vegetais.....	28
2 JUSTIFICATIVA.....	33
3 OBJETIVOS.....	37
3.1 Objetivo geral.....	37
3.2 Objetivos específicos.....	37
4 RESULTADOS E DISCUSSÃO.....	38
5 CONSIDERAÇÕES FINAIS.....	77

6 REFERÊNCIAS.....	80
7 ANEXOS.....	90

RESUMO

Sementes de mamona (*Ricinus communis*) contêm grande quantidade de triacilglicerídios (TAGs) ricos em ácido ricinoléico (18:1OH), o qual é produzido pela conversão do ácido oléico (18:1) em ácido ricinoléico em uma reação catalizada pela enzima oleato hidroxilase (FAH12). Como resultado de suas propriedades físico-químicas, o óleo de mamona e seus derivados apresentam inúmeras aplicações industriais. A expressão heteróloga de FAH12 na oleaginosa modelo *Arabidopsis thaliana* resultou na produção de apenas 17% de ácidos graxos hidroxilados no óleo da semente. Este nível é muito inferior ao encontrado no óleo da mamona (90%), sugerindo que genes adicionais e maior conhecimento em relação à biossíntese de óleos em sementes são necessários antes de plantas serem manipuladas geneticamente visando a produção de óleos industrialmente importantes. Neste estudo, realizamos pesquisas no banco de dados genômico da mamona (<http://castorbean.jcvi.org>) com o objetivo de identificar genes da biossíntese de TAGs. No total, vinte e seis genes codificando seis distintas classes de enzimas envolvidas na biossíntese de TAGs foram identificados. A caracterização *in silico* e análises filogenéticas permitiram a identificação de isoformas plastidiais das famílias de enzimas glicerol-3-fosfato aciltransferase e ácido lisofosfatídico aciltransferase, envolvidas na rota procariótica da síntese de lipídios, que formam um *cluster* à parte das isoformas citoplasmáticas, envolvidas na rota eucariótica. Além disso, dois distintos polipeptídios da enzima diacilglicerol aciltransferase foram identificados. Adicionalmente genes codificantes das enzimas ácido fosfatídico fosfatase, fosfolipídio: diacilglicerol aciltransferase e colina fosfotransferase também foram identificados. A análise do padrão de expressão quantitativa demonstrou variação no perfil de expressão ao longo do desenvolvimento da semente de mamona tanto entre genes de diferentes classes de enzimas quanto entre genes de uma mesma classe de enzima. No geral, entretanto, existe uma tendência para um máximo nível de expressão nas fases intermediárias do desenvolvimento da semente, embora isso não tenha sido observado para todos os genes analisados. Os nossos resultados representam uma visão global da regulação da expressão de genes pertencentes a seis famílias enzimáticas envolvidas na biossíntese de TAGs em mamona ao longo do desenvolvimento da semente. Estes resultados podem vir a contribuir na escolha de possíveis genes-alvo para estratégias biotecnológicas que objetivem a produção de óleos desejáveis do ponto de vista nutricional e industrial.

ABSTRACT

Castor bean (*Ricinus communis*) seed oil contains high amounts of ricinoleic acid (18:1OH) rich triacylglycerols (TAGs), which is produced by the conversion of oleic acid (18:1) to ricinoleic acid in a reaction catalyzed by the enzyme oleate hydroxylase (FAH12). As result of its physical and chemical properties, castor oil and its derivatives are used for numerous bio-based products. Heterologous expression of FAH12 in the model oilseed plant *Arabidopsis thaliana* produced only up to 17% of hydroxy fatty acids in seed oils. This level is much lower than that found in castor oil (90%), suggesting that additional genes and significantly more knowledge of seed oil biosynthesis are needed before plants could be engineered to produce industrially important oils. In this study, we survey the Castor Bean Genome Database (<http://castorbean.jcvi.org>) to report the identification of TAG biosynthesis genes. A set of twenty-six genes encoding six distinct classes of enzymes involved in TAGs biosynthesis were identified. Interestingly, *in silico* characterization and phylogenetic analysis allowed us to identify plastidic isoforms of glycerol-3-phosphate acyltransferase and lysophosphatidate acyltransferase enzyme families, involved in prokaryotic lipid biosynthesis pathway, that form a cluster apart from the cytoplasmic isoforms, involved in eukaryotic pathway. Besides, two distinct membrane bound polypeptides of diacylglycerol acyltransferase enzyme were identified. Additionally genes encompasses phosphatidate phosphatase, phospholipid:diacylglycerol acyltransferase and choline-phosphotrasferase were also identified. Quantitative expression pattern analyses demonstrated variations on gene expressions along the castor seed development as much among enzyme classes as into the genes of same enzyme class. In general, however, there is a tendency of maximum expression level at the middle of seed development, although this is not observed for all the genes analyzed. These results represent snapshots of global expression regulation of genes encompassing six enzyme families involved in castor bean TAG biosynthesis along seed development. These results can contributed for the choice of genes that will be target to biotechnological approaches for the production of nutritionally and industrially desirable oils.

1 INTRODUÇÃO

1.1 Aspectos históricos do cultivo da mamona

A mamoneira (*Ricinus communis L.*) é uma planta xerófila e heliófila da família Euphorbiácea, provavelmente originária da África. A origem desta planta ainda é muito controversa, já que existem relatos, em épocas bastante longínquas, de seu cultivo na Ásia e na África. A diversificação de um grande número de variedades desta planta, encontradas tanto no continente africano como no asiático, impossibilita qualquer tentativa de estabelecer uma procedência efetiva da mamona. A mamoneira destaca-se entre as euforbiáceas por produzir um óleo glicerídico solúvel em álcool em temperatura ambiente e com outras propriedades singulares, como ser o mais viscoso de todos os óleos vegetais (11 vezes mais viscoso que o óleo de soja, por exemplo). Esta planta apresenta grande variabilidade, com seis subespécies e 25 variedades botânicas, além de inúmeras cultivares comerciais híbridas em todo o mundo, em especial nos principais países produtores: Índia, China, Brasil e Rússia (EMBRAPA, 2004).

A mamoneira é uma planta de hábito arbustivo, de portes diversos, com variadas colorações de caule, folhas e racemos (cachos), podendo ou não possuir cera no caule e pecíolo (Figura 1). As sementes apresentam-se com tamanho e formato variados, além de grande variabilidade de coloração (Figura 2) (Chiaradia, 2005).

A importância da mamona é ressaltada na obra *Historiarum Mundi*, de Plínio, conhecida há 1.900 anos. Neste documento encontra-se o seguinte trecho no qual são descritas as qualidades do óleo da mamona:

"O óleo de mamona bebe-se com igual quantidade de água morna para purgar o corpo. Diz-se particularmente que purga os intestinos. É útil nas moléstias das articulações, em todas as inflamações ao

ouvido e às assaduras. Com a cinza de peixe é usada na cura da sarna. Dá boa cor à pele e faz nascer o cabelo abundantemente. Nenhum animal é capaz de comer a semente da qual se extrai o óleo. Das bagas fazem-se lâmpadas de particular claridade. As folhas cozidas no vinho, se usadas como emplastro e açafão nas inflamações e, colocados por três dias no rosto, por três dias o purgam"

Plínio



Figura 1. Variabilidade morfológica (inflorescência) em diferentes variedades de mamona.

Foto: EMBRAPA-Pelotas-RS.

Em nosso país, esta planta recebe diversas denominações tais como mamoneira, mamona, rícino, carrapateira, bafureira, baga e palma-criste; na Inglaterra e Estados Unidos é conhecida pelo nome de "*castor bean*" e "*castor seed*". Segundo autores clássicos, o uso da mamona já era conhecido na época dos antigos egípcios, que a apreciavam como planta milagrosa, sendo igualmente utilizada na Índia para os mais diversos fins, sobretudo na medicina popular. No Brasil, sua introdução se deu durante a colonização portuguesa com a finalidade de utilizar seu óleo para iluminação e para

lubrificação dos eixos das carroças. Seu cultivo, historicamente, foi uma das opções para o semi-árido Nordeste, juntamente com o algodão. Nas décadas de 40 a 60, o Ceará chegou a plantar áreas significativas, variando de 30.000 a 57.000 hectares. Contudo, a partir do início da década de 80, ocorreu um contínuo declínio da área cultivada, resultando na desativação de indústrias beneficiadoras e na redução da importância da atividade na economia regional (Vital *et al.*, 2005).



Figura 2. Variabilidade morfológica em sementes de diferentes variedades de mamona.

Foto: EMBRAPA-Pelotas-RS.

Posteriormente, o cultivo da mamoneira desenvolveu-se nas regiões Sudeste e Sul e revitalizou-se no Nordeste do Brasil. Nas regiões Sudeste e Sul, para garantir a competitividade com outros produtos concorrentes, foi necessário o desenvolvimento de técnicas que facilitassem a mecanização da colheita e o desenvolvimento de variedades

mais rentáveis. Deste modo, foram obtidas variedades de baixo porte e indeiscentes, cujos frutos têm maturação sincronizada, o que permite colheita única. Segundo dados fornecidos pelo Instituto Brasileiro de Geografia e Estatística – IBGE - (2004), a produção brasileira de mamona está concentrada na Região Nordeste, especialmente no Estado da Bahia. Nesta região, a maioria dos cultivos é feito com base em cultivares locais, resultantes da miscigenação de variedades com frutos deiscentes, pouca ou nenhuma resistência a pragas e doenças, e baixa produtividade (Coelho, 1979; Kouri *et al.*, 2004).

Como já mencionado, o Estado da Bahia – principalmente o município de Irecê – é o maior produtor nacional, com mais de 90% da produção. Com exclusão total de plantas daninhas, pragas e doenças, e com um sistema de irrigação com fertilização, a mamoneira pode atingir seu ápice de produção de 8.500 kg de bagas por hectare plantado. Embora o potencial teórico seja de 10.000 kg/ha de bagas, a produção na Bahia, consorciado com feijão de ciclo curto, alcança apenas 10% deste valor, com picos de 1200 kg/ha. Esta produção encontra-se dentro da média mundial que é baixa, menos de 650 kg/ha. Pode-se justificar, em parte, a baixa produtividade ao sistema de agricultura familiar adotado. Este se caracteriza pelo baixo uso de insumos, sobretudo os mais modernos, como fertilizantes inorgânicos concentrados, herbicidas e inseticidas, entre outros. O uso de máquinas é pequeno, exceto no preparo do solo, em que parte dos produtores usa o trator, mas a maioria, mais de 60%, utiliza o cultivador à tração animal para o preparo do solo. Alguns produtores ainda usam uma mistura de sementes de cultivares e de tipos locais na mesma área, o que leva a redução de produtividade (EMBRAPA, 2004).

Hoje o Brasil é o terceiro produtor mundial de bagas e é o segundo exportador mundial de óleo de mamona, e poderia a vir a ser o primeiro produtor em pouco tempo, caso mantivesse a produtividade média maior que a do maior produtor atual que é a Índia. No entanto, nos últimos anos a produção brasileira sofreu forte queda, perdendo mercado exterior. A diminuição da produção no Brasil deve-se a uma série de fatores estruturais que levaram a uma redução significativa na área plantada e no volume de produção. Todavia, a mamona ainda representa uma alternativa interessante para a fabricação de biodiesel e novos investimentos vêm permitindo a recuperação da cultura. Embora seu custo de produção ainda seja elevado, a mamona surge como alternativa viável em função da forte demanda por seu óleo (EMBRAPA, 2004; Amorim, 2005; Chiaradia, 2005).

A mamona é hoje considerada uma planta de excelente potencial biotecnológico para a produção de biodiesel. Por isso, o Governo Federal criou dois programas indicando que o fruto de mamona deva ser a principal oleaginosa, no processo de substituição do diesel brasileiro: o "Programa Brasileiro de Desenvolvimento Tecnológico do Biodiesel - PROBIODIESEL", coordenado pelo Ministério de Ciência e Tecnologia e o "Programa Combustível Verde", coordenado pelo Ministério de Minas e Energia. Entretanto, apesar desta sinalização positiva, a cultura ainda enfrenta problemas na base da cadeia produtiva, principalmente na produção, distribuição e adoção de sementes de genótipos melhorados, com alto potencial produtivo (Barreto, 2004). Isso faz com que estudos mais aprofundados se tornem essenciais para o correto delineamento de estratégias efetivas de melhoria da cultura de mamona no Brasil (Santos-Filho *et al.*, 2004).

1.2 Combustíveis alternativos

Na era moderna, o petróleo exerce grande influência em inúmeros aspectos da sociedade. O consumo mundial de petróleo em 2005 foi de aproximadamente 4.200 milhões de toneladas, o que representa cerca de 85 milhões de barris de óleo por dia. Cerca de 90% deste óleo é usado como combustível para sustentar os setores energético e de transportes, enquanto os 10% restantes são usados na indústria petroquímica para a produção de plásticos, fibras e outros componentes de uso diário (Dyer and Mullen, 2008).

Entretanto, a diminuição na produção de óleo, a volatilidade dos preços e a necessidade de independência energética associada à preocupação com o aumento nos níveis de dióxido de carbono (CO_2) tem intensificado a busca por fontes renováveis de energia que possam substituir nosso atual consumo e dependência por combustíveis fósseis. A necessidade de desenvolvimento de uma fonte mais segura de energia tem levado muitos países a explorar fontes alternativas de combustíveis e matéria-prima industrial, dentre as quais se destaca a busca por combustíveis derivados de óleos vegetais (Durrett *et al.*, 2008; Dyer and Mullen, 2008).

O biodiesel, produzido a partir de óleo vegetal, vem atraindo cada vez mais atenção e interesse, apresentando-se como fonte renovável e ambientalmente mais “amigável” de combustível (Vasudevan and Briggs, 2008).

Nesse contexto, uma das aplicações de grande valor econômico do óleo de mamona é a sua utilização como biocombustível alternativo (Durrett *et al.*, 2008; Dyer and Mullen, 2008; Vasudevan and Briggs, 2008).

O óleo vegetal é o bioproduto quimicamente mais similar ao óleo fóssil e, portanto, apresenta grande potencial para substituir o petróleo na indústria química (Abbadi *et al.*, 2004).

Com o acordo estabelecido pelo protocolo de Kyoto e nas diretrizes do MDL (Mecanismo de Desenvolvimento Limpo), a produção de mamona e biodiesel além de causar ganhos decorrentes da redução da emissão de CO₂ por queima de um combustível mais limpo (esta redução está estimada em 2,5 toneladas de CO₂ por tonelada de biodiesel) pode gerar renda pelo seqüestro de carbono da atmosfera. Por exemplo, uma plantação de mamona pode absorver CO₂ na atmosfera na proporção de 33 toneladas de gás carbônico por hectare de mamona (Savy-Filho, 2005). Para a substituição de 1% do diesel mineral são necessárias 348 mil toneladas de mamona o que corresponde a 740 mil hectares de mamona, ou seja, anualmente poderiam ser absorvidos mais de 6 milhões de toneladas de carbono pela lavoura de mamona (Biodiesel, 2008) (www.biodieseleco.org.br).

Já está bem caracterizado que óleos armazenados em sementes apresentam cadeias carbônicas quimicamente semelhantes às de origem fóssil (petróleo), representando uma fonte renovável de combustível e matéria-prima para a indústria química. Os benefícios da produção industrial de óleos vegetais incluem decréscimo na dependência de importação de óleo cru, criação de uma nova fonte de derivados químicos que não são facilmente obtidos pelas tradicionais rotas sintéticas de produção e a expansão da competitividade e viabilidade do setor agrícola na sociedade (Dyer and Mullen, 2008).

A busca por variedades com excelente produtividade para obtenção de biodiesel é premente, pois, além da demanda do mercado, a agricultura atual exige o uso racional da área agriculturável e menores custos de produção.

1.3 Óleos vegetais

Lipídios de reserva em sementes são compostos, principalmente, por três cadeias de ácidos graxos esterificadas em uma molécula de glicerol (figura 3), estrutura que passa então a ser denominada triacilglicerídeo (TAG) (Durrett *et al.*, 2008).

TAGs são acumulados durante a fase de maturação do embrião ou endosperma. Durante a germinação da semente, TAGs são então catabolizados, fornecendo a energia necessária para o desenvolvimento e estabelecimento da plântula (Voelker and Kinney, 2001).

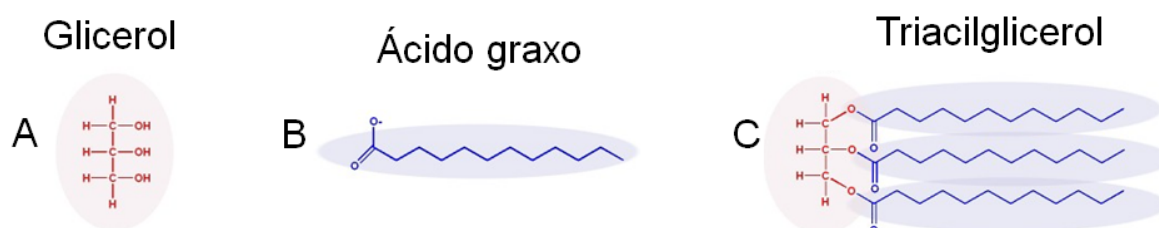


Figura 3. Triacilglicerol. Uma molécula de triacilglicerol (C) é formada por três ácidos graxos (B) esterificados em uma molécula de glicerol (A).

Diversos estudos têm demonstrado que o conteúdo de óleo em sementes é controlado por inúmeras QTLs (*quantitative trait loci*), 10 QTLs em *Brassica napus* (Delourme, 2006) e 4 QTLs em *Arabidopsis thaliana* (Hobbs, 2004), por exemplo, e também por interações entre o genótipo e o ambiente. Porém, interessante, a

manipulação genética de genes-alvo tem resultado em linhagens transgênicas com substancial aumento em seu teor de óleo (Weselake *et al.*, 2009).

Atualmente, grande parte da produção mundial de óleos vegetais vem sendo utilizada com propósitos industriais. As culturas oleaginosas apresentam-se como a maior fonte agrícola de *commodities*, com uma produção de mais de 380 milhões de toneladas de óleo produzido no ano de 2005. Cerca de 85% do total de óleo vegetal extraído é utilizado para fins alimentícios, enquanto cerca de 15% é empregado na indústria de lubrificantes, tintas, plásticos, biodiesel e outros subprodutos. Entretanto, um significativo problema encontrado durante o uso de óleos vegetais para propósitos industriais é que os componentes dos ácidos graxos não são especialmente reativos (Dyer and Mullen, 2008).

Esta baixa reatividade ocorre porque as espécies cultivadas vêm sendo domesticadas visando à produção de óleos de sementes para uso na alimentação e não para uso industrial. Conseqüentemente, a composição destes ácidos graxos atende, primeiramente, às necessidades nutricionais humanas. Os óleos de sementes, portanto, necessitam de tratamento químico antes de serem utilizados industrialmente, o que contribui significativamente para o aumento dos custos e impactos ambientais resultantes de sua utilização (Metzger and Bornscheuer, 2006).

A compreensão de como muitas espécies de plantas ditas “não-domesticadas” acumulam ácidos graxos estruturalmente diferentes na composição de seus óleos, muitos dos quais, inclusive, aptos para uso direto na indústria, é um tema de grande importância. A maioria destes óleos é enriquecida em um único ácido graxo, “não usual”, como no caso da mamona, (onde o ácido ricinoléico representa 90% da composição

total de ácidos graxos da semente), o que simplifica seu processamento e, em muitos casos, permite o uso direto do óleo em aplicações industriais (Dyer and Mullen, 2008).

1.3.1 Ácidos graxos não usuais

Ácidos graxos “não usuais” podem ser definidos como aqueles que possuem estruturas químicas que diferem das apresentadas pelos principais ácidos graxos constituintes de óleos armazenados em sementes da maioria das culturas convencionais: ácidos palmítico (16:0), esteárico (18:0), oléico (18:1), linoléico (18:2) e linolênico (18:3). Muitos ácidos graxos “não usuais” apresentam propriedades que são desejáveis do ponto de vista industrial, incluindo comprimento da cadeia carbônica, posição e configuração de duplas ligações e/ou presença de grupos funcionais (como grupos hidroxil ou epóxi) no ácido graxo (Jaworski and Cahoon, 2003).

Infelizmente, o principal problema associado ao uso destes óleos para propósitos industriais é que as espécies de plantas que produzem estes tipos de óleos industrialmente importantes, geralmente, apresentam características agrônômicas limitantes, tais como sementes pequenas, baixa produtividade, toxicidade ou área de distribuição geográfica limitada. Tais características dificultam a obtenção de óleo suficiente, de uma maneira competitiva e financeiramente viável, para seu uso na indústria. Devido a isto, um dos principais objetivos da pesquisa nesta área, nos últimos anos, tem sido estudar estas espécies selvagens e identificar genes da rota de biossíntese de ácidos graxos, e então transferi-los, com o auxílio de ferramentas biotecnológicas, para espécies cultivadas e altamente produtivas (Jaworski and Cahoon, 2003; Cahoon *et al.*, 2007; Napier, 2007).

As plantas possuem uma grande gama de diferentes ácidos graxos, dentre os quais ácidos graxos hidroxilados, epoxilados e acetilenados (Tabela 1). A heterogeneidade apresentada na composição destes ácidos graxos é, entretanto, restrita aos TAGs armazenados em lipídios de sementes. Portanto semelhante ao que ocorre durante o acúmulo de ácidos graxos não usuais na forma de TAGs em procaríotos, estes TAGs devem excluídos da composição da bicamada lipídica dos vegetais, evitando assim disfunções na membrana ocasionadas pela presença de tais ácidos graxos não usuais (Dyer and Mullen, 2008).

Tabela 1. Composição do óleo (porcentagem de cada ácido graxo) em espécies cultivadas e selvagens. As propriedades dos ácidos graxos não usuais presentes nas espécies não domesticadas são fornecidas entre parênteses na última coluna da tabela. (Dyer and Mullen, 2008).

Plantas	Ácido graxo											Nota
	8:0	10:0	12:0	14:0	16:0	18:0	18:1	18:2	18:3	20:1	Não usual	
Grandes culturas												
<i>Óleo de palma</i>				5	36	2	50	8				
<i>Óleo de soja</i>					11	4	23	54	8			
<i>Óleo de canola</i>					4	2	60	21	10	1		Alto teor de ácido oléico
<i>Óleo de girassol</i>					7	5	19	68				Alto teor de ácido linoléico
<i>Óleo de linhaça</i>					6	2	19	24	47			Alto teor de ácido linoléico
<i>Óleo de côco</i>	7	7	48	18	9	3	6	2				Alto teor de ácido láurico
Pequenas culturas ou espécies selvagens												
<i>Borago officinalis</i>					10	4	16	38	1	4	23	y-linolênico (isômero de α-linolênico)
<i>Echium plantagineum</i>					6	3	14	13	33			y-linolênico (isômero de α-linolênico)
<i>Cuphea hookeriana</i>	50	25	4	1	7		4	5				Ácido caprílico (ácido graxo de cadeia média)
<i>Cuphea pulcherrima</i>	96	2			1			1				Ácido caprílico (ácido graxo de cadeia média)
<i>Cuphea lanceolata</i>		83	2	2	3		3	5				Ácido caprílico (ácido graxo de cadeia média)
<i>Ricinus communis</i>					1	1	3	4			89	Ácido ricinoléico (ácido graxo hidroxilado)
<i>Coriandrum sativum</i>					4	3	8	17	1		66	Ácido petroselinico (isômero de ácido oléico)
<i>Crepis alpina</i>				1	4	1	2	18			74	Crepénico (ácido graxo com tripla ligação)
<i>Vernonia galamensis</i>					3	3	15	9			67	Ácido vernólico (ácido graxo epoxilado)
<i>Momordica charantia</i>					2	17	15	14	9	7	57	Ácido eleosteárico (ácido graxo conjugado)
<i>Brassica napus</i>					5	1					45	Ácido erúxico (ácido graxo de cadeia longa)
<i>Crambe abyssinica</i>					2	1					56	Ácido erúxico (ácido graxo de cadeia longa)

Muitos dos ácidos graxos não usuais encontrados em óleos de sementes de diversas plantas são sintetizados no citoplasma, usando como precursores ácidos

graxos comuns, exportados dos plastídios. Estas reações incluem a elongação dos ácidos graxos, formando ácidos graxos de cadeia longa (com mais de 18 carbonos). Tais etapas de elongação são realizadas enquanto um grupo acila é ligado à Coenzima A (CoA) em uma seqüência de reações análoga à síntese *de novo* de ácidos graxos nos plastídios, exceto que todas as enzimas extra-plastidiais encontram-se associadas à membrana, enquanto nos plastídios elas são solúveis e utilizam *Acyl-Acyl Carrier Protein* (Acil-ACP) como substrato (Domergue *et al.*, 2000).

Embora todos os ácidos graxos não usuais conhecidos sejam sintetizados enquanto ligados à fosfatidilcolina (PC), eles são rápida e especificamente removidos deste fosfolípido e canalizados para formação de TAGs, em um processo que se acredita ser realizado por fosfolipases acil-específicas (Stahl *et al.*, 1995), PDAT (fosfolípido: diacilglicerol aciltransferase) (Dahlqvist *et al.*, 2000) ou pela ação reversa da CPT (colina fosfotransferase), que converte fosfadidilcolina em diacilglicerol (um substrato para a subsequente formação de TAGs).

Interessantemente, os níveis de ácidos graxos não usuais presentes na fosfatidilcolina em plantas transgênicas de soja (*Glycine Max (L). Merr.*) e *Arabidopsis (Arabidopsis thaliana)* são substancialmente maiores dos que os observados nas plantas selvagens das quais esses ácidos graxos são originados, o que suporta a idéia de que o pequeno acúmulo destes ácidos graxos em plantas transgênicas (na forma de TAGs) é, primeiramente, limitado por sua ineficiente remoção da PC bem como por sua provável ineficiente passagem pela rota de Kennedy (Singh *et al.*, 2001; Cahoon *et al.*, 2006). A eventual rota para o acúmulo de altos níveis de ácidos graxos não usuais em plantas transgênicas deve, portanto, envolver a descoberta e introdução de genes codificando

formas especializadas de enzimas chaves da síntese de ácidos graxos e seu acúmulo na forma de TAGs.

1.3.2 Ácido ricinoléico

No caso da mamona, o ácido ricinoléico (ácido 12-hidroxi-octadecis-9-enoico: 18:1-OH), o qual possui um grupo funcional hidroxila (OH) no carbono 12 de sua cadeia, é um ácido graxo não usual de interesse econômico (Figura 4) (Dyer and Mullen, 2008). Óleos vegetais ricos em ácido ricinoléico são de particular interesse para a indústria oleoquímica uma vez que o grupo funcional (hidroxila) presente neste óleo diminui os custos gerais da produção de óleo-derivados. Estes custos são reduzidos significativamente quando se utiliza uma fonte natural deste grupo funcional ao invés da necessidade de sua síntese (transformação química) (Bafor *et al.*, 1991). Este ácido graxo é encontrado como constituinte de óleos de reserva em sementes de 12 gêneros dentro de 10 famílias de plantas superiores (van de Loo *et al.*, 1995).

Porém, a única fonte comercial de ácido ricinoléico conhecida é o endosperma de mamona. A presença de ácido ricinoléico em altos níveis (aproximadamente 90% do conteúdo total de ácidos graxos) dá ao óleo de mamona propriedades únicas, agregando-lhe grande valor industrial (Rojas-Barros *et al.*, 2004).

A síntese do ácido ricinoléico em sementes de mamona tem sido estudada *in vitro* em preparações de membranas microsomais de endosperma em desenvolvimento. O hidroxiácido graxo é formado pela hidroxilação do ácido oléico (18:1) esterificado na

posição Sn-2 da fosfatidilcolina (Bafor *et al.*, 1991). A reação é catalisada por uma oleato-12-hidroxilase (FAH12, figura 4) do retículo endoplasmático (RE).

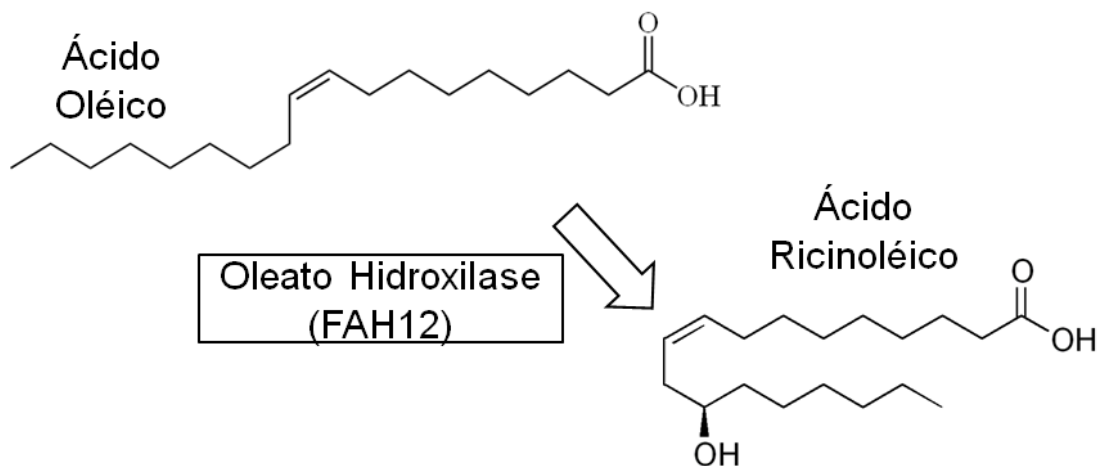


Figura 4. Síntese de ácido ricinoléico. A enzima oleato hidroxilase (FAH12) adiciona um grupamento hidroxila (OH) no ácido oléico originando ácido ricinoléico.

1.4 Biossíntese de Óleos vegetais

A biossíntese de óleos de reserva em sementes dos cinco ácidos graxos mais comuns ocorre em dois compartimentos subcelulares distintos. A síntese *de novo* dos ácidos palmítico (16:0) e esteárico (18:0) e a desaturação do ácido esteárico originando ácido oléico (18:1) ocorre nos plastídios, com o ácido graxo nascente sendo esterificado em uma proteína solúvel denominada Proteína Carreadora de Acil (ACP). Estes três ácidos graxos são então exportados ao citosol onde formam um complexo com coenzima A (CoA) formando Acil-CoA. O ácido oléico é incorporado à fosfatidilcolina (PC) no retículo endoplasmático tornando-se apto a sofrer conversão para ácido linoléico (18:2) e então ácido linolênico (18:3) pela ação de enzimas desaturases substrato-específicas. Uma parte dos ácidos graxos produzidos nos plastídios e modificados ou não no retículo endoplasmático é retida para a formação de componentes estruturais de

membranas e o restante pode, finalmente, ser armazenado na forma de TAGs no endosperma das sementes (Figura 5) (Dyer and Mullen, 2008).

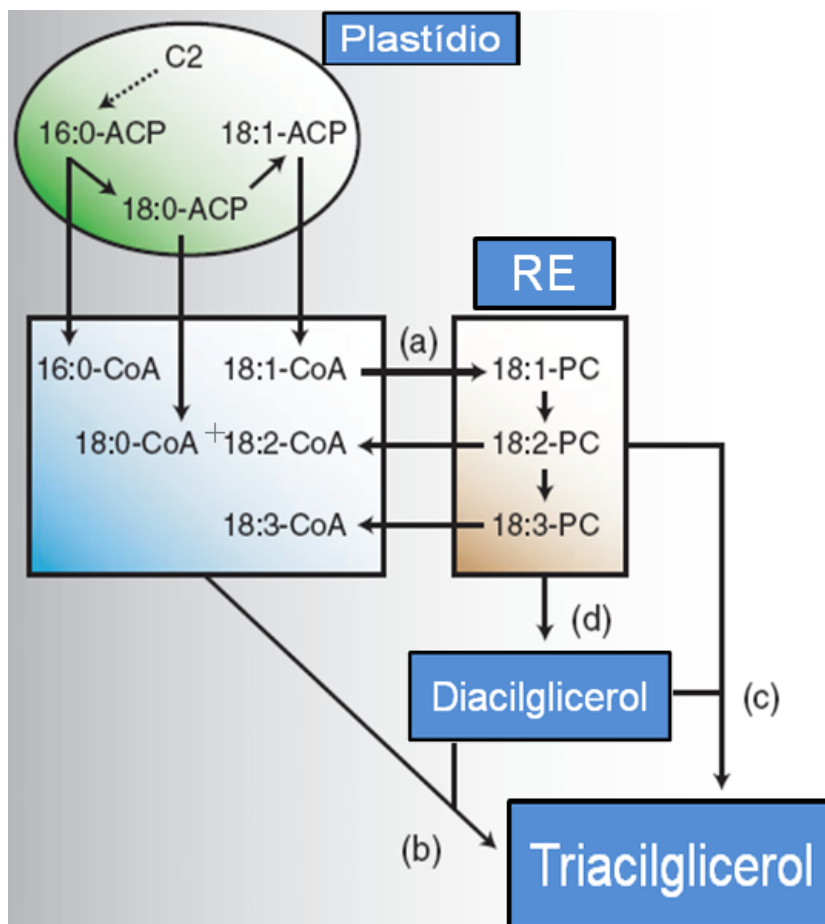


Figura 5. Síntese dos principais ácidos graxos encontrados nos vegetais.

Os primeiros dois ácidos graxos (16:0, 18:0) são produzidos nos plastídios onde também ocorre a reação de desaturação do ácido esteárico originando ácido oléico (18:1). Estes ácidos graxos são, então, exportados para o citoplasma onde são complexados à CoA, formando Acil-CoA. O ácido oléico é transferido para a fosfatidilcolina (PC) no retículo endoplasmático (RE) (a) para a produção de ácido linoléico (18:2) e ácido linolênico (18:3).

Estes ácidos graxos podem ser incorporados à molécula de glicerol, formando TAGs, por três potenciais rotas: (i) a tradicional rota de Kennedy (b), (ii) direta transferência dos ácidos graxos da PC para diacilglicerol (c) para produzir TAG, ou (iii) remoção do grupamento fosfato da fosfatidilcolina produzindo diacilglicerol (d), que é utilizado nas reações b ou c para produzir TAG. Fonte: (Dyer and Mullen, 2008).

TAGs são normalmente armazenados na forma de gotas de lipídios localizadas no citoplasma ou em estruturas especializadas denominadas oleossomos ou corpos lipídicos (Figura 6) (Voelker and Kinney, 2001).

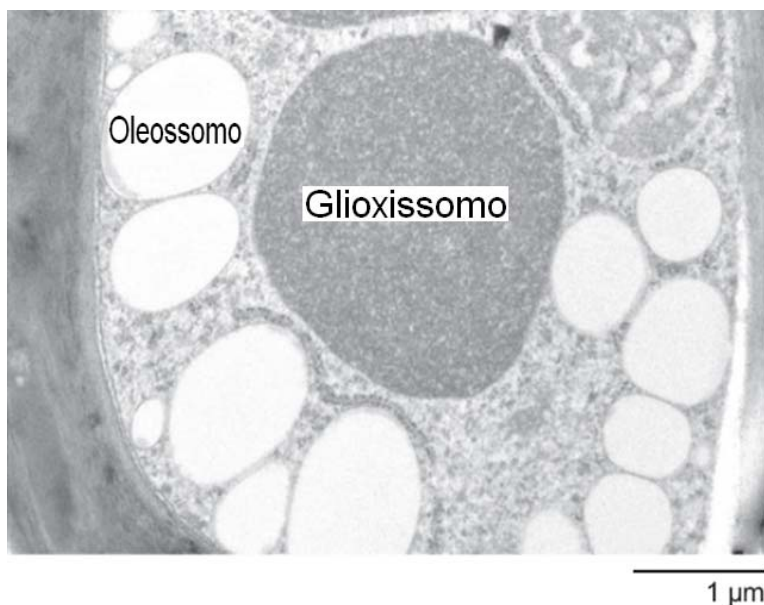


Figura 6. Oleossomo. Micrografia de transmissão de elétrons de semente de *A. thaliana* mostrando oleossomo. Fonte: (Graham, 2008).

Oleossomos consistem em *pools* de TAGs circundados por uma membrana de camada única. O modelo mais aceito (Figura 7) afirma que tais estruturas são originadas do desprendimento da membrana externa do retículo endoplasmático. Segundo esse modelo, proteínas envolvidas no metabolismo de lipídios neutros acumulam-se em certas regiões do retículo endoplasmático. Entre esses polipeptídios, encontram-se enzimas que catalisam a formação de TAGs e esterol ésteres (SE) (Figura 7a). Devido ao fato de que tais lipídios neutros recém formados são inaptos para se integrar à bicamada lipídica da membrana eles formam um *cluster* e agrupam-se na região hidrofóbica entre as duas camadas que compõem a membrana do retículo endoplasmático (Figura 7b). Durante a síntese de TAGs e SEs as gotas crescem e formam uma estrutura em forma de botão (Figura 7c). Após atingir certo tamanho, estes

“botões de gordura” desprendem-se do retículo endoplasmático e tornam-se estruturas independentes no citosol (Figura 7d) (Athenstaedt and Daum, 2006).

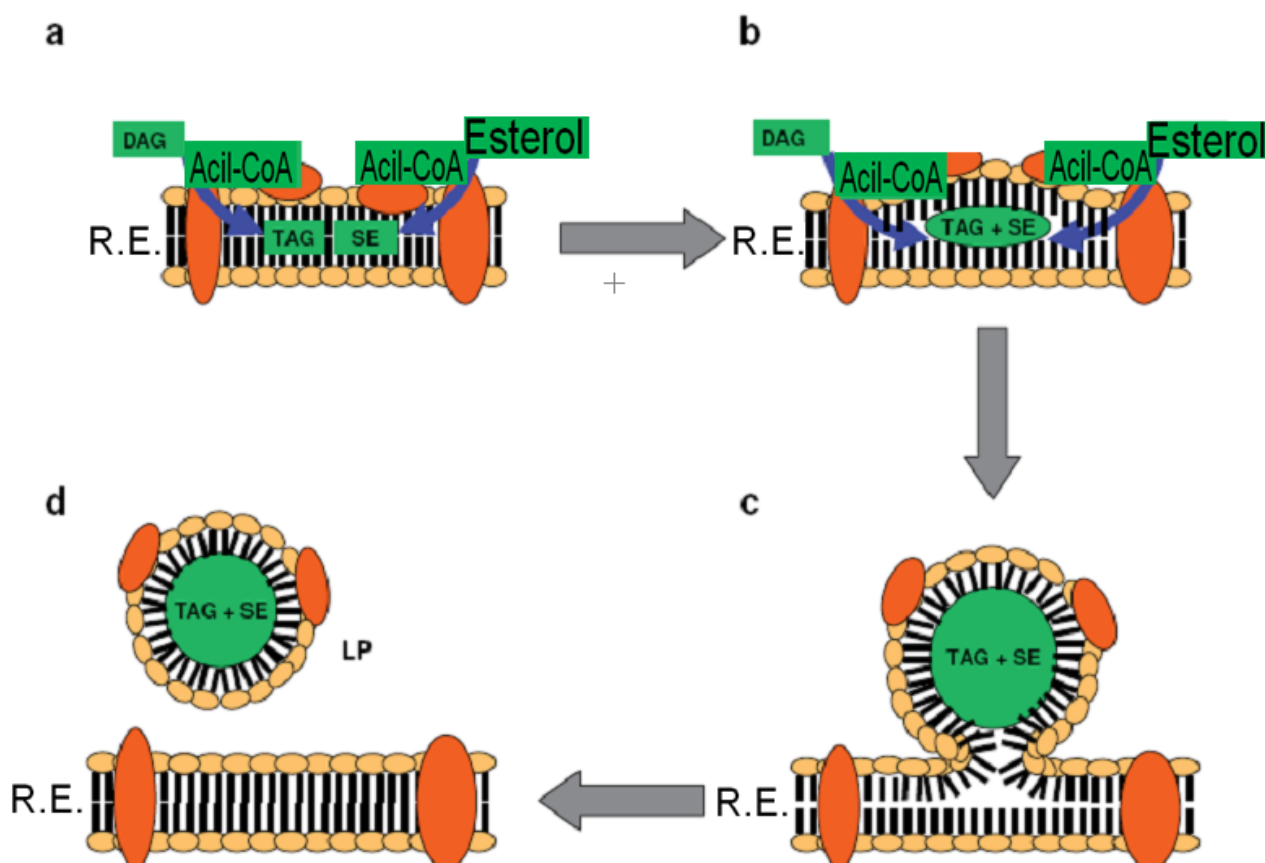


Figura 7. Modelo para a formação de oleossomos. (a) A síntese de triacilglicerídeos (TAG) e esterol ésteres (SE) por enzimas do retículo endoplasmático leva ao (b) acúmulo de TAGs e SEs entre as duas camadas da bicamada fosfolipídica do retículo (R.E.). A contínua síntese de lipídios neutros resulta no acúmulo de TAGs e SEs, formando gotas (c), que se desprendem do retículo após atingir certo tamanho (d). DAG, diacilglicerol; acil-CoA, ácido graxo ligado à coenzima A. Fonte: (Athenstaedt and Daum, 2006).

A membrana externa que envolve os corpos lipídicos contém proteínas conhecidas como oleosinas, cuja função é estabilizar o oleossomo durante a dissecação da semente (Figura 8) (Huang, 1996; Voelker and Kinney, 2001).

Ao contrário da composição dos ácidos graxos das membranas celulares, a qual é restrita a uma pequena classe de ácidos graxos conservados, a composição dos ácidos graxos dos TAGs pode ser altamente divergente, com mais de 300 estruturas diferentes. Esses ácidos graxos podem variar no comprimento, número e posição de ligações duplas e triplas, e podem conter grupos funcionais adicionais, como grupos hidroxil e epóxi, por exemplo (Dyer and Mullen, 2008).

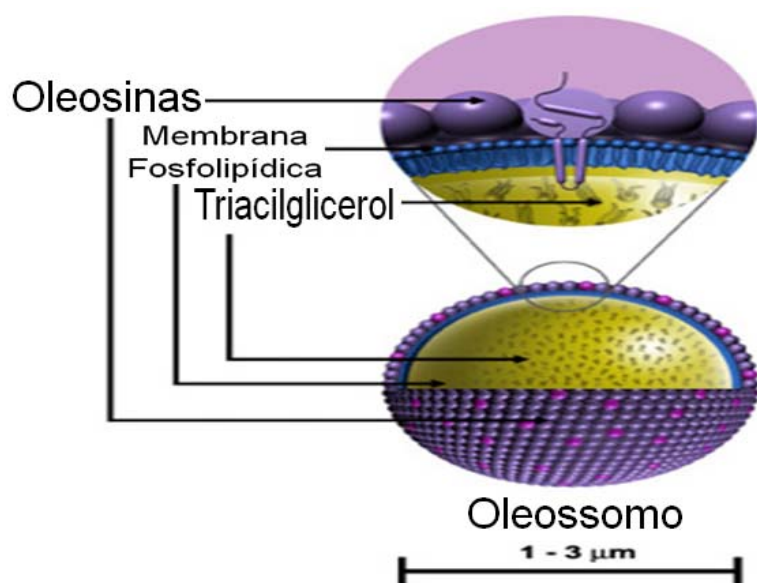


Figura 8. Estrutura de uma oleossomo. O oleossomo é recoberto por uma membrana fosfolipídica revestida, por sua vez, por uma camada de proteínas oleosinas.

Duas rotas metabólicas distintas para produção de TAGs têm sido elucidadas, uma rota dependente de acil-CoA e outra independente de acil-CoA (Figura 9).

A rota dependente de acil-CoA, comumente denominada “rota de Kennedy”, utiliza acil-CoA como substrato para sucessivas reações de acilação da molécula de glicerol (Kennedy, 1961).

A primeira enzima da rota de Kennedy, glicerol-3-fosfato aciltransferase (G3PAT), catalisa a acilação de glicerol-3-fosfato em sua posição *sn*-1, formando ácido lisofosfatídico (LPA). A segunda enzima aciltransferase da cascata enzimática, ácido

lisofosfatídico aciltransferase (LPAAT), é a responsável pela acilação na posição *sn*-2 do LPA originando ácido fosfatídico (PA). PA é então defosforilado por ação da enzima ácido fosfatídico fosfatase (PAP) produzindo diacilglicerol (DAG). A terceira aciltransferase, diacilglicerol aciltransferase (DGAT), catalisa a acilação de DAG levando à produção de triacilglicerol (TAG). É importante salientar que a especificidade destas aciltransferases desempenha um papel decisivo na determinação da composição final do conteúdo de TAGs em órgãos de reserva (Kocsis and Weselake, 1996).

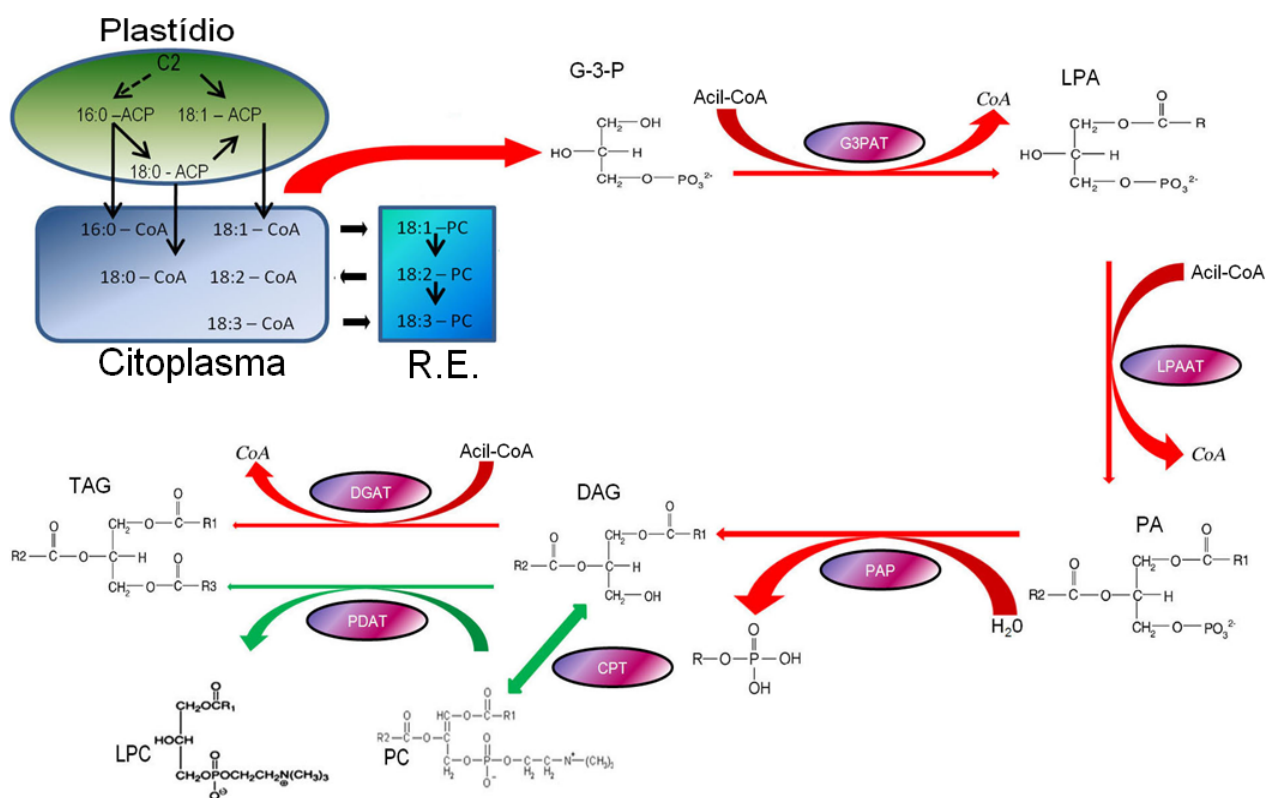


Figura 9. Rotas de biossíntese de triacilgliceróis (TAGs)

Esquema geral da biossíntese de TAGs em sementes de oleaginosas. A síntese de ácidos graxos ocorre nos plastídios. Ácidos graxos sintetizados são transportados ao citoplasma onde são esterificados junto à Coenzima A (CoA) formando um complexo denominado acil-CoA. Em seguida, ocorre a subsequente incorporação dos ácidos graxos na molécula de glicerol levando a formação de TAG. Setas vermelhas indicam a rota dependente de Acil-CoA (rota de Kennedy). Setas verdes indicam rotas alternativas para a produção de TAGs (independentes de Acil-CoA). Enzimas da biossíntese de TAG são representadas pelas esferas ovaladas: glicerol-3-fosfato aciltransferase (G3PAT), ácido lisofosfatídico aciltransferase (LPAAT), ácido fosfatídico fosfatase (PAP), diacilglicerol aciltransferase (DGAT), fosfolípido: diacilglicerol

aciltransferase (PDAT) e colina fosfotransferase (CPT). ACP, proteína carreadora de acil; G-3P, glicerol-3-fosfato; LPA, ácido lisofosfatídico; PA, ácido fosfatídico, DAG, diacilglicerol; TAG, triacilglicerol; PC, fosfatidilcolina; LPC, lisofosfatidilcolina; RE, retículo endoplasmático.

A rota independente de acil-CoA (figura 9) usa uma enzima alternativa pra a última reação de acilação, denominada fosfolípido: diacilglicerol aciltransferase (PDAT). PDAT transfere um grupo acila diretamente da fosfatidilcolina (PC) para DAG formando TAG (Dahlgvist *et al.*, 2000).

Em outra rota de transacilação independente de acil-CoA (Stobart and Stymne, 1985), é possível que uma enzima denominada diacilglicerol transacilase (DGTA) utilize duas moléculas de DAG para a formação de TAG. Entretanto, até o presente momento nenhum gene codificando uma DGTA foi identificado em nenhum organismo (Kroon *et al.*, 2006).

Adicionalmente, a enzima colina fosfotransferase (CPT) pode contribuir para a biossíntese de TAG catalisando a transferência reversível de PC para DAG (Somerville and Browse, 1991). Em seguida, o DAG resultante pode ser convertido em TAG pelas enzimas DGAT ou PDAT (figura 9).

1.5 Aplicações industriais dos óleos vegetais

Os óleos vegetais vêm sendo utilizados para gerar calor e luz desde tempos remotos. Os óleos vegetais possuem um conteúdo energético maior do que outras fontes bioenergéticas, tais como etanol. Além disso, apresentam cerca de 90% da capacidade de produção de calor do diesel derivado de petróleo e exibem uma proporção *input/output* de energia que varia de 1:2 a 1:4 (quantidade de energia

necessária para a produção da matéria-prima comparada à quantidade de energia obtida pelo uso do óleo vegetal) (Agarwal, 2007).

Os óleos vegetais utilizados nas indústrias alimentícia e química são, na grande maioria, compostos por uma variada proporção dos cinco mais comuns e nutricionalmente importantes ácidos graxos vegetais: ácidos palmítico, esteárico, oléico, linoléico e linolênico (Figura 10) (Durrett *et al.*, 2008).

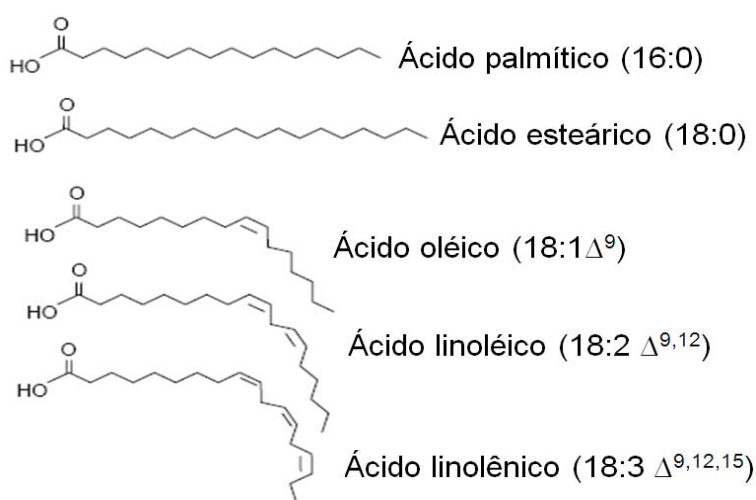


Figura 10: Principais ácidos graxos encontrados nos óleos vegetais. Ácidos palmítico (16:0), esteárico (18:0), oléico (18:1), linoléico (18:2) e linolênico (18:3). Fonte: (Jaworski and Cahoon, 2003).

Atualmente, os óleos vegetais são produzidos, principalmente, a partir de quatro culturas: palma, soja, brassicáceas e girassol, que juntas respondem por aproximadamente 80% do total de óleo vegetal produzido no mundo (Figura 11) (Dyer and Mullen, 2008).

As propriedades dos óleos vegetais dependem, em grande parte, da composição de seus ácidos graxos, sendo certas composições mais desejáveis para um uso específico do que outras. Por exemplo, óleos de cozinha normalmente contêm uma grande proporção de ácidos graxos mono-insaturados (como ácido oléico, por exemplo), que são mais estáveis sob alta temperatura enquanto margarinas e manteigas são geralmente enriquecidas em ácidos graxos saturados (ácidos palmítico e esteárico, por

exemplo). Outros óleos, como os usados no tempero de saladas, contêm uma quantidade maior de ácidos graxos poli-insaturados (ácidos linoléico e linolênico). Tradicionalmente, a produção de óleos para um uso específico tem sido atingida pela mistura de óleos de várias plantas ou por sua parcial hidrogenação, na qual as duplas ligações dos ácidos graxos são removidas para tornar o óleo mais saturado. Entretanto, o processo de hidrogenação também introduz ácidos graxos trans-esterificados, que apresentam efeitos indesejáveis à saúde humana, contribuindo, principalmente para aumento dos riscos de problemas cardíacos (Kinney, 1996; Ascherio, 2006).

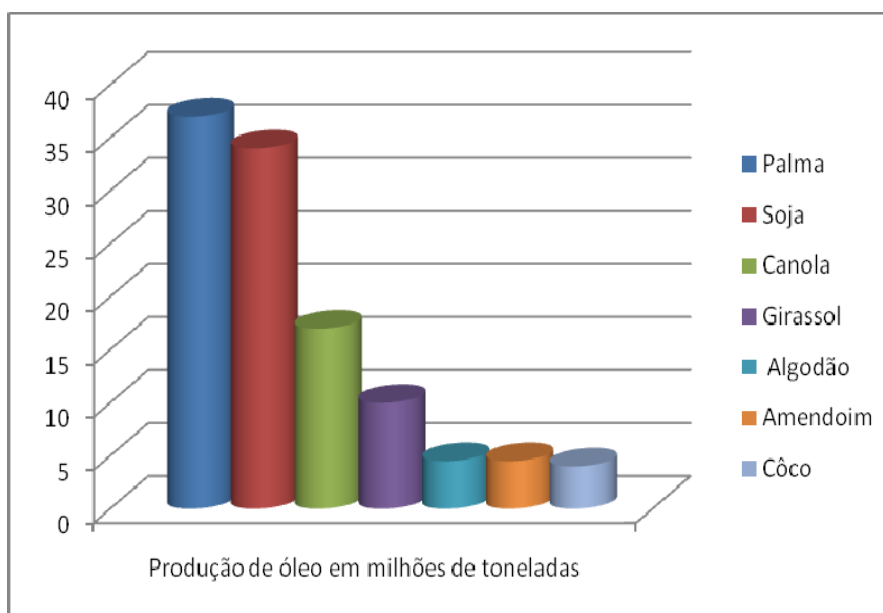


Figura 11. Produção mundial de oleaginosas. Produção mundial de óleo em milhões de toneladas (eixo y) para cada uma das principais oleaginosas cultivadas. Figura adaptada de (Durrett *et al.*, 2008).

As duplas ligações presentes nos ácidos graxos também representam um excelente ponto de início para modificações na cadeia de hidrocarbonetos para a produção de novos tipos de matérias-primas. Por exemplo, os óleos vegetais podem ser tratados com uma variedade de químicos para converter as duplas ligações de seus ácidos graxos em grupamentos hidroxila e o produto resultante pode ser misturado a

compostos tais como isocianeto, originando poliuretanos. Estes ambientalmente corretos e renováveis poliuretanos, alternativos aos poliuretanos derivados de petróleo, apresentam excelentes características físicas e alto desempenho em uma grande variedade de funções, tais como na fabricação de espumas em *spray*, espumas rígidas, espumas flexíveis usadas na forração do interior de carros, selantes, adesivos, dentre outros. Estes e outros tipos de modificações químicas em óleos vegetais podem criar novos e originais compostos com propriedades interessantes, principalmente do ponto de vista comercial. Entretanto, o tratamento de óleos vegetais com químicos contribui significativamente para o aumento dos custos de produção e para os danos ao meio ambiente resultante da utilização destes óleos para propósitos industriais (Metzger and Bornscheuer, 2006).

Devido à grande quantidade de ácido ricinoléico, o óleo de mamona, extraído através da prensagem das sementes, apresenta amplo potencial para utilização industrial, tornando a cultura da mamoneira importante economicamente para o país. O óleo apresenta em sua cadeia carbônica sítios em que são realizadas diversas reações químicas, obtendo-se assim uma variedade de subprodutos, derivados das modificações das cadeias carbônicas, conforme mostra a tabela 2 (Henriques, 2009).

Tabela 2. Aplicações industriais do óleo de mamona (IAC, 2008).

Sítio de Reação Química	Derivado	Aplicação
Ligação Éster	Metilricinoleato	Nylon-11 (Fios, Tubos, Indústria Automobilística, Aeronáutica.)
Dupla Ligação	Óleo Hidrogenado	Ceras, lubrificantes, Cosméticos, Plásticos
	Óleo Oxidado	Plasticizante, Protetores, Tintas, Adesivos
Grupo Hidroxila	Óleo Desidratado	Sicativo
	Óleo Sulfonado	Indústria Têxtil
	Ácido Sebácico	Lubrificantes, Nylon 6-10
	Óleo Etoilado	Cosméticos, Detergentes, Lubrificantes de Superfície, Óleo de Corte, Fluido Hidráulico, Ind. Têxtil
	Poliuretanos	Telecomunicações, Materiais Elétricos, Produtos Biomédicos, Filtros Industriais
	Transesterificação	Biodiesel

O óleo da mamona é utilizado em vários processos industriais, tais como, fabricação de desinfetantes, óleos lubrificantes, tintas e corantes; manufatura de fungicidas, e como base para inseticidas (Vital *et al.*, 2005). No Brasil, o mercado de mamona é oligopsônico, o que não permite muita elasticidade de preço pago ao produtor. Embora já existam em grande quantidade, outros produtores continuam surgindo. Além dos pequenos produtores isolados, associações e cooperativas, os assentamentos feitos pelo governo podem tornar-se centros de produção de mamona, com o objetivo de ser matéria-prima para a produção do biodiesel. Acredita-se no forte crescimento deste mercado nos próximos anos, com a redução da área plantada com esta euforbiácea pela China, e a grande possibilidade de se tornar uma das matérias-primas para a produção de biodiesel, com a vantagem de ter elevado rendimento (mais de 99,0% na conversão do óleo em biodiesel). A reação ocorre a frio devido à solubilidade do óleo da mamona em álcool (metílico ou etílico) (EMBRAPA, 2004).

2 Justificativa

As plantas apresentam-se como uma excelente alternativa para a substituição do uso do petróleo porque elas constituem-se numa fonte renovável (parte das sementes colhidas pode ser usada para restituir todo o processo). Além disso, sementes são biofábricas ideais para a produção de óleo, proteína e amidos, muitos dos quais são usados para propósitos energéticos ou industriais (Scheller and Conrad, 2005; Metzger and Bornscheuer, 2006).

Nos últimos anos, esforço considerável foi realizado no melhoramento de culturas agronomicamente bem estabelecidas, visando à obtenção de óleos vegetais com propriedades atrativas à indústria. Um foco particular desse esforço tem sido os estudos de identificação e da expressão heteróloga de genes que codificam enzimas envolvidas na síntese de ácidos graxos ditos “não usuais” que, posteriormente, serão acumulados na forma de TAGs (Jaworski and Cahoon, 2003).

A produção de óleos industrialmente importantes em plantas cultivadas é um dos maiores objetivos da biotecnologia de plantas e representa um importante passo na transição de uma sociedade baseada na utilização de óleos fósseis para uma economia mais sustentável, baseada na utilização de fontes mais limpas e renováveis de energia, denominada bioeconomia (Dyer and Mullen, 2008).

Bioeconomia pode ser definida como toda atividade econômica derivada do (1) desenvolvimento de um melhor entendimento dos mecanismos moleculares de sistemas biológicos e (2) a aplicação deste conhecimento para expandir e diversificar produtos obtidos a partir de fontes agrícolas, químicas ou industriais (Dyer and Mullen, 2008).

Apesar de muitos dos genes envolvidos no metabolismo de lipídios terem sido identificados e poderosas ferramentas para quantificação de produtos lipídicos estarem disponíveis, existe um vazio significativo no que se refere à compreensão dos processos biológicos envolvidos na expressão desses genes e na produção de seus respectivos lipídios (Dyer and Mullen, 2008).

Além disso, um dos mais importantes problemas associados às sementes de mamona é a presença da toxina ricina e da substância alergênica albumina 2S que apresentam alto potencial de danos à saúde tanto na etapa de produção quanto de beneficiamento (Chen *et al.*, 2007). Portanto, é extremamente desejável o desenvolvimento de uma fonte segura para a produção de ácido ricinoléico.

Esforços para a transferência de genes codificantes de proteínas responsáveis pela biossíntese de ácidos graxos não-usuais em espécies agronomicamente mais produtivas tem, geralmente, apresentando sucesso limitado (Shockey *et al.*, 2006).

Por exemplo, o cDNA codificando a oleato 12-hidroxilase (FAH12) de mamona foi clonado e a seqüência deduzida de aminoácidos da proteína mostra homologia com a oleato 12-desaturase (FAD2) do retículo endoplasmático (ER) de plantas superiores (van de Loo *et al.*, 1995). A expressão deste cDNA em sementes de *Arabidopsis thaliana* resultou na síntese de 4 novos hidroxiácidos de ácidos graxos (Smith *et al.*, 2003) e num acúmulo de cerca de 17% de ácido ricinoléico no conteúdo total de ácidos graxos. Esse nível é muito inferior ao encontrado na mamona (cerca de 90%) (Broun and Somerville, 1997). Estes resultados indicam que genes adicionais e um maior conhecimento sobre a biossíntese de óleos em sementes são necessários antes de plantas serem submetidas

à engenharia genética visando a produção de óleos industrialmente importantes (Shockey *et al.*, 2006).

Se a produção de ácidos graxos industrialmente importantes em culturas altamente produtivas for desenvolvida através de estratégias de engenharia genética, haverá um enorme volume de mercado disponível. Assumindo que estes óleos vegetais possam competir eficientemente com derivados de petróleo, tanto no que se refere às propriedades funcionais quanto aos preços, um maior potencial de crescimento para matérias-primas industriais derivadas de óleos vegetais poderá ser criado (Dyer and Mullen, 2008).

Tanto ácidos graxos comuns quanto não usuais presentes em óleos vegetais são capazes de satisfazer a demanda de uma grande variedade de mercados existentes e hoje servidos pela indústria petroquímica. Entretanto, o sucesso no desenvolvimento de produtos baseados em oleoquímicos e ácidos graxos nutricionalmente importantes para o mercado mundial é extremamente dependente de estratégias eficazes para produção e processamento, a baixos custos, dessa matéria-prima, bem como a produção destes ácidos graxos em culturas altamente produtivas. Nesse contexto, ferramentas biotecnológicas podem contribuir muito para que tais objetivos sejam alcançados (Dyer and Mullen, 2008).

A elucidação da atividade das enzimas responsáveis pelo controle do fluxo de ácidos graxos levando a formação de TAGs poderá contribuir não apenas para um melhor entendimento do processo como também para a possível transferência de ácidos graxos de interesse para óleos de reserva.

Portanto, é de especial interesse a identificação e a caracterização dos genes envolvidos no metabolismo dos lipídios em mamona, especialmente aqueles envolvidos no acúmulo de ácido ricinoléico na forma de triacilglicerídios.

A caracterização do perfil de transcritos dos genes do metabolismo lipídico em mamona poderá revelar como e quando cada um desses genes é expresso, contribuindo para um melhor entendimento do processo que culmina no acúmulo de ácido ricinoléico em níveis extremamente elevados e discrepantes aos observados em culturas tradicionais, além de auxiliar na escolha de potenciais genes-alvos para abordagens biotecnológicas.

3 OBJETIVOS

3.1 Objetivo geral:

Identificar genes relacionados à síntese de triacilglicerídios em mamona (*Ricinus communis*) e caracterizar o padrão de expressão dos genes identificados ao longo do desenvolvimento da semente.

3.2 Objetivos específicos:

- a) Identificar os genes envolvidos na biossíntese de triacilglicerídios em sementes de mamona;
- b) Caracterizar *in silico* os genes identificados;
- c) Desenhar oligonucleotídeos específicos para cada um dos genes estudados e validar a amplificação dos mesmos;
- d) Analisar o padrão de expressão dos genes identificados em diferentes etapas do desenvolvimento/amadurecimento da semente de mamona

4 RESULTADOS E DISCUSSÃO

Este item será apresentado na forma de artigo científico submetido à revista *Phytochemistry*, em setembro de 2009.

**Identification and expression analysis of castor bean (*Ricinus communis*) genes
encoding enzymes from the triacylglycerol biosynthesis pathway**

Title:

Identification and expression analysis of castor bean (*Ricinus communis*) genes encoding enzymes from the triacylglycerol biosynthesis pathway

Authors:

Alexandro Cagliari ¹, Guilherme Loss², Matheus Etges², Márcia Margis-Pinheiro ¹ and Rogério Margis^{2§}

Affiliations:

¹ Departamento de Genética, Universidade Federal do Rio Grande do Sul, Brasil

² Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Brasil

§Corresponding Author

Email addresses:

AC: alexandrocagliari@yahoo.com.br;

GL: guilherme.loss@gmail.com

ME: matheus.etges@hotmail.com;

MMP: marcia.margis@ufrgs.br

RM: rogerio.margis@ufrgs.br

Keywords: castor bean; lipid metabolism; lipid biosynthesis; seed development; ricinoleate; triacylglyceride; TAG

Abstract

Castor bean (*Ricinus communis*) seed oil contains high amounts of ricinoleic acid rich triacylglycerols (TAGs). As result of its physical and chemical properties, castor oil and its derivatives are used for numerous bio-based products. An oleate hydroxylase (FAH12) catalyzes the hydroxylation that converts oleate to ricinoleic acid. Heterologous expression of FAH12 in the model oilseed plant *Arabidopsis thaliana* produced only up to 17% of hydroxy fatty acids in seed oils. This level is much lower than that found in castor oil (90%), suggesting that additional genes and significantly more knowledge of seed oil biosynthesis are needed before plants can be engineered to produce industrially important oils. In this study, we survey the Castor Bean Genome Database to report the identification of TAG biosynthesis genes. A set of twenty-six genes encoding six distinct classes of enzymes involved in TAGs biosynthesis were identified. Interestingly, *in silico* characterization and phylogenetic analysis allowed us to identify plastidic isoforms of glycerol-3-phosphate acyltransferase and lysophosphatidate acyltransferase enzyme families, involved in prokaryotic lipid biosynthesis pathway, that form a cluster apart from the cytoplasmic isoforms, involved in eukaryotic pathway. Besides, two distinct membrane bound polypeptides of diacylglycerol acyltransferase enzyme were identified. Additionally genes encompasses phosphatidate phosphatase, phospholipid:diacylglycerol acyltransferase and choline-phosphotrasferase were also identified. Quantitative expression pattern analyses demonstrated variations on gene expressions along the castor seed development as much among enzyme classes as into the genes of same enzyme class. In general, however, there is a tendency of maximum expression level at the middle of seed development, although this is not observed at all the genes analyzed. Our results represent snapshots of global transcriptional activities of genes encompassing six enzyme families involved in castor bean TAG biosynthesis along seed development. This results can contributed for the choice of genes that will be target to biotechnological approaches for the production of nutritionally and industrially desirable oils.

1 Introduction

Triacylglycerols (TAGs) are the major storage lipids in living organisms. As a highly reduced form of carbon, they serve as an important energy reserve in plant seeds, providing for subsequent germination and seedling development. These storage lipids are of great nutritional and nutraceutical value and, thus, are a common source of edible oils for human consumption and industrial purposes (Lung and Weselake, 2006).

According to the current knowledge of castor seed TAG biosynthesis, oleic acid is synthesized in the plastid and then exported to the cytoplasm following the standard fatty acid biosynthesis pathway. Oleic acid is activated to oleoyl-CoA in the cytoplasm and imported into the endoplasmic reticulum (ER) with the help of coenzyme A (CoA). In the ER, it undergoes a hydroxylation process, through the activity of the oleate hydroxylase (FAH12) enzyme, yielding ricinoleic acid (C18:OH) that will be incorporated into a glycerol backbone by successive acylation reactions, yielding TAG (Fig.1B) (Somerville C *et al.*, 2000).

Two metabolic pathways for the production of TAGs have been elucidated: an acylCoA-dependent pathway and an acyl-CoA-independent pathway (Kroon *et al.*, 2006). The acyl-CoA-dependent pathway, commonly known as the Kennedy pathway, utilizes acyl-CoA as a substrate for successive acylation reactions of the glycerol backbone (Kennedy, 1961). Glycerol-3-phosphate acyltransferase (G3PAT), the first enzyme of the Kennedy pathway, catalyzes the transfer of a fatty acid to glycerol-3-phosphate (G3P) at the *sn*-1 position to form lysophosphatidic acid (LPA). The second acyltransferase, lysophosphatidic acid acyltransferase (LPAAT), is responsible for *sn*-2 acylation of LPA to yield phosphatidic acid (PA). PA is then dephosphorylated by phosphatidic acid phosphatase (PAP) to yield diacylglycerol (DAG). The third acyltransferase,

diacylglycerol acyltransferase (DGAT), catalyzes the acylation of DAG that leads to the production of triacylglycerol (Fig.1). The specificities of these acyltransferases for fatty acids with different lengths and degrees of unsaturation play an important role in determining the final TAG composition (Kocsis and Weselake, 1996).

The acyl-CoA-independent pathway uses an alternative enzyme for the final acylation reaction, termed phospholipid:diacylglycerol acyltransferase (PDAT). PDAT directly transfers an acyl group from phosphatidylcholine (PC) to DAG, producing TAG (Fig.1) (Dahlqvist *et al.*, 2000). In another acyl-CoA-independent transacylation reaction described, a DAG transacylase (DAGTA) uses two molecules of DAG to form TAG. However, at present, no gene encoding a DAGTA has been identified (Stobart and Stymne, 1985; Kroon *et al.*, 2006). Choline phosphotransferase (CPT) can also contribute to TAG biosynthesis by catalyzing the reversible conversion of PC to DAG (Somerville and Browse, 1991). Afterward, the resulting DAG can be converted to TAG by the DGAT or PDAT enzyme (Fig.1B). Some plant species are known to accumulate high levels of unusual fatty acids during TAG synthesis. These unusual fatty acid structures include medium chain lengths (C8-C12), very long chain lengths (C20-C24), or substitutions such as hydroxyl, epoxy, and cyclopropane groups (Stobart and Stymne, 1985; Gunstone, 1987; Bafor *et al.*, 1991; Somerville and Browse, 1991). Castor seed triacylglycerols contain about 90% ricinoleic acid (12-hydroxy-oleate). This unusual fatty acid possesses a hydroxyl group (-OH) responsible for the unique chemical and physical properties that make castor oil a vital industrial raw material for numerous products such as cosmetics, paints, coatings, plastics and anti-fungal compounds (Caupin, 1997).

However, one of the most important problems associated with castor seeds is the presence of the toxin ricin and hyper-allergenic 2S albumins, which are detrimental to

growers and processors (Chen *et al.*, 2007). Therefore, it is highly desirable to develop a safe source for castor oil production. Furthermore, efforts to transfer genes encoding the proteins responsible for unusual fatty acid biosynthesis to higher yielding plants have generally produced limited success. A castor plant gene for a fatty acid hydroxylase (*FAH12*), which is directly responsible for synthesis of ricinoleic acid, has been successfully isolated. Its over expression in the heterologous model oilseed *Arabidopsis thaliana* produced only 17% ricinoleic acid content. This accumulated oil level is much lower than that observed in native plant species (90%) (Broun and Somerville, 1997), and might be resulted of an inefficient incorporation of this novel fatty acid along the TAG biosynthesis in heterolog model. These results suggest that additional genes are involved and that significantly better knowledge of seed oil biosynthesis, especially TAG biosynthesis, is needed before plants can be engineered to produce industrially important oils (Shockey *et al.*, 2006).

In order to abroad the knowledge about the TAG biosynthesis in castor bean, in this study, we inspected the available Castor Bean Genome Database (<http://castorbean.jcvi.org>) and reported the identification of twenty-six genes representing six classes of enzymes that participate in different steps of TAG biosynthesis. Furthermore, we characterized the expression profiles of these genes during seed development and the accumulation of ricinoleic acid and TAG. Functional roles of the different genes encoding members of the diverse classes of enzymes are discussed. Our results may contribute for future studies involving design of strategies for manipulation of oil production in plants.

2 Results and Discussion

A series of Basic Local Alignment Search Tool (BLAST) searches, using known *A. thaliana* lipid enzymes as queries, identified twenty-six genes encompassing six distinct metabolic lipid families involved in TAG biosynthesis. Amino acid sequence alignment revealed homologous regions among identified genes, allowing us to identify motifs and conserved characteristic signatures of each class of enzyme (Supplementary Data 1 to 6). In addition, phylogenetic analysis was conducted in order to confirm the identity of the different genes and to determine the clustering pattern representing the genetic variability of each enzyme class (Fig.2).

The use of quantitative real-time polymerase chain reaction (qPCR) allowed for determination of the steady-state levels of mRNA transcripts of these genes in samples from each of the five seed developmental stages. For convenience, these mRNA levels will be referred to as expression levels throughout this article. Expression profiles of TAG biosynthesis genes are shown in Fig.3.

2.1 Glycerol-3-phosphate-1-acyltransferase (G3PAT)

G3PAT is a soluble enzyme that catalyzes the incorporation of an acyl group from either acyl-acyl carrier proteins (acyl-ACPs) or acyl-CoA molecules into the *sn*-1 position of glycerol-3-phosphate (G3P) to yield lysophosphatidic acid (LPA). In several plants, two G3PAT isoforms have been characterized, found in the plastidial and cytoplasmic cellular compartments. In *Arabidopsis*, plastidic G3PAT is encoded by the gene *At1g32200*. The corresponding *Arabidopsis* cDNA has been expressed in *Escherichia coli*, confirming the activity of the encoded protein as a G3PAT. Using the amino acid sequence of *At1g32200* as a BLAST query allowed for the identification of a single,

putative castor bean plastidic G3PAT gene, named *G3PAT1* (Fig.2 and Supplementary Data 1A and 7). *G3PAT1* contains two conserved acyltransferase motifs (ATMs I and II), harboring the amino acid residues HxxxxD (Lewin *et al.*, 1999), and one trans-membrane domain (TmD, see Supplementary Data 1). The role of conserved histidine and aspartic residues in conserved ATMs is discussed above.

It was reported the existence of an *A. thaliana* cytoplasmic G3PAT gene family that encodes membrane-bound G3PATs (Zheng *et al.*, 2003). The *AtG3PAT* family consists of seven members that share high sequence identity to each other but no significant homology with plastidic *At1g32200*. Using seven cytoplasmic *A. thaliana* G3PAT sequences (Supplementary Data 1B), we performed BLAST analyses that resulted in the identification of eight putative cytoplasmic castor bean G3PAT genes (*G3PAT2* to 9) (Fig.2 and Supplementary Data 1B and 7). Amino acid sequence alignment revealed that all cytoplasmic G3PATs contain the previously defined four acyltransferase motifs (ATMs I to IV) (Lewin *et al.*, 1999) and two putative transmembrane domains (Supplementary Data 1B). The histidine and aspartic residues in ATM I and the proline residue in ATM IV, all of which have been shown to be catalytically important (Lewin *et al.*, 1999), are completely conserved (black stars, Supplementary Data 1B) (Zheng *et al.*, 2003). It has been proposed that the invariant histidine in ATM I acts as a general base to remove a proton from the hydroxyl group at the *sn*-1 position of glycerol 3phosphate, facilitating nucleophilic attack on the thioester of palmitoyl-CoA. The invariant aspartate in ATM I may act in a charge relay system with the histidine residue to increase the nucleophilicity of the glycerol-3-phosphate hydroxyl group (Lewin *et al.*, 1999). Furthermore, residues critical for G3P binding, including the arginine in ATM II and the glutamine and serine in ATM III, also are invariant among the

members of the cytoplasmic *Ricinus G3PAT* family (black arrows in Supplementary Data 1B) (Zheng *et al.*, 2003). The majority of cytoplasmic G3PAT paralogs identified in castor bean have orthologous genes in *Arabidopsis*, as observed in the five cluster groups in the main G3PAT clade (Fig.2). Interestingly, phylogenetic analysis showed that all cytoplasmic G3PAT genes formed a cluster apart from plastidic G3PAT genes (Fig.2). This cluster separation can be explained by differences in amino acid sequence and by functional divergence in this class of enzyme.

Analyses of gene expression throughout seed maturation were carried out after relative expression determination in five stages of seed development (Fig.1A). The results of qPCR analysis demonstrated that the castor bean plastidic *G3PAT1* mRNA reached its highest level at stages 3 and 4 (Fig.3A), with a decrease at stage 5 to a level similar to that observed at the initial stages (1 and 2). This pattern was the same as that presented by oleate hydroxylase (data not shown) and coincided with the stages of highest oil accumulation in castor seeds (Chen *et al.*, 2007). In general, plastidic *G3PAT1* demonstrated the largest relative expression level of the G3PAT family during castor seed development (Fig.3A). This result indicates that plastidic *G3PAT1* has a key role in the first acylation reaction of castor bean TAG biosynthesis, at least at the level of mRNA accumulation.

During seed development, the cytoplasmic G3PATs presented different expression patterns (Fig.3A). *G3PAT2* had a biphasic pattern, with high cytoplasmic G3PAT expression in the majority of seed developmental stages (except for stage 2). *G3PAT2* showed high mRNA accumulation during stage 1, followed by a decrease in stage 2 and a subsequent increase in stages 3 and 4, followed again by a decline at stage 5. *G3PAT3* mRNA accumulated at high levels in the earlier stages and then

decreased to lower levels over time. *G3PAT9* showed a drastic increase, with peak expression at stage 2 followed by a gradual reduction at later stages.

The differential expression patterns shown by the G3PAT family may indicate functional and temporal complementation activities. G3PAT genes presented maximal expression levels at different castor seed developmental stages, indicating that the expression of some genes is more important in certain stage than others. It seems that the temporal succession of G3PAT gene expression results in the maintenance of enzymatic activity throughout castor seed development.

The expression of cytoplasmic *G3PATs* 4 to 8 was not detected at any stage of seed development. These genes may demonstrate tissue-specific or temporally controlled expression patterns or be induced in response to specific conditions, such as environmental changes or stress, that differ from our experimental conditions. Besides, many of the undetectable genes might be involved in the membrane homeostasis and might not be specific to the seed and biosynthesis of TAG, making difficult their detection in our experimental approach.

2.2 *Lysophosphatidic acid acyltransferase (LPAAT)*

LPAAT is the enzyme responsible for the acylation of 1-acyl-*sn*-G3P (LPA), leading to the formation of phosphatidic acid (PA). In plants, LPAAT activity is associated with multiple membrane systems, including chloroplasts, endoplasmic reticulum (ER) and the outer membrane of mitochondria, suggesting the presence of several different isoforms (Bourgis *et al.*, 1999; Yu *et al.*, 2004). In traditional oil seed crops, LPAAT possesses strong specificity and selectivity for unsaturated C18 acyl groups but is able to discriminate acyl groups having longer or shorter chain lengths. LPAATs that can

incorporate unusual acyl groups are found in the developing seeds of certain plant species, such as castor bean, which accumulates oils esterified with unusual fatty acids (Weier *et al.*, 1998).

ATS2 locus was identified as the *At4g30580* gene of *Arabidopsis*, encoding the plastidic isoform of LPAAT. Introduction of *ATS2* cDNA into *E. coli* JC 201, which is temperature-sensitive and carries a mutation in its LPAAT gene *plsC*, restored this mutant to nearly wild type growth at high temperature (Yu *et al.*, 2004). Two putative LPAAT genes were characterized in *A. thaliana* that encode non-plastidic LPAATs (*AtLPAAT2* and *AtLPAAT3*) (Kim *et al.*, 2005). *AtLPAAT2* (At3g57650) encodes a ubiquitous ER-localized LPAAT, which functionally complement a bacterial mutant with defective LPAAT. *AtLPAAT2* was ubiquitously expressed in diverse tissues, while *AtLPAAT3* (At1g51260) was preferentially expressed in pollen (Kim *et al.*, 2005).

We identified one castor bean plastidic LPAAT isoform (*LPAAT1*) using *ATS2* as a BLAST query (Fig.2 and Supplementary Data 2A and 7). *LPAAT1* possesses a conserved acyltransferase motif (ATM, HxxxD) and also contains three predicted transmembrane domains (TmD in Supplementary Data 2A). qPCR analysis demonstrated that *LPAAT1* has a bell-shaped expression pattern, with a linear rise in the initial stages (S1 and S2), reaching maximal expression at stage 3, followed by a gradual decrease during the last stages of seed development (S4 and S5, Fig.3B).

We identified three putative castor bean LPAAT non-plastidic isoform encoding genes (*LPAAT2*, 3 and 4) using *AtLPAAT2* and *AtLAAT3* as queries for BLAST searches (Fig.2 and Supplementary Data 2B and 7). All non-plastidic LPAATs contain four domains (ATMs I to IV) conserved across LPAATs in diverse organisms and five transmembrane domains (TmD in Supplementary Data 2B). The analysis of gene

expression reveals non-plastidic LPAATs have distinct expression patterns during castor seed development (Fig.3B). *LPAAT2* has the highest relative expression within the non-plastidic LPAAT family. *LPAAT2* shows a bell-shaped expression pattern, with a linear rise in the initial stages (1 and 2), achieving peak expression at stage 3, followed by a decrease during stage 4 to a level maintained during stage 5. *LPAAT3* presents maximal expression in stages 1 and 2, followed by a drop over time. *LPAAT4* has higher expression at stage 2 and then shows a significant reduction in later stages; it could not be detected at stage 4 or 5 (Fig.3B).

Phylogenetic analysis showed that non-plastidic LPAATs form a cluster separate from plastidic LPAAT, similar to what has been reported for the G3PAT family (Fig.2). Interestingly, both plastidic G3PAT and LPAAT formed an independent cluster. These data may suggest a common origin for these two classes of genes with divergent functions during their evolution (Fig.2). However, more detailed analyses are necessary to confirm this hypothesis.

Glycerolipids in seed plants can be synthesized via two similar pathways occurring in distinct subcellular compartments. These two pathways have been described as the prokaryotic and eukaryotic systems on the basis of their evolutionary origins. The prokaryotic and eukaryotic glycerolipid synthesis pathways are similar in the two-step enzymatic conversion of G3P to PA but differ in the subsequent conversion of PA into structural, storage and signaling lipids (Banas *et al.*, 2000).

Although the two-step enzymatic acylation reactions of the prokaryotic and eukaryotic systems are similar, they are catalyzed by enzymes unique to their respective system (Kim *et al.*, 2005). Plastidic enzymes are responsible for the prokaryotic system, and cytoplasmic enzymes are used in the eukaryotic system. These enzymes present

structural differences, depending on their subcellular localizations. This would explain, at least in part, the genetic divergence between plastidic and cytoplasmic isoforms and the similarity between plastidic G3PAT and LPAAT enzymes, as observed in our phylogenetic approach.

2.3 Phosphatidic acid phosphatase (PAP)

PAP catalyzes the dephosphorylation of phosphatidic acid (PA) to yield *sn*-1,2 diacylglycerol (DAG) and inorganic phosphate. DAG produced from the hydrolysis of PA is not only a direct precursor of TAG, but also a substrate for the synthesis of membrane phospholipids (Nakamura *et al.*, 2007). PAP involved in glycerolipid synthesis appears to exist in both soluble and particulate forms (membrane-associated fractions), with the distribution of these forms being affected by cellular metabolic status. The intracellular translocation of PAP may help to regulate lipid metabolism (Kocsis and Weselake, 1996). PAP activity located in the microsome-associated fractions of safflower (*Carthamus tinctorius* L.) and avocado (*Persea americana*) is thought to be implicated in the so-called eukaryotic pathway of lipid biosynthesis, while the activity associated with the inner membrane of the plastid envelopes of spinach (*Spinacia oleracea*) and pea (*Pisum sativum*) is thought to be implicated in the prokaryotic pathway of glycolipid synthesis (Franca *et al.*, 2008).

Heterologous expression of *Arabidopsis* PAP (*AtLPP1*, At2g01180) cDNA in yeast showed that the encoded protein is indeed a lipid phosphate phosphatase that exhibits diacylglycerol pyrophosphate phosphatase (DGPP) and PAP activities. A second gene (*AtLPP2*, At1g15080), also encoding a PAP enzyme, has been identified in the *Arabidopsis* database (Pierrugues *et al.*, 2001).

Using four known *A. thaliana* PAP enzymes as BLAST queries (including AtLPP1, AtLPP2, At3g02600 and At3g18220), we identified four putative castor bean PAP genes, namely *PAP1-4* (Fig.2 and Supplementary Data 3 and 7). Phylogenetic analysis demonstrated that all castor PAP genes formed a single cluster, with branches showing their similarity to the *Arabidopsis* orthologs (Fig.2).

All PAPs have six transmembrane domains (TmD) and three conserved phosphatase sequence motifs (KxxxxxxRP, PSGH and SRxxxxxHxxxD), which were named PAPm1-3. These motifs are also shared among several lipid phosphatases, mammalian glucose-6-phosphatases and a collection of bacterial nonspecific acid phosphatases (Stukey and Carman, 1997) and shown in Supplementary Data 3. The importance of PAP motifs has been confirmed by mutational analysis of the *S. cerevisiae* *DPP1* and mouse *LPP1* genes (Pierrugues *et al.*, 2001). It was demonstrated by a direct mutational approach that single amino acid changes in each phosphatase consensus domain in mouse PAP are sufficient to produce a severe loss of PAP activity (Zhang *et al.*, 2000). The mutations K¹⁶⁸R, R¹⁷⁵K, P¹⁷⁶I, S²¹⁵T, H²¹⁷L, R²⁶⁵K and H²⁷¹L (mutated residues are indicated by black arrows in Supplementary Data 3) from the three domains decreased PAP activity by more than 95% compared to the wild-type mouse PAP (Zhang *et al.*, 2000).

The qPCR analysis showed that castor bean *PAP1* mRNA accumulated to its highest level at stage 1 of seed development, followed by a linear decrease over time (Fig.3C). *PAP3* presented a linear increase in expression until stage 3, followed by a decrease in stages 4 and 5 (Fig.3C). These results indicate the importance of *PAP1* in initial castor seed development stages. From stage 3 onward, *PAP1* showed the same mRNA accumulation pattern as *PAP3*.

With respect to *PAP2* and *PAP4*, we were unable to detect any expression in samples from developing seeds under our experimental conditions. The involvement of PAP with other cellular process might explain, at least in part, their undetectable expression along seed development. For example, besides its important role in glycerolipid synthesis, PAP is also involved in general phospholipid degradation and turnover. Some studies have shown that PA accumulates in plants in a transient manner in response to various forms of stress. The dephosphorylation activity of PAP results in the attenuation of the signaling function of PA through its conversion to DAG. This activity can be important in the response to stress, especially in the modeling of membrane composition (Kocsis and Weselake, 1996; Pierrugues *et al.*, 2001).

2.4 Diacylglycerol acyltransferase (DGAT)

DGAT catalyzes the third acyl-CoA-dependent acylation reaction that leads to the production of TAG. DGAT is an integral ER protein and has also been shown to be present in oil bodies and plastids (Siloto *et al.*, 2009). The substrate selectivity of DGAT depends on several factors, including the acyl composition of the DAG pool, acyl-CoA concentration and temperature. The temperature effect on acyl-CoA selectivity of DGAT has been attributed to differential changes in substrate-enzyme affinity and solubility of the substrate (Lung and Weselake, 2006).

DGAT activity is found in at least two distinct membrane-bound polypeptides, referred to as DGAT1 and DGAT2 (Cases *et al.*, 1998; Cases *et al.*, 2001). The first DGAT gene was cloned from mouse and is a member of the DGAT1 family, which has high sequence similarity to sterol:acyl-CoA acyltransferase (Cases *et al.*, 1998). Several homologues of DGAT1 have been found in animals and plants, and their functions have

been verified by both overexpression and deletion approaches (Katavic *et al.*, 1995; Hobbs *et al.*, 1999; Zou *et al.*, 1999; Bouvier-Nave *et al.*, 2000). Mutation of the DGAT1 gene in *Arabidopsis* (*AtDGAT1*) severely affects seed development, causing the appearance of wrinkled and incompletely filled seeds and aberrant seedling growth, reducing seed oil content and increasing the DAG/TAG ratio and the beta-oxidation of elevated FA levels (Lung and Weselake, 2006; Xu *et al.*, 2008).

Through BLAST searches using *AtDGAT1* as a query, we were able to identify one putative castor bean *DGAT1* gene (Fig.2; Supplementary Data 4A and 7), which is identical to the two castor *DGAT1* sequences already deposited in Genebank: AY366496 (He *et al.*, 2004) and gi 170676815 (Burgal *et al.*, 2008). Predicted structural features of the castor bean *DGAT1* revealed a region spanning R¹¹⁸-SHAGLF-K¹⁵⁷ (Cons. Reg., Supplementary Data 4A) that is highly conserved among other *DGAT1*s. This region contains the putative acyl-CoA binding signature spanning residues R¹¹⁸-G¹³⁴ (ABM; Supplementary Data 4A) as well as the putative active site catalytic residues R¹⁴⁹LIEN¹⁵⁴ (CR, Supplementary Data 4A) (Zou *et al.*, 1999). The importance of this motif has been demonstrated through insertional mutation (81 bp) in the *A. thaliana DGAT1* gene, which resulted in a 27-amino-acid tandem repeat in this region, leading to a decrease in seed oil content (Katavic *et al.*, 1995; Zou *et al.*, 1999). *DGAT1* also contains a putative DAG/phorbol ester binding motif (P.E. Dom, Supplementary Data 4A) that is a signature sequence unique to *DGAT* and absent from acylCoA:cholesterol acyltransferases, as previously reported (Billheimer *et al.*, 1990). This motif contains a phenylalanine residue (black star, Supplementary Data 4A) that has been demonstrated to be a key determinant of oil content and composition in maize (*Zea mays*) and to be conserved within the *DGAT* family (Zheng *et al.*, 2008). *DGAT1* possesses a consensus sequence

(xL(Y)²⁰⁰xK²⁰²xxS²⁰⁵xxxV²⁰⁹) identified as a targeting motif typical of members of the SnRK1 protein kinase family (see Supplementary Data 4A). SnRK1 proteins are a class of serine-threonine protein kinases that have been increasingly implicated in the global regulation of carbon metabolism in plants (Halford and Hardie, 1998; Zou *et al.*, 1999). A putative thiolase acyl-enzyme intermediate-binding motif is also present (THIOm, in Supplementary Data 4A). It contains an invariant Pro²²⁴ at the *N*-terminus of the motif that has been suggested to participate in presenting the fatty acyl group to the active site for esterification into DAG. There is also a fatty acid binding protein signature spanning sequence (FAS), which contains a putative tyrosine phosphorylation site (Y, indicated by a black asterisk in Supplementary Data 4A) (Lewin *et al.*, 1999). Additionally, *DGAT1* presents two potential DGAT motifs (Supplementary Data 4A), H257xxxD261 and H348xxxxD353, having three or four amino acids, respectively, between what have been postulated to be critical histidine and aspartic acid residues (Daniel *et al.*, 2004), a signature which is also described for other acyltransferase family members (Saha *et al.*, 2006; Xu *et al.*, 2006; Xu *et al.*, 2008).

A second family of DGAT genes (*DGAT2*) was first identified in the oleaginous fungus *Mortierella ramanniana*, having no sequence similarity to *DGAT1* (Lardizabal *et al.*, 2001; Xu *et al.*, 2008). A recent study of tung tree DGATs has suggested that, in plants containing unusual fatty acids, *DGAT2* may play a more important role in channeling unusual fatty acids into seed storage oils (Shockey *et al.*, 2006; Xu *et al.*, 2008). The expression of castor bean *DGAT2* (GI 170676815) selectively increases the hydroxylated-fatty-acid (HFA) content of the seed oil from 17% to nearly 30% of total fatty acids in *A. thaliana* (Kroon *et al.*, 2006). Moreover, in enzymatic assays, recombinant *DGAT2* from castor bean exhibited an approximately 10-fold higher preference for 18:1-

OH-DAG over 18:1- or 18:2-DAG, suggesting that this protein may indeed have co-evolved in castor bean for the efficient transfer of HFA into oil.

We identified a putative DGAT2 gene in castor bean using *A. thaliana* DGAT2 (*AtDGAT2*) as a query for BLAST searching (Fig.2 and Supplementary Data 4B and 7). DGAT2, as identified in the Castor Bean Genome Database, differs from the DGAT2 sequence deposited in GenBank (GI 170676815, EU391592, DQ923084 and AY916129) in 29 amino acids at the *N*-terminal region. Like many other DGAT2 isoforms, *DGAT2* contains two closely spaced membrane-spanning domains (TmD) (Shockey *et al.*, 2006) and a *C*-terminal aromatic ER retrieval motif (RE in Supplementary Data 4B) (McCartney *et al.*, 2004).

DGAT2 presented higher mRNA accumulation than *DGAT1* in all five stages of seed development (Fig.3D). *DGAT1* showed low expression in stage 1, followed by increased mRNA accumulation in stages 2 and 3 to levels that were maintained up to the end of stages 4 and 5. On the other hand, *DGAT2* presented an increasing expression pattern until stage 3, followed by decreasing expression in stage 4 and stage 5 (Fig.3D).

The roles of *DGAT1* and *DGAT2* in oil production can be species-dependent. In some plants, *DGAT1* may play a more dominant role depending on gene expression patterns and protein accumulation, whereas in plants containing unusual fatty acids, *DGAT2* may play a more important role (Shockey *et al.*, 2006). The latter seems to describe the situation in castor beans. Considering that castor oil contains about 90% ricinoleic acid and that *DGAT2* presents a high affinity for unsaturated fatty acids, it is plausible that *DGAT2* would show higher relative expression than *DGAT1*, especially in the final stages of seed development. Therefore, our results support the hypothesis that *DGAT2* plays a more critical role in the final step of castor bean TAG biosynthesis

(Shockey *et al.*, 2006).

The high DGAT activity in developing seeds has been demonstrated to coincide with the phase of rapid lipid accumulation and subsequently to be reduced when the lipid content in the seeds reaches a plateau (Lung and Weselake, 2006). It has been suggested that DGAT activity may represent one of the rate-limiting steps in the accumulation of plant storage lipids and, therefore, represent a potential biotechnological target for modulating the fatty acid composition of seed oil crops, especially given the recent reports on the specific involvement of DGAT2 in channeling specific fatty acids into TAG (Xu *et al.*, 2008; Siloto *et al.*, 2009).

2.5 Phospholipid diacylglycerol acyltransferase (PDAT)

In the acyl-CoA-independent pathway of TAG biosynthesis, a phospholipid such as phosphatidylcholine (PtdCho) or phosphatidylethanolamine (PtdEtn) can donate its *sn*-2 acyl group to 1,2-diacyl-*sn*-glycerol (DAG), resulting in the formation of TAG. This transacylase activity is catalyzed by an enzyme termed phospholipid: diacylglycerol acyltransferase (PDAT), which is homologous to mammalian lecithin cholesterol acyltransferase (LCAT) (Banas *et al.*, 2000; Mhaske *et al.*, 2005).

The specificities of PDATs from different plant sources toward acyl donors are species dependent. Studies of PDAT activity have shown preferential removal of ricinoleic acid and vernolic acid from PtdCho in microsomal fractions from *Ricinus communis* and *Crepis palaestina*, respectively (Dahlqvist *et al.*, 2000; Lung and Weselake, 2006). PDAT has been implicated in the selective channeling of ricinoleic acid from phospholipids into TAG in the developing endosperm of the castor bean. Ricinoleate and other unusual fatty acids produced by oleate hydroxylase (FAH12) are synthesized

from acyl groups linked to phospholipids and then rapidly transferred from those phospholipids to TAG (Bafor *et al.*, 1991; Dahlqvist *et al.*, 2000; Ghosal *et al.*, 2007).

Using *A. thaliana* PDAT (AtPDAT, At5g13640) as a query for BLAST searches, we were able to identify six putative castor bean *PDAT* genes, *PDAT1* to 6 (Fig.2 and Supplementary Data 5 and 7). Amino acid alignment, followed by phylogenetic analysis, revealed that castor *PDATs* formed a single cluster, as shown in Fig.2. *PDAT1*, 2 and 3 possess C-terminal aromatic amino acid-rich ER retrieval motifs (RE, Supplementary Data 5), reported in *Arabidopsis* PDAT-like proteins (McCartney *et al.*, 2004) and also possess signal peptides directing them to the ER (SP, Supplementary Data 5), as previously described (Krogh *et al.*, 2001) (Schrag and Cygler, 1997; Peelman *et al.*, 1998; Krogh *et al.*, 2001). These two motifs, however, are not conserved in DGAT4, 5 and 6 from castor bean. The conserved residues S274, D607, and H660 are present in all castor bean PDAT proteins (black arrows in Additional File 5). In human LCAT (HsLCAT) it has been demonstrated, by site-directed mutagenesis, that these residues have catalytic importance and are involved in the phospholipase A activity of the enzyme (Peelman, 1998). PDATs also contain so-called lid domains. It has been suggested that this domain destabilizes the lipid bilayer and facilitates both the binding of the hydrophobic substrate and its diffusion into the active site cavity of the enzyme (Peelman *et al.*, 1999). A Trp¹⁵⁴ residue present in the lid (black star in Supplementary Data 5) is thought to bind the cleaved fatty acid in the active site of the enzyme (Martinelle *et al.*, 1996; Stahl *et al.*, 2004).

The expression profile of this gene family was also analyzed. *PDAT1* was not detectable in any stage of seed development. On the other hand, *PDAT2*, 3 and 5 were expressed and showed similar, bell-shaped expression profiles. Their mRNA levels

increased until stage 3, where peak expression was achieved, followed by continuous decreases in stages 4 and 5 (Fig.3E). *PDAT4* expression was detected, but at a very low level as compared with other *PDAT* genes. *PDAT4* showed maximum expression in stage 2 and decreased over time. *PDAT6* showed a linear increase until stage 4 and then presented a reduction of expression in the final developmental stage (Fig.3E). *PDAT6* also presented the highest relative expression of any castor bean *PDAT* gene at all stages of seed development, with the exception of stage 3, where *PDAT2* presented a higher relative expression (Fig.3E).

The differences in the expression patterns showed by these *PDAT* genes may indicate that divergent temporal control among the members of this gene family during seed development may result in a thin adjustment of membrane turnover and/or TAG accumulation, revealing a well controlled subfunctionalization of different genes of this family. *PDAT* enzymes in oilseed plants may play a critical role in the removal of unusual fatty acids from membrane phospholipids by the selective transfer of these fatty acids into TAGs (Lung and Weselake, 2006). The balance of the expression of *PDAT* genes, in addition to their importance in TAG biosynthesis, may represent a mechanism for membrane lipid homeostasis (by contributing to membrane lipid turnover) in specific circumstances or sites or for the removal of DAG, a product of the phosphatidylinositol signaling pathway, critical for plant stress responses (Mhaske *et al.*, 2005).

2.6 Choline phosphotransferase (CPT)

CPT, or amino-alcohol phosphotransferase (AAPT), is an integral membrane bound protein able to use both an insoluble substrate, DAG, and a soluble substrate, CDP-choline (Slack *et al.*, 1985). CPT is a key enzyme in oilseed metabolism because it

catalyzes both the synthesis of PC from DAG and the reverse reaction, one route by which polyunsaturated fatty acids (PUFAs) are made available for incorporation into TAG molecules (Somerville C *et al.*, 2000). It has been hypothesized that the continuous reversible transfer of DAG into PC may control, in part, the PUFA content of the seed oil. These PUFAs often contain unusual fatty acids (Stobart and Stymne, 1985; McMaster and Bell, 1997). As assessed by the association of specific activities of CPT with organelle marker enzymes, most membrane fractionation studies have concluded that the bulk of CPT activity resides in the ER (McMaster and Bell, 1997).

We have identified one putative castor bean *CPT* gene using *A. thaliana* *CPT* enzymes (At1g13560 and At3g25585) as queries (Fig.2 and Supplementary Data 6 and 7). Alignment of the predicted protein sequence of *CPT* derived from the known DNA sequences of yeast, soybean, *A. thaliana* and *Brassica campestris* *CPT* genes revealed high conservation. Sequence comparison analysis identified a completely conserved CDP alcohol phosphotransferase motif: DGxxARxxxxxxxxGxxxDxxxD (McMaster and Bell, 1997; Williams and McMaster, 1998; Goode and Dewey, 1999). This motif is also present in castor bean *CPT* (Supplementary Data 6). The catalytic role of the phosphotransferase motif in *CPT* from *Brassica napus* (BnAAPT1) was confirmed by scanning alanine mutagenesis (Qi *et al.*, 2003). The authors demonstrated that Asp-Ala, Gly-Ala, Ala-Gly, Arg- Ala and Gly-Ala substitutions resulted in 14%, 78%, 44%, 65% and 88% reduction in BnAAPT1 enzyme activity, respectively, and that Asp¹²⁶Ala and Asp¹³⁰Ala substitutions effectively abolished BnAAPT1 activity. These conserved residues are present in castor bean *CPT* (Supplementary Data 6, black arrows). Additionally, seven potential transmembrane domains (TmD) are predicted to be present in the *CPT* enzyme (Supplementary Data 6).

The qPCR analysis demonstrated that castor bean *CPT* shows increased expression during stage 2 (about four fold higher than in stage 1) to a level that was maintained in stage 3, followed by a gradual drop during the last stages of seed development (stages S4 and S5, Fig.3F).

Molecular and biochemical characterization of CPT from soybean (AAPT1) and *Arabidopsis* (AtAAPT1p and AtAAPT2p) revealed that these plant AAPTs can utilize both CDP-choline and CDP-ethanolamine as substrates, but with different preferences and CMP inhibition properties (Bowden and Lord, 1975). The same occurs in the castor bean endosperm, where a single enzyme apparently possesses choline and ethanolamine phosphotransferase activities (Bowden and Lord, 1975). These results further suggest that, in higher plants, CPT is a dual-function enzyme (Bowden and Lord, 1975; Dewey *et al.*, 1994; Goode and Dewey, 1999).

3 Concluding remarks

Modification of seed oil composition could contribute to the production of nutritionally and industrially desirable oils (Kocsis and Weselake, 1996). However, this requires manipulation of fatty acid and TAG synthesis in such a way that a specific fatty acid is synthesized at high rates and effectively incorporated into each position of TAG (Weier *et al.*, 1998). Although genes encoding suitable fatty acid-modifying enzymes are available from many species in the wild, progress has been limited because the expression of these genes in transgenic plants has produced low yields of the desired oils (Burgal *et al.*, 2008).

TAG biosynthesis requires coordinated regulatory control by seed developmental programs at multiple levels, such as endosperm cellularization, gene transcription and posttranscriptional regulation. In this scenario, gene transcription may exert primary control over castor oil biosynthesis.

Our results represent snapshots of the global transcriptional activities of twenty-six genes encompassing six enzyme families involved in castor bean TAG biosynthesis throughout castor seed development. The fatty acid specificity and selectivity of these enzymes determines the fatty acid pattern in TAG molecules and, thus, their quality and suitability for use in defined markets (Weier *et al.*, 1998). Certainly, further studies encompassing posttranscriptional analyses and enzymatic assays are required to ascertain the activity of the enzymes encoded by these genes. However, the results presented here can contribute to the identification of potential targets of future biotechnological approaches that may lead to the development of genetic strategies to manipulate oil fatty acid synthesis.

4 Experimental

4.1 *Plant material*

Castor seeds (Al Guarany commercial variety) were obtained from EMBRAPA-CPACT, Pelotas, Rio Grande do Sul State in Brazil. Dissected seeds were frozen immediately in liquid nitrogen and stored at -80°C . Expression analysis was performed at five stages of seed development. The criteria adopted to discriminate the five stages were based on morphologic characteristics, such as size, color and texture of the seed (Fig.1A). The first stage, S1, corresponds to seeds that are not totally expanded and are less than 1 cm in length. Fully expanded and non-pigmented seeds were grouped in stage S2. Stage S3 seeds presented only slight pigmentation, approximately less than 20% of the final pigmentation. Totally pigmented seeds from immature and green fruits were placed in the S4 stage. S5 seeds were collected from mature dehiscent fruits.

4.2 *Isolation of RNA and cDNA synthesis*

Total RNA was extracted from castor bean seeds using the NucleoSpin[®] RNA Plant kit (Macherey-Nagel), and mRNA was isolated using the PolyATtract[®] mRNA Isolation System III (Promega), according to the manufacturer's instructions. Reverse transcription of first-strand cDNA was carried out by using a T25V primer and 200 units of M-MLV reverse transcriptase (Promega) in a final volume of 50 μL . The cDNA products were diluted 1:10 and stored at -80°C .

4.3 *BLAST searches*

BLAST searches (tBLASTx) were performed against the Castor Bean Genome

Database (<http://castorbean.jcvi.org>) using known *A. thaliana* enzymes from *Plant Metabolic Network* (PMN – <http://www.plantcyc.org>) as queries (Supplementary Data 1 to 6), in order to identify all genes related to those involved in TAG biosynthesis. The choice of queries only from PMN databank was based in the fact of their enzyme activities was experimentally tested. The inclusion criteria for putative genes required more than 50% amino acid identity with *Arabidopsis* queries. The initially selected castor bean genes were used for a second tBLASTx round against the Castor Bean Genome Database, using the same criteria as the first BLAST approach. This extra step permitted us to find any other castor bean paralogs excluded by their dissimilarity to the *Arabidopsis* orthologs.

4.4 Phylogenetic analysis

Deduced amino acid sequences from coding sequences of the identified genes were aligned using Molecular Evolutionary Genetics Analysis (Mega 4.1) (Tamura *et al.*, 2007) and edited with GeneDoc software (<http://www.psc.edu/biomed/genedoc>). *In silico* characterization of protein domains and characteristic signatures of enzyme classes was performed using available literature data. Transmembrane domain prediction was carried out using the TmPred server (http://www.ch.embnet.org/software/TMPRED_form.html). Phylogenetic trees were drawn using Mega 4.1 software. Phylogenetic relationships were inferred from neighbor-joining amino acid alignment after 1000 bootstrap replicates, using pairwise deletion amino acid p-distance parameters.

4.5 Primer design and quantitative polymerase chain reaction

Primers were designed using Primer3 software

(http://frodo.wi.mit.edu/primer3/primer3_code.html) to produce amplicons between 100 and 250 bp. qPCR was carried out with an initial step of 5 minutes at 94°C followed by 40 cycles of 10 seconds at 94°C, 15 seconds at 60°C and 15 seconds at 72°C. The melting curve profiles for all PCR products were obtained after incubation at 40°C for 2 minutes for reannealing followed by a step-wise warming from 55°C to 99°C to acquire relative denaturizing curve data. The qPCRs were carried out using 10 µL of cDNA sample (diluted 1:100), 2.5 µL of 10X PCR buffer (100 mM Tris/HCl, pH 8.0), 1.2µL of 50 mM MgCl₂, 0.4 µL of 5 mM dNTPs, 0.2 µL each of forward and reverse primers (shown in Supplementary Data 7), 3.45 µL of H₂O, 2.0 µL of SYBR-Green (1:10.000) and 0.05 µL of Platinum Taq DNA Polymerase (5 U/µL; Invitrogen), in a final volume of 20 µL. As internal reference genes, we took advantage of recent work by our group that analyzed the expression pattern of 14 candidate house-keeping genes in leaf and castor seed tissues (unpublished data). The two genes with the most stable expression pattern throughout seed development were used as normalizing house-keeping genes in the present study. They were EF1B/elongation factor 1-beta (Castor ID, 29785.m000934; forward primer 5'-GCAGTTCGGAGCATTGAGAT-3' and reverse primer 5'-GTCGTCCACAATGGTCATCA-3') and Ubi9/ubiquitin (Castor ID, 30169.m006323; forward primer 5'-ATCGATCGAATCAAGGAACG-3' and reverse primer 5'-CACCTCAATGTTGTAGTCACG-3'). For each selected gene, two independent experiments were done, with eight biological samples for each of the five stages of castor seed development. Each biological sample consisted of a pool of two to three castor seeds. Quadruplicate experimental sets of qPCR reaction samples, including the reference control genes and quadruplicate negative controls (reaction samples without cDNA templates), were prepared and run in 96-well plates. SYBR

fluorescence was analyzed by StepOne Software version 2.0 (Applied Biosystems), and the CT value for each sample was calculated and reported using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Authors' contributions

AC have made substantial contributions to conception and acquisition of data, and also to analysis, interpretation of data and writing the draft of the manuscript. GL and ME participated in the sample collection and in the preparation of RNA and cDNA. RM participated in the experimental design, sequence alignment and acquisition and interpretation of qPCR data. MMP conceived of the study, and participated in its coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This research was supported by FINEP (Financiadora de Estudos e Projetos) and FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul) “projeto Estruturante de Agroenergia”. G. Loss and A. Cagliari have Master fellowships from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and M. Margis-Pinheiro and R. Margis are recipients of CNPq research fellowships number 308708/20067 and 303967/2008-0, respectively. The authors also thank Dr. Sergio Delmar dos Anjos e Silva from EMBRAPA for castor bean seed supply.

5 Figures

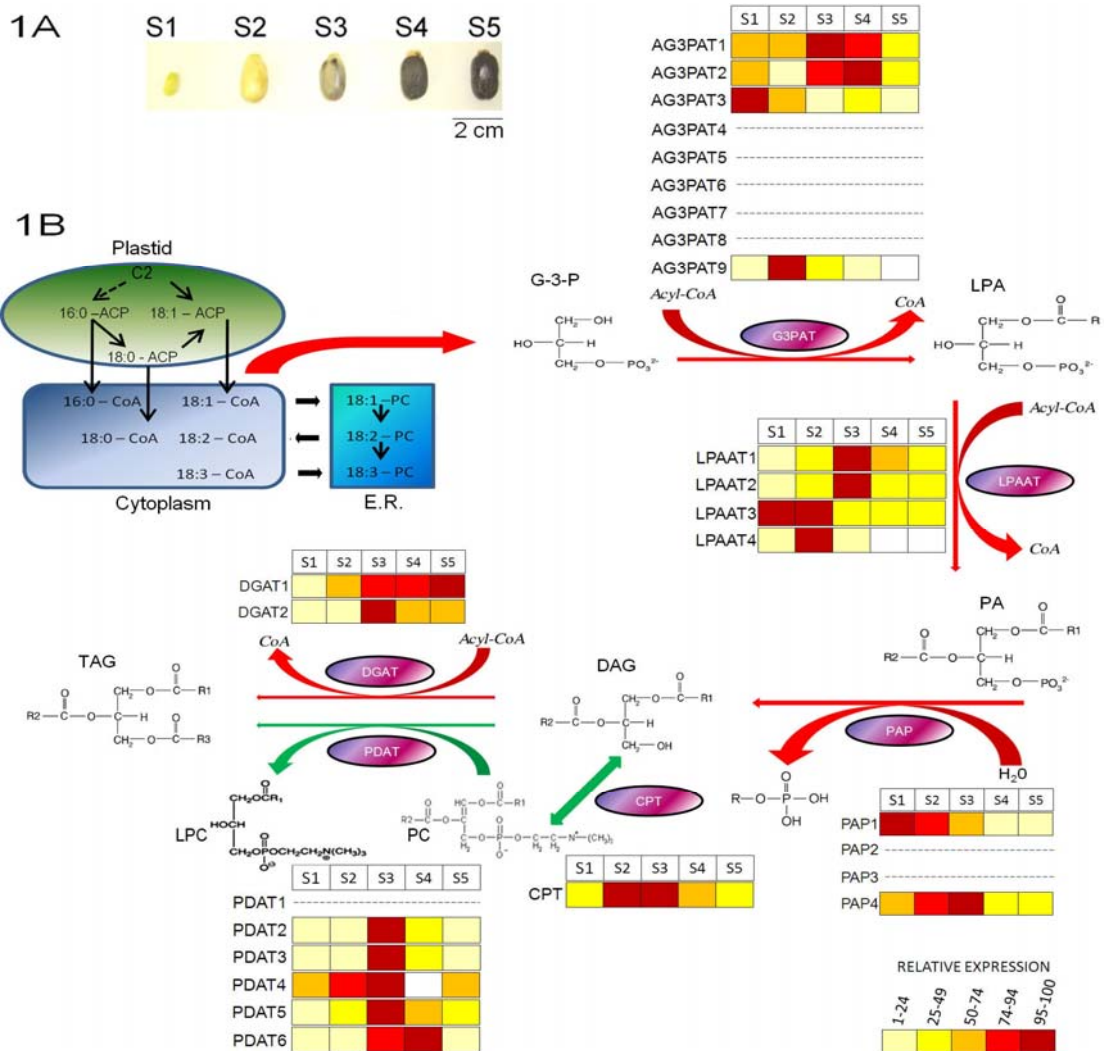


Figure 1. Pathway of triacylglycerol biosynthesis.

(A) The five stages of castor bean seed development (S1 to S5). **(B)** General scheme of TAG synthesis in plants. Fatty acid (FA) synthesis occurs in plastids. The *de novo* synthesized FA enters the cytosolic pool in an esterified form known as acyl-CoA. In the next step, FA is incorporated into a glycerol backbone, leading to formation of TAG. Red arrows indicate the acyl-CoA-dependent pathway (Kennedy pathway). Green arrows indicate alternative TAG biosynthesis pathways (acylCoA-independent). Enzymes involved in TAG biosynthesis are shown as oval shapes: glycerol-3-phosphate acyltransferase (G3PAT), lysophosphatidic acid acyltransferase (LPAAT), phosphatidate phosphatase (PAP), diacylglycerol acyltransferase (DGAT), phospholipid:diacylglycerol acyltransferase (PDAT) and choline phosphotransferase (CPT). Colored boxes indicate the relative mRNA expression level of genes (%) at each stage of seed development (S1 to S5), in comparison with the maximal expression level observed, from two independent expression experiments. Dotted lines indicate seed stages where gene expression was not detected under our experimental conditions. ACP, acyl carrier protein; G3P, glycerol-3-phosphate; LPA, lysophosphatidic acid; PA, phosphatidic acid, DAG, diacylglycerol; TAG, triacylglycerol; PC, phosphatidylcholine; LPC, lysophosphatidylcholine.

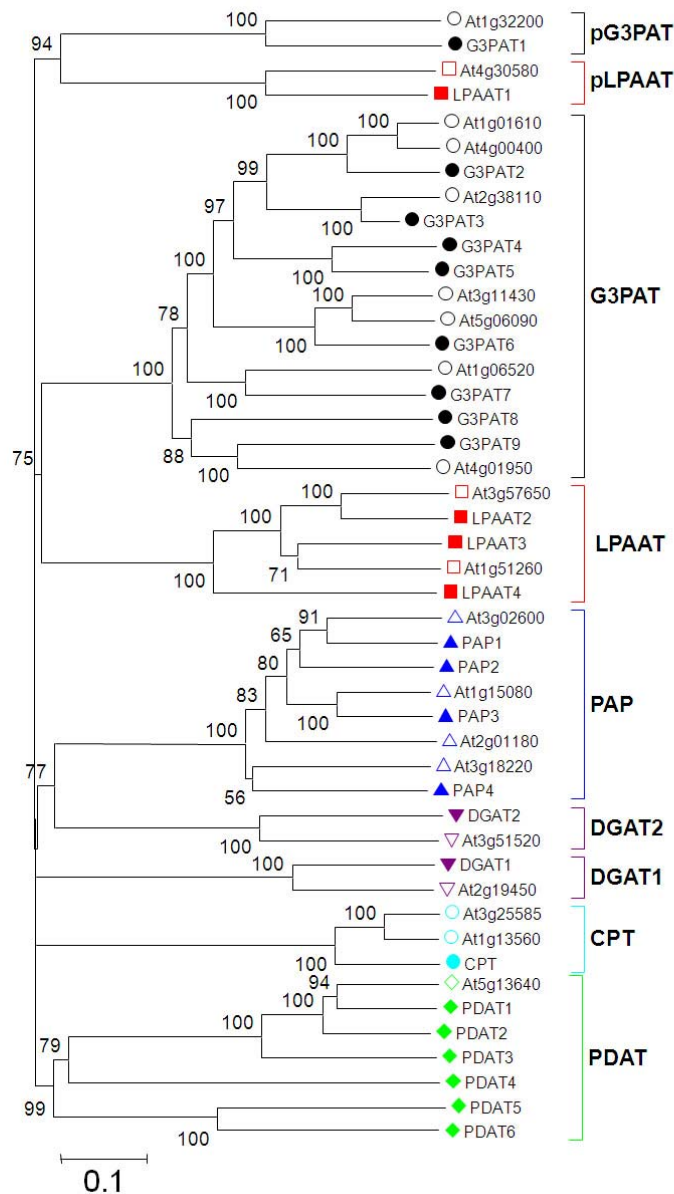


Figure 2. Phylogenetic analysis of genes encoding enzymes from the TAG biosynthetic pathway. The phylogenetic tree was constructed by the neighbor-joining method, using the MEGA 4.1 program, after 1,000 bootstrap replications, with pairwise deletion and amino acid p-distance as the main parameters. Bootstrap values higher than 70% are given at the respective nodes. The scale bar indicates genetic distance. *Arabidopsis* enzyme queries are designed by their respective *Arabidopsis thaliana* Genome Database accession numbers (for example, At1g32200). pG3PAT, plastidic glycerol-3-phosphate acyltransferase; G3PAT, cytoplasmic glycerol-3-phosphate acyltransferase; pLPAAT, plastidic lysophosphatidic acid acyltransferase; LPAAT, cytoplasmic lysophosphatidic acid acyltransferase; PAP, phosphatidic acid phosphatase, DGAT, diacylglycerol acyltransferase; PDAT, phospholipid diacylglycerol acyltransferase; CPT, choline phosphotransferase.

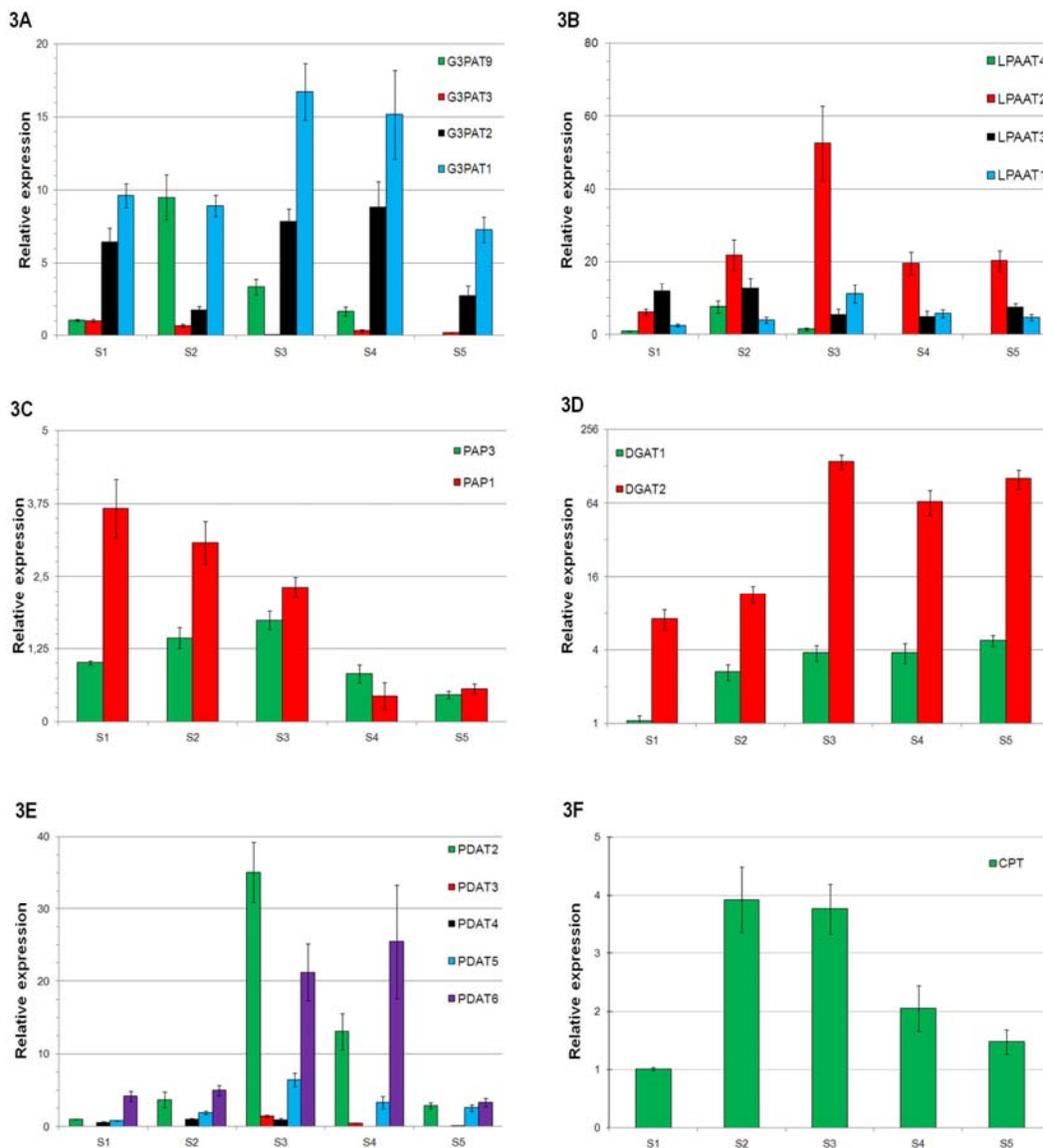
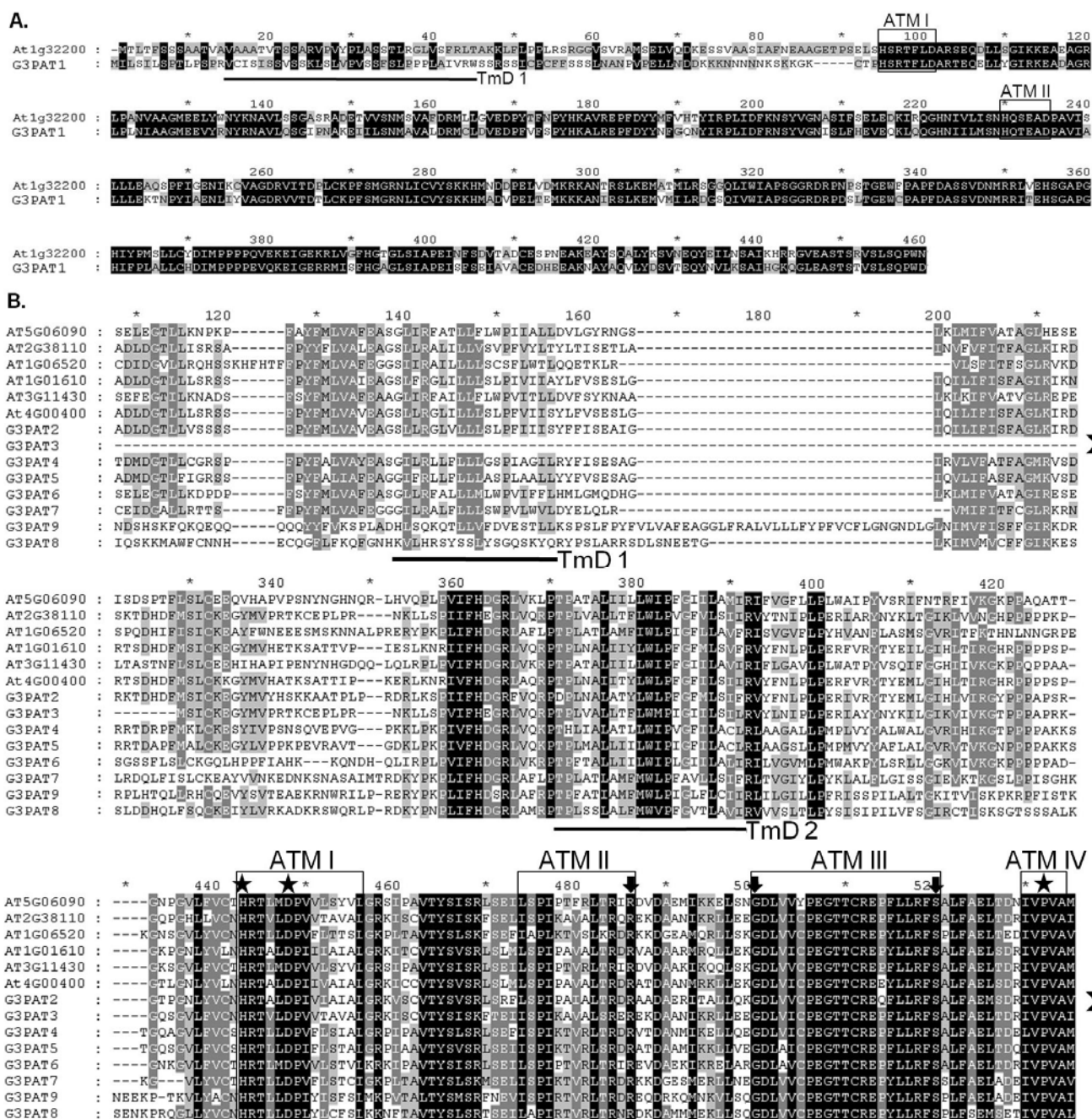


Figure 3. Relative expression pattern of genes encoding enzymes related to TAG biosynthesis. Relative expression of genes involved in TAG biosynthesis during castor seed development (stages S1 to S5). Each data point represents the mean \pm standard deviation of two independent experiments, with eight biological and four experimental replicates. (A) glycerol-3-phosphate acyltransferase: G3PAT1 is a plastidic isoform, while G3PAT2, 3 and 9 correspond to cytoplasmic isoforms; (B) lysophosphatidic acid acyltransferase: LPAAT1 is a plastidic isoform, while LPAAT2, 3 and 4 correspond to cytoplasmic isoforms; (C) phosphatidic acid phosphatase, PAP; (D) diacylglycerol acyltransferase, DGAT (in this panel, the ordinate, corresponding to the relative expression of genes is plotted in log base 2 scale to allow a better comparison of groups of genes presenting large differences of values in their expressions); (E) phospholipid diacylglycerol acyltransferase, PDAT and (F) choline phosphotransferase, CPT.



Supplementary Data 1 - Deduced amino acid sequence alignment of the members of the *Arabidopsis* G3PAT enzyme family. *Arabidopsis* enzymes queries, as designated by their respective database accession numbers (for example, At1g32200). G3PAT(n): castor bean genes. Identical and similar amino acid residues are shaded in black and gray, respectively. The transmembrane (TmD) regions are underlined. Acyltransferase motifs (ATMs) are boxed. (A) plastid G3PAT. (B) cytoplasmic G3PAT. The black stars indicate catalytically important residues, while black squares highlight residues critical for G3P binding. Double black arrows indicate interruptions in the alignment sequence.

A.

```

AT4G30580 : MDVASARSISSHPFYGKPCSSQSSLLIRISRDKVCCEGRISNGMTSETTSLHAVPSEKEMGSTRRTGIQMSNRSLRHDFRELDKRS PRSSQARDDITVADLSGAAT
LPAAT1    : -----MMAAESSELCIPRKLVAVEWAEELNPHEDSRDLHRNKLPRKTVVRSSELVGNAN

*          20          *          40          *          60          *          80          *          100
AT4G30580 : PDSESEPEEPIKLSRRLRGIEPEYVAGHSANFLIYLMIIHQPFVLLRDPYRRKPHHFIAKLWASISIIYPFYKINI EGLENLPSSTPAVYVSNHOSFLDIYTLTLLSGKSF
LPAAT1    : EESAHLSEDIKLSAKVRGICEYAVPATAAHELEVLMLVCHPFVLLRDRYRRKTOEPIAKIWAIMITVAPFEKIEEGLENLPSSTPAVYVSNHOSFLDIYTLTLLSGKSF

*          120         *          140         *          160         *          180         *          200         *          220
AT4G30580 : KFIKSTGIFVPIIGWAMSMGVLPLKRMDSRSQVDCIKRRCMELLKKGASVFFPEGTRSKDGRLGSPKKGAFVAAKTVAVVPITLMGTGKIMPGSSSILNHNVR
LPAAT1    : KFIKSTGIFVPIIGWAMSMGVLPLKRMDSRSQVDCIKRRCIELVKKGASVFFPEGTRSKDGRLGSPKKGAFVAAKTVAVVPITLMGTGKIMPGSSSILNHNVR

*          240         *          260         *          280         *          300         *          320
AT4G30580 : VTIHKPTHGSKADVLCNEARSKIAESMDI*-- :
LPAAT1    : VTIHKPVSSNPVLCNEARNTIASVNLNQQS* :

*          340

```

B.

```

At1G51260 : MKIPAAWVPIPVGVLELISGLIVNIHQLFPIIVPESSRSLYRRINRNVAEILLWQLIWLFDWVAQIKINLIVDAETLELIGKEHALVSNHRSDIETWLVGVMQAQVVG
At3G57650 : -MVIAAAVTVPIGLPEHSGLAVNLECAVYVLIIPHSKTYRKNIRVABTLWLELVWVVDWVAGKICVADNETENRNGKEHALVQNHRSDIETWLVGVIQAQRSG
LPAAT3    : MAIPAAALVPIPIGLPEVSGEVNLIQAVLPIILVFPVSKSLRRINRNVAEILLWLELIWLVDDWVAHKIEVVADETEPEEMGKEHALLIQNHRSDIETWLVGVIQAQRSG
LPAAT2    : MVAVAAVVILPELGVLEPISGLVNVLPALCFVLIIPHSKSYRMINRVAEILLWLELVWVVDWVAKIKIIVITDRETLCMGKEHALLIQRSDIETWLVGVMQAQRAG
LPAAT4    : MBVRAVLSIPIPEFVLELISGLIINLIQACAYILLVPIPSKRAHRRVIGAVTEILLWLELIFLMDWVSGEKVRLHTDPEETQIMGKEHALLMNHICDAMIFIMVLLAQRSS

*          20          *          40          *          60          *          80          *          100
At1G51260 : CLGSLAIMMKREAVLPIIGWAMSMGVLPLKRMDSRSQVDCIKRRCMELLKKGASVFFPEGTRSKDGRLGSPKKGAFVAAKTVAVVPITLMGTGKIMPGSSSILNHNVR
At3G57650 : CLGSLAVMKKSSKFLPVIGWAMSMGVLPLKRMDSRSQVDCIKRRCMELLKKGASVFFPEGTRSKDGRLGSPKKGAFVAAKTVAVVPITLMGTGKIMPGSSSILNHNVR
LPAAT3    : CLGSLAVMKKQAVLPIIGWAMSMGVLPLKRMDSRSQVDCIKRRCMELLKKGASVFFPEGTRSKDGRLGSPKKGAFVAAKTVAVVPITLMGTGKIMPGSSSILNHNVR
LPAAT2    : CLGSLAVMKKSSKFLPVIGWAMSMGVLPLKRMDSRSQVDCIKRRCMELLKKGASVFFPEGTRSKDGRLGSPKKGAFVAAKTVAVVPITLMGTGKIMPGSSSILNHNVR
LPAAT4    : CLRGALMVKRSMYIPIIGWAMSMGVLPLKRMDSRSQVDCIKRRCMELLKKGASVFFPEGTRSKDGRLGSPKKGAFVAAKTVAVVPITLMGTGKIMPGSSSILNHNVR

*          120         *          140         *          160         *          180         *          200         *          220
At1G51260 : PAIYDCTVITVHNQETPTLLRMSGSSSINLQMRHKMSSELPEPDDGIAQWQDDFFIKDAQLEMYPTKDFVSDLEVHCINRPIPLVIVVITLGLFVFGPKLLOQL
At3G57650 : PAIYDNTVITVETSTPEPTMLRLEKGSVYVWHYKCHSMKLDLPEDDALAQWCRDQFVAKDALLDQHIADDFPGQEQNIGRPIKSLAVVLSLAQVLTCAKPLHQA
LPAAT3    : PAIYDCTVVDKSQLAPMLRIFRRQSSVINLQIKRHSQOEVPETADGIAQWQDDFFIKDAQLEMYPTKDFVSDLEVHCINRPIPLVIVVITLGLFVFGPKLLOQL
LPAAT2    : PAIYDNTVIAHQRNSQPTMLRLEKGSVYVWHYKCHSMKLDLPEPDDAVAQWCRDQFVAKDALLDQHIADDFPGQEQNIGRPIKSLAVVLSLAQVLTCAKPLHQA
LPAAT4    : PAIYDNTLAVRGLRTPSLLGLFGRBETAVCHIKRYSKRELPESEDAVAQWCRDQFVAKDALLDQHIADDFPGQEQNIGRPIKSLAVVITLGLFVFGPKLLOQLRY

*          240         *          260         *          280         *          300         *          320
At1G51260 : SIWASRKHILLVFPVVIATITMOLLIQSSESQRSTPAKRPLOEQQLISA*-----
At3G57650 : QIFSSNRGITISALGLIITLQMOILLIRSSQSERSTPAKVVPAKPKDNHHPSSQTEKEK*-----
LPAAT3    : SEFNSTQVLSFVTLVVLVITMOLLHSSSEHSTPEPKLTFSDQAKERLLQK*-----
LPAAT2    : SLLSSKCIALSMSALAVVIMHLLIRFQSEHSTPAQVAPANHAQVAPENPQNGGEPSEKAENKQD*
LPAAT4    : SLLSTRKVIISLAAGALDAVFLHAVIEYIKLPQARKISMGNLTVQLKN*-----

*          340         *          360         *          380         *

```

Supplementary Data 2 -Deduced amino acid sequence alignment of the members of the LPAAT family. *Arabidopsis* enzymes queries are designated by their respective database accession numbers. LPAAT(n): castor bean genes. Identical and similar amino acid residues are shaded black and gray, respectively. The transmembrane (TmD) regions are underlined. Acyltransferase motifs (ATMs) are boxed. (A) plastidic LPAAT. (B) cytoplasmic LPAAT.

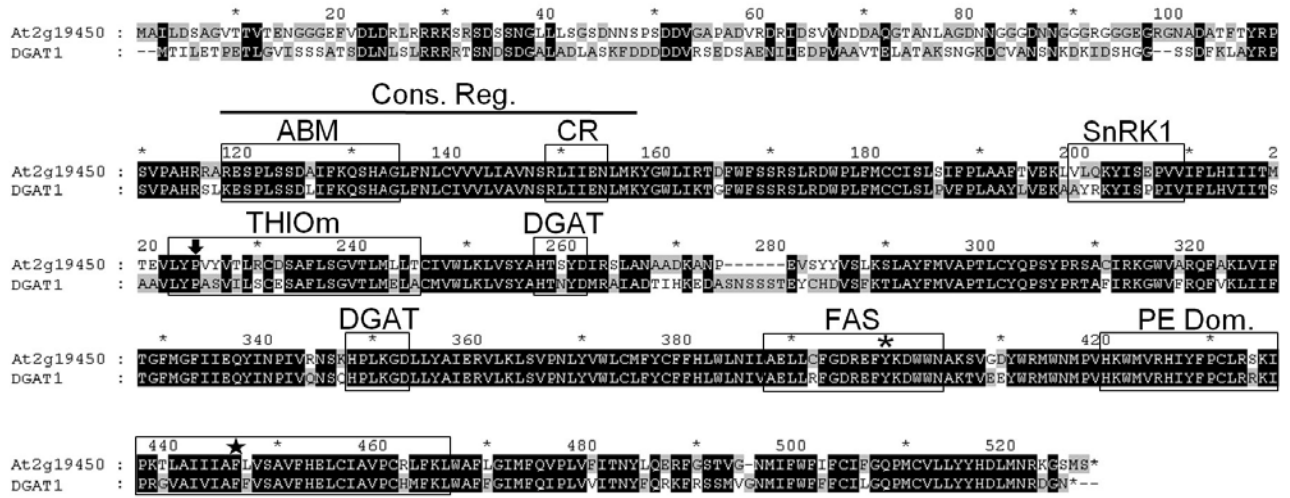
```

At3g02600 : MARFSFPCF PNF GGFNAQAVNTRGPEISETADNWNVSPSDIPLIEPNSKEHRMRBAQLGGHTLRSHGTVARTHMHDWIIILVLLVTECCVLLIHPFFRYVGRDMMDDL
At2g01180 : -----MTIGSFFSLLFWRNSQDQEAQRGRMQEIDLSVHTIKSHGGRVASKKHKDIIILVLLIATBGLNLISSPFRYVGRDMMDDL
At3g18220 : -----MAKMLGSHSVKSHGTVARBEHICDIIILVLLVCIIDIVLAVIIPFFRYVGRDMMDDL
At1g15080 : -----MPETHLGAHTIRSHGTVARFHMHDWIIILVLLVTECCVLLIHPFFRYVGRDMMDDL
PAP1 : -----MWRTEQLGSHTVRSHGTVARTHMHDWIIILVLLVTECCVLLIHPFFRYVGRDMMDDL
PAP2 : -----MAWKRLVSCFPLQNFPGFQQRRMREVDLGAHTIRSHGTVARFHMHDWIIILVLLVTECCVLLIHPFFRYVGRDMMDDL
PAP4 : -----MDALVLSLVELGRMPEVWGLSHTVRSHGTVARITHCDDIIILVLLVTECCVLLIHPFFRYVGRDMMDDL
PAP3 : -----MPETHLGAHTVRSHGTVARTHMHDWIIILVLLVTECCVLLIHPFFRYVGRDMMDDL
TmD 1
PAPm1
At3g02600 : SYPLKSNVPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
At2g01180 : KYPFKDNVPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
At3g18220 : TRFFEYEDTPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
At1g15080 : RYPLQDNVPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
PAP1 : RYPMKSNVPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
PAP2 : KYPFKDNVPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
PAP4 : TYPLKDNVPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
PAP3 : SYPLKSNVPELMSVPVYAMLLPLVIEPIFVIRRRRDVYDHHAVLGLLYSVLVAVITDAIKNAVGRPRPDDFFWRCPF PDGKALYLSLG-DVICHGDKSIVREGHKSEF
TmD 2 TmD 3
PAPm2
At3g02600 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
At2g01180 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
At3g18220 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
At1g15080 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
PAP1 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
PAP2 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
PAP4 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
PAP3 : SGHTSSEAGLFLSYLSGKIAPFDGQGHVAKLCIVLPLLEALVGI SRVDDYWHHWQDVFAGGLGATSTIAYLOEFPFPMHTEGGGEMAYFQVLEAARVQGA
TmD 4 TmD 5 TmD 6
At3g02600 : ANGAVQQPPPVQVNNGEED---GGFMGLHLVDNPTMRREEDVETGRG*-
At2g01180 : SSQNGDALRAMSLQMDSTSLNEMSGTSTAPR*-----
At3g18220 : TMTNRTGSRGMLGNDVEPGNSASSPHDRHRESTDSD*-----
At1g15080 : DSAGMNHLSVRQTELESVR*-----
PAP1 : ASSAINLLNSEIREAQVENEDDSNGFLGLHADASSFPMGKEDIERGR*
PAP2 : NREQTGSSLNVAQMDAQ*-----
PAP4 : GRRANSIRKRGSTDVEAAMPFDPEMETPASRS SQTSLD*-----
PAP3 : SSNNANCLNIQQSELSQVYIDSQHQSVISRDTSPILEGTESEGRH*-----

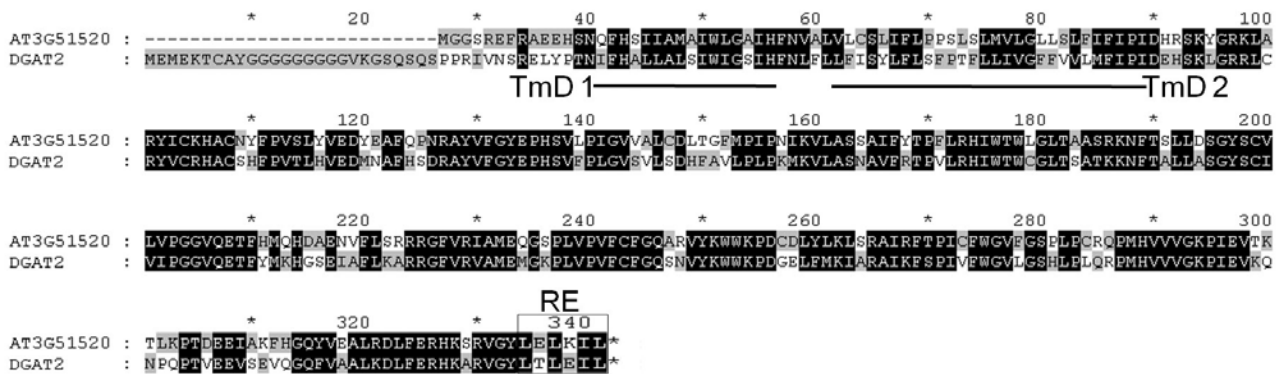
```

Supplementary Data 3 -Deduced amino acid sequence alignment of the members of the PAP family. *Arabidopsis* enzymes queries are designated by their respective database accession numbers. PAP(n): castor bean genes. Identical and similar amino acid residues are shaded black and gray, respectively. The transmembrane (TmD) regions are underlined. PAP motifs (PAPMs) are boxed. Black arrows indicate catalytically important residues.

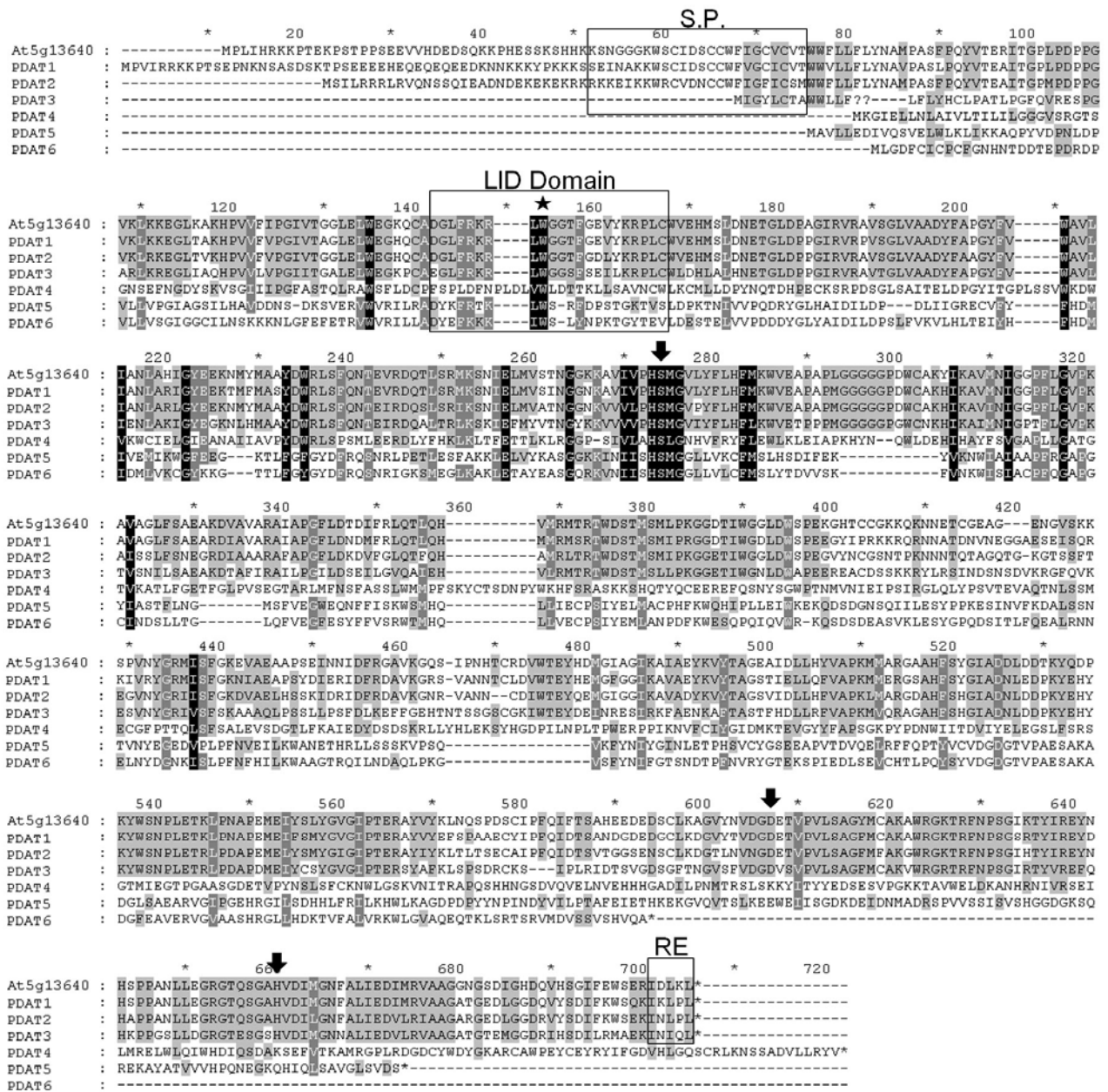
A.



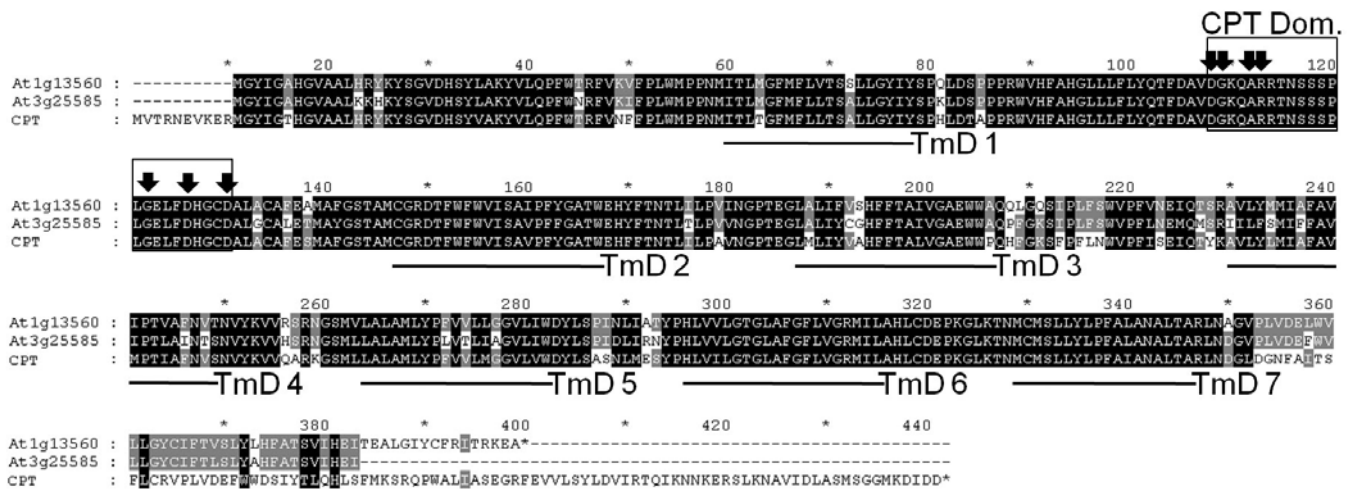
B.



Supplementary Data 4 -Deduced amino acid sequence alignment of the members of the DGAT family. *Arabidopsis* enzymes queries are designated by their respective database accession numbers. DGAT(n): castor bean genes. Identical and similar amino acid residues are shaded black and gray, respectively. The transmembrane (TmD) regions are underlined.



Supplementary Data 5 -Deduced amino acid sequence alignment of the members of the PDAT family. *Arabidopsis* enzymes queries are designated by their respective database accession numbers. PDAT(n): castor bean genes. Identical and similar amino acid residues are shaded in black and gray, respectively. RE, C-terminal aromatic amino acid-rich ER retrieval motif. SP represents the signal peptide that directs the protein to the ER. The lid domain is indicated by a black star, and conserved catalytic residues are marked with black arrows.



Supplementary Data 6 -Deduced amino acid sequence alignment of the members of the choline phosphotransferase family. *Arabidopsis* enzymes queries are designated by their respective database accession numbers. CPT: castor bean gene. Identical and similar amino acid residues are shaded black and gray, respectively. The transmembrane (TmD) regions are underlined. The CPT domain (CPT Dom.) is boxed. Black arrows indicate catalytically important residues.

Summary of TAG biosynthesis genes, primer sequences and size of amplification products

LIPID GENE ACTIVITY	ACRONYM	CASTOR BEAN ID	FORWARD PRIMER (5' 3') REVERSE PRIMER (5' 3')	AMPLICON Size (bp)
G3PAT	G3PAT1	30068.m002660	TCCTTTCATGGAGCTGGATT CATGGTTGCGACAATGAGAC	212
	G3PAT2	30174.m008615	CAAAGTTTTGGGAACGGAGA TCCTTGCAAATGGACATGAA	189
	G3PAT3	29969.m000267	CGACAGAATTGTGCCTGTTG CGATCGACGATTTACCACCT	182
	G3PAT4	29822.m003441	GCATAAAGGGAGAGCGACTG CCCGTCATGGAAAACATG	244
	G3PAT5	29736.m002070	CGCACTGATGCTCCCTTTAT AGGAGTGATCCTGCTGCAAT	224
	G3PAT6	27568.m000266	TCGGATTGCTACAGGTTTT CGGATTAGCTGGTGGTCATT	205
	G3PAT7	28350.m000105	CCGTGACCAGCTCTTCATCT ACAGCAAAGGAAGCCACAT	195
	G3PAT8	30076.m004618	TGTCGGAGTCATGGAAGAGA ATTTGTCTCTGGCAATCGT	200
	G3PAT9	29908.m005967	GGGTAGGGAGTTGATGGTGA CGCCAATTTCTTTCTCTGC	229
LPAAT	LPAAT1	29687.m000572	CCAGAGGGAACACGGAGTAA TGCTTCCTTCTACCGCTTA	198
	LPAAT2	27810.m000646	CGTTTGTCCAGCCATTTAT CACCATTGAGCAACAGCATC	169
	LPAAT3	30169.m006433	GGACGCAATGTTGGAGAAAT GACGATGCCGAAGATAAGGA	130
	LPAAT4	30169.m006432	ATTTCAAGCGAAAGGCACAT TACTGCATCAAGTGCCAACC	193
PAP	PAP1	29586.m000620	GCAATTCTTTCCTCCACCAT AATGTAGCCCCAAAAATCCA	183
	PAP2	29660.m000759	TTGTCGGGCAAAATTAAGC CATGATATGGAGCTGGGAAGA	211
	PAP3	29660.m000760	TGCAATTCTTTCACCTCCT CTCTGTTCCCTCCAGAATGG	212
	PAP4	29747.m001075	TTGCAGTCTTCCCTTATCCA GGATCGTGATGCAGGAGTTT	193
DGAT	DGAT1	29912.m005373	TCATGATGTGAGCTTTAAGACTTTG TGATAAATCCCATGAATCCTGT	158
	DGAT2	29682.m000581	TACCGCTACAACGTCCAATG TAGCCAACCCGTGCTTTATG	158
PDAT	PDAT1	29706.m001305	GGAGCGTGGTAGTGCTCATT TAGCACTCAGCAGCAGGAGA	203
	PDAT2	29912.m005286	AGCTTATGGCACGTGGAGAT GGAATAGCGCATTCACTGGT	216
	PDAT3	29991.m000626	CGGGGCTCATTCTCTCAT GGAGCGGAATACTCTTGCAC	210
	PDAT4	29637.m000766	TGGGTCTGACGTACAAGTGG TGCCATATCTGAAGCCACAA	225
	PDAT5	30060.m000520	GGCATGCCCACATTTAAGT GGGACATCCTCACCCATATA	180
	PDAT6	29929.m004538	AAATTCAGTTTGGCGCAAG GGCAGTTGTGCATCATTGAG	217
CPT	CPT	30138.m003845	TGTCTGCATCTAATTTAATGGAAAGT GTTTTCAATCCCTTGGTTCA	129

Supplementary Data 7 -Summary of TAG biosynthesis genes, primer sequences and sizes of amplification products.

5 CONSIDERAÇÕES FINAIS E PERSPECTIVAS

No presente trabalho buscamos contribuir para um melhor entendimento do processo de síntese de triacilgliceróis (TAGs) em sementes de mamona ao longo do desenvolvimento do endosperma desta oleaginosa.

Para tanto, 26 genes codificando enzimas envolvidas na síntese de TAGs foram identificados no genoma da mamona. Tais genes podem ser classificados como participantes das rotas dependente ou independente de acil-CoA, divisão esta, baseada no doador do último ácido graxo (acil-CoA ou fosfolipídios) que será incorporado ao diacilglicerol (DAG) originando TAG.

Na rota dependente de acil-CoA identificamos 9 genes para a enzima glicerol-3-fosfato aciltransferase, 4 para ácido lisofosfatídico aciltransferase, 4 para ácido fosfatídico fosfatase e 2 para diacilglicerol aciltransferase. Quanto à rota independente de acil-CoA identificamos 6 genes para fosfolipídio:diacilglicerol aciltransferase e 1 para colina fosfotransferase.

Análises de expressão por qRT-PCR demonstraram que alguns destes genes apresentam distintos padrões de expressão ao longo do desenvolvimento da semente. Estas diferenças também foram observadas para genes pertencentes a uma mesma classe de enzima e podem vir a indicar complementariedade funcional e temporal ao longo do processo de síntese de TAGs. Contudo, existe uma tendência para uma máxima expressão nos estágios medianos do desenvolvimento da semente (principalmente durante estágio 3). Tal tendência está em acordo com a fase de maior acúmulo de óleo em mamona, previamente descrita na literatura como sendo durante as etapas intermediárias de desenvolvimento do endosperma (Chen *et al.*, 2007).

Cabe ressaltar que, de acordo com a literatura vigente, existe uma preferência de utilização da rota independente de acil-CoA para a síntese de TAGs em espécies vegetais que apresentam ácidos graxos não usuais na composição de seus óleos, como é o caso da mamona. Este fato torna os genes identificados pertencentes a esta rota extremamente atrativos para possíveis estratégias biotecnológicas que visem produzir ácidos graxos não usuais em espécies agronomicamente importantes.

Atualmente, estamos em fase de ampliação de nossas pesquisas para genes envolvidos na biossíntese de TAGs presentes em outras espécies de oleaginosas: canola (*Brassica napus* L. var *oleifera*) e soja (*Glycine max* (L). Merr.).

Pela análise comparativa da organização desses genes, de suas histórias evolutivas, da expressão e da estrutura dos seus produtos gênicos pretende-se contribuir para a elucidação dos mecanismos que controlam o metabolismo de lipídios, em análise comparativa entre estas três espécies de oleaginosas.

A elucidação da atividade das enzimas responsáveis pelo controle do fluxo de ácidos graxos levando a formação de TAGs em oleaginosas poderá contribuir não apenas para o esclarecimento do processo do ponto de vista molecular como também para a identificação e possível manipulação da composição de ácidos graxos de interesse econômico, presentes em óleos vegetais de reserva.

Além disso, os resultados obtidos no presente trabalho somados aos que serão produzidos com a ampliação dos estudos para outras oleaginosas nos permitirão obter uma visão das especificidades dessa etapa do metabolismo lipídico em cada uma dessas plantas, o que poderá permitir o delineamento de

novas estratégias para o melhoramento das oleaginosas de interesse econômico para o Brasil.

6 REFERÊNCIAS

Abbadì, A., Domergue, F., Bauer, J., Napier, J.A., Welti, R., Zahringer, U., Cirpus, P. and Heinz, E. (2004) Biosynthesis of very-long-chain polyunsaturated fatty acids in transgenic oilseeds: constraints on their accumulation. *Plant Cell*, **16**, 2734-2748.

Agarwal, A.K. (2007) Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. *Prog. Energ. Combust.*, **33**, 233-271.

Amorim, P.Q.R. (2005) Perspectiva Histórica da cadeia da mamona e a introdução da produção de biodiesel no semi-árido brasileiro sob o enfoque da teoria dos custos de transação. *ESALQ – Escola Superior de Agricultura “Luiz de Queiroz”*, Piracicaba-SC: ESALQ-USP. **GRADUAÇÃO.**

Ascherio, A. (2006) Trans fatty acids and blood lipids. *Atheroscler Suppl*, **7**, 25-27.

Athenstaedt, K. and Daum, G. (2006) The life cycle of neutral lipids: synthesis, storage and degradation. *Cell Mol Life Sci*, **63**, 1355-1369.

Bafor, M., Smith, M.A., Jonsson, L., Stobart, K. and Stymne, S. (1991) Ricinoleic acid biosynthesis and triacylglycerol assembly in microsomal preparations from developing castor-bean (*Ricinus communis*) endosperm. *Biochem J*, **280 (Pt 2)**, 507-514.

Banas, A., Dahlqvist, A., Stahl, U., Lenman, M. and Stymne, S. (2000) The involvement of phospholipid:diacylglycerol acyltransferases in triacylglycerol production.

Biochem Soc Trans, **28**, 703-705.

Barreto, A.N. (2004

) Quantificação de água necessária para a mamoneira irrigada com base nas constantes hídricas do solo. *I Congresso Brasileiro da Mamona*, Campina Grande-PB.

Billheimer, J.T., Cromley, D.A. and Kempner, E.S. (1990) The functional size of acyl-coenzyme A (CoA):cholesterol acyltransferase and acyl-CoA hydrolase as determined by radiation inactivation. *J Biol Chem*, **265**, 8632-8635.

Biodiesel, E. (2008). from www.biodieseleco.org.br.

Bourgis, F., Kader, J.C., Barret, P., Renard, M., Robinson, D., Robinson, C., Delseny, M. and Roscoe, T.J. (1999) A plastidial lysophosphatidic acid acyltransferase from oilseed rape. *Plant Physiol*, **120**, 913-922.

Bouvier-Nave, P., Benveniste, P., Oelkers, P., Sturley, S.L. and Schaller, H. (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *Eur J Biochem*, **267**, 85-96.

Bowden, L. and Lord, J.M. (1975) Development of phospholipid synthesizing enzymes in castor bean endosperm. *FEBS Lett*, **49**, 369-371.

Broun, P. and Somerville, C. (1997) Accumulation of ricinoleic, lesquerolic, and densipolic acids in seeds of transgenic *Arabidopsis* plants that express a fatty acyl

hydroxylase cDNA from castor bean. *Plant Physiol*, **113**, 933-942.

Burgal, J., Shockey, J., Lu, C., Dyer, J., Larson, T., Graham, I. and Browse, J. (2008) Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnol J*, **6**, 819-831.

Cahoon, E.B., Dietrich, C.R., Meyer, K., Damude, H.G., Dyer, J.M. and Kinney, A.J. (2006) Conjugated fatty acids accumulate to high levels in phospholipids of metabolically engineered soybean and Arabidopsis seeds. *Phytochemistry*, **67**, 1166-1176.

Cahoon, E.B., Shockey, J.M., Dietrich, C.R., Gidda, S.K., Mullen, R.T. and Dyer, J.M. (2007) Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Curr Opin Plant Biol*, **10**, 236-244.

Cases, S., Smith, S.J., Zheng, Y.W., Myers, H.M., Lear, S.R., Sande, E., Novak, S., Collins, C., Welch, C.B., Lusic, A.J., Erickson, S.K. and Farese, R.V., Jr. (1998) Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. *Proc Natl Acad Sci U S A*, **95**, 13018-13023.

Cases, S., Stone, S.J., Zhou, P., Yen, E., Tow, B., Lardizabal, K.D., Voelker, T. and Farese, R.V., Jr. (2001) Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. *J Biol Chem*, **276**, 38870-38876.

Caupin, H. (1997) Products from castor oil: past, present, and future. *Lipid Technologies and Applications*, pp. 787-795.

Chen, G.Q., Turner, C., He, X., Nguyen, T., McKeon, T.A. and Laudencia-Chingcuanco, D. (2007) Expression profiles of genes involved in fatty acid and triacylglycerol synthesis in castor bean (*Ricinus communis* L.). *Lipids*, **42**, 263-274.

Chiaradia, J.J. (2005) Avaliação agrônômica e fluxo de gases do efeito estufa a partir de solo tratado com resíduos e cultivado com mamona (*Ricinus communis* L.) em área de reforma canavial. *Escola Superior de Agricultura Luiz de Queiroz*, Piracicaba: ESALQ/USP. **Tese de doutorado.**

Coelho, I. (1979) Avaliação das exportações tradicionais baianas: caso de sisal e mamona.

Dahlqvist, A., Stahl, U., Lenman, M., Banas, A., Lee, M., Sandager, L., Ronne, H. and Stymne, S. (2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc Natl Acad Sci U S A*, **97**, 6487-6492.

Daniel, J., Deb, C., Dubey, V.S., Sirakova, T.D., Abomoelak, B., Morbidoni, H.R. and Kolattukudy, P.E. (2004) Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in *Mycobacterium tuberculosis* as it goes into a dormancy-like state in culture. *J Bacteriol*, **186**, 5017-5030.

Dewey, R., Wilson, R., Novitzky, W. and Goode, J. (1994) The AAPT1 gene of soybean complements a cholinephosphotransferase-deficient mutant of yeast. *Plant Cell* **6**, 1495-1507.

Domergue, F., Chevalier, S., Creach, A., Cassagne, C. and Lessire, R. (2000) Purification of the acyl-CoA elongase complex from developing rapeseed and characterization of the 3-ketoacyl-CoA synthase and the 3-hydroxyacyl-CoA dehydratase. *Lipids*, **35**, 487-494.

Durrett, T.P., Benning, C. and Ohlrogge, J. (2008) Plant triacylglycerols as feedstocks for the production of biofuels. *Plant J*, **54**, 593-607.

Dyer, J.M. and Mullen, R.T. (2008) Engineering plant oils as high-value industrial feedstocks for biorefining: the need for underpinning cell biology research. *Physiol Plant*, **132**, 11-22.

EMBRAPA (2004) Informações sobre os Sistemas de Produção Utilizados na Ricinocultura na Região Nordeste, em Especial o Semi-Árido e outros aspectos ligados a sua cadeia, Campina Grande, Paraíba.

Franca, M.G., Matos, A.R., D'Arcy-Lameta, A., Passaquet, C., Lichtle, C., Zuily-Fodil, Y. and Pham-Thi, A.T. (2008) Cloning and characterization of drought-stimulated phosphatidic acid phosphatase genes from *Vigna unguiculata*. *Plant Physiol Biochem*, **46**, 1093-1100.

Ghosal, A., Banas, A., Stahl, U., Dahlqvist, A., Lindqvist, Y. and Stymne, S. (2007) *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyl transferase (PDAT) devoid of its membrane anchor region is a soluble and active enzyme retaining its substrate

specificities. *Biochim Biophys Acta*, **1771**, 1457-1463.

Goode, J. and Dewey, R. (1999) Characterization of aminoalcoholphosphotransferases from *Arabidopsis thaliana* and soybean. *Plant Physiol*, **37**, 445-457.

Graham, I.A. (2008) Seed storage oil mobilization. *Annu Rev Plant Biol*, **59**, 115-142.

Gunstone, F.D. (1987) Fatty acids and glycerides. *Nat Prod Rep*, **4**, 95-112.

Halford, N.G. and Hardie, D.G. (1998) SNF1-related protein kinases: global regulators of carbon metabolism in plants? *Plant Mol Biol*, **37**, 735-748.

He, X.H., Turner, C., Chen, G.Q., Lin, J.T. and McKeon, T.A. (2004) Cloning and characterization of a cDNA encoding diacylglycerol acyltransferase from castor bean. *Lipids*, **39**, 311-318.

Henriques, R.M. (2009) Potencial para Geração de Energia Elétrica no Brasil com Resíduos de biomassa através da Gaseificação. *UFRJ/COPPE*,
, Rio de Janeiro: UFRJ/COPPE,
. **Doctorate**, 2007.

Hobbs, D.H., Lu, C. and Hills, M.J. (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. *FEBS Lett*, **452**, 145-149.

Huang, A.H. (1996) Oleosins and oil bodies in seeds and other organs. *Plant Physiol*, **110**, 1055-1061.

IAC (2008). <http://www.iac.sp.gov.br/tecnologias/mamona/mamona.htm> Retrieved julho de 2008, 2008.

Jaworski, J. and Cahoon, E.B. (2003) Industrial oils from transgenic plants. *Curr Opin Plant Biol*, **6**, 178-184.

Katavic, V., Reed, D.W., Taylor, D.C., Giblin, E.M., Barton, D.L., Zou, J., Mackenzie, S.L., Covello, P.S. and Kunst, L. (1995) Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferase activity. *Plant Physiol*, **108**, 399-409.

Kennedy, E.P. (1961) Biosynthesis of complex lipids. *Fed Proc*, **20**, 934-940.

Kim, H.U., Li, Y. and Huang, A.H. (2005) Ubiquitous and endoplasmic reticulum-located lysophosphatidyl acyltransferase, LPAT2, is essential for female but not male gametophyte development in *Arabidopsis*. *Plant Cell*, **17**, 1073-1089.

Kinney, A.J. (1996) Designer oils for better nutrition. *Nat Biotechnol*, **14**, 946.

Kocsis, M.G. and Weselake, R.J. (1996) Phosphatidate phosphatases of mammals, yeast, and higher plants. *Lipids*, **31**, 785-802.

Kouri, J., Santos, R.F. and Santos, J.W. (2004) Evolução da Cultura da Mamona no Brasil. *CONGRESSO BRASILEIRO DE MAMONA*, Campina Grande-PB: EMBRAPA ALGODÃO.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, **305**, 567-580.

Kroon, J.T., Wei, W., Simon, W.J. and Slabas, A.R. (2006) Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. *Phytochemistry*, **67**, 2541-2549.

Lardizabal, K.D., Mai, J.T., Wagner, N.W., Wyrick, A., Voelker, T. and Hawkins, D.J. (2001) DGAT2 is a new diacylglycerol acyltransferase gene family: purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. *J Biol Chem*, **276**, 38862-38869.

Lewin, T.M., Wang, P. and Coleman, R.A. (1999) Analysis of amino acid motifs diagnostic for the sn-glycerol-3-phosphate acyltransferase reaction. *Biochemistry*, **38**, 5764-5771.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(-Delta Delta C(T)) Method. *Methods*, **25**, 402-408.

Lung, S.C. and Weselake, R.J. (2006) Diacylglycerol acyltransferase: a key mediator of plant triacylglycerol synthesis. *Lipids*, **41**, 1073-1088.

Martinelle, M., Holmquist, M., Clausen, I.G., Patkar, S., Svendsen, A. and Hult, K. (1996) The role of Glu87 and Trp89 in the lid of *Humicola lanuginosa* lipase. *Protein Eng*, **9**, 519-524.

McCartney, A.W., Dyer, J.M., Dhanoa, P.K., Kim, P.K., Andrews, D.W., McNew, J.A. and Mullen, R.T. (2004) Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. *Plant J*, **37**, 156-173.

McMaster, C.R. and Bell, R.M. (1997) CDP-choline:1,2-diacylglycerol cholinephosphotransferase. *Biochim Biophys Acta*, **1348**, 100-110.

Metzger, J.O. and Bornscheuer, U. (2006) Lipids as renewable resources: current state of chemical and biotechnological conversion and diversification. *Appl Microbiol Biotechnol*, **71**, 13-22.

Mhaske, V., Beldjilali, K., Ohlrogge, J. and Pollard, M. (2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid: diacylglycerol transacylase gene (At5g13640). *Plant Physiol Biochem*, **43**, 413-417.

Nakamura, Y., Tsuchiya, M. and Ohta, H. (2007) Plastidic phosphatidic acid phosphatases identified in a distinct subfamily of lipid phosphate phosphatases with

prokaryotic origin. *J Biol Chem*, **282**, 29013-29021.

Napier, J.A. (2007) The production of unusual fatty acids in transgenic plants. *Annu Rev Plant Biol*, **58**, 295-319.

Peelman, F., Vinaimont, N., Verhee, A., Vanloo, B., Verschelde, J.L., Labeur, C., Seguret-Mace, S., Duverger, N., Hutchinson, G., Vandekerckhove, J., Tavernier, J. and Rosseneu, M. (1998) A proposed architecture for lecithin cholesterol acyl transferase (LCAT): identification of the catalytic triad and molecular modeling. *Protein Sci*, **7**, 587-599.

Peelman, F., Verschelde, J.L., Vanloo, B., Ampe, C., Labeur, C., Tavernier, J., Vandekerckhove, J. and Rosseneu, M. (1999) Effects of natural mutations in lecithin:cholesterol acyltransferase on the enzyme structure and activity. *J Lipid Res*, **40**, 59-69.

Pierrugues, O., Brutesco, C., Oshiro, J., Gouy, M., Deveaux, Y., Carman, G.M., Thuriaux, P. and Kazmaier, M. (2001) Lipid phosphate phosphatases in Arabidopsis. Regulation of the AtLPP1 gene in response to stress. *J Biol Chem*, **276**, 20300-20308.

Qi, Q., Huang, Y.F., Cutler, A.J., Abrams, S.R. and Taylor, D.C. (2003) Molecular and biochemical characterization of an aminoalcoholphosphotransferase (AAPT1) from *Brassica napus*: effects of low temperature and abscisic acid treatments on AAPT expression in Arabidopsis plants and effects of over-expression of BnAAPT1 in transgenic Arabidopsis. *Planta*, **217**, 547-558.

Rojas-Barros, P., Haro, A., Munoz, J. and Fernandez-Martinez (2004) Isolation of a natural mutant in castor with high oleic/low ricoleic acid content in the oil. *Crop Sci.*, **44**, 76-80.

Saha, S., Enugutti, B., Rajakumari, S. and Rajasekharan, R. (2006) Cytosolic triacylglycerol biosynthetic pathway in oilseeds. Molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *Plant Physiol*, **141**, 1533-1543.

Santos-Filho, S.V., B., L., Oliveira, M. and Santos, M.A. (2004) Absorção de nutrientes durante a fase vegetativa mamoneira em três solos da região de Mossoró, RN sob diferentes níveis de salinidade da água de irrigação. *I Congresso Brasileiro da Mamona*, Campina Grande-PB.

Savy-Filho, A. (2005) Mamona tecnologia agrícola. *EMOPI*, Campinas-SP.

Scheller, J. and Conrad, U. (2005) Plant-based material, protein and biodegradable plastic. *Curr Opin Plant Biol*, **8**, 188-196.

Schrag, J.D. and Cygler, M. (1997) Lipases and alpha/beta hydrolase fold. *Methods Enzymol*, **284**, 85-107.

Shockey, J.M., Gidda, S.K., Chapital, D.C., Kuan, J.C., Dhanoa, P.K., Bland, J.M., Rothstein, S.J., Mullen, R.T. and Dyer, J.M. (2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different

subdomains of the endoplasmic reticulum. *Plant Cell*, **18**, 2294-2313.

Siloto, R.M., Truksa, M., Brownfield, D., Good, A.G. and Weselake, R.J. (2009) Directed evolution of acyl-CoA:diacylglycerol acyltransferase: Development and characterization of *Brassica napus* DGAT1 mutagenized libraries. *Plant Physiol Biochem*, **47**, 456-461.

Singh, S., Thomaus, S., Lee, M., Stymne, S. and Green, A. (2001) Transgenic expression of a delta 12-epoxygenase gene in *Arabidopsis* seeds inhibits accumulation of linoleic acid. *Planta*, **212**, 872-879.

Slack, C., Roughan, P., Browse, J. and Gardiner, S. (1985) Some properties of cholinephosphotransferase from developing safflower cotyledons. *Biochim Biophys Acta*, **833**.

Smith, M.A., Moon, H., Chowrira, G. and Kunst, L. (2003) Heterologous expression of a fatty acid hydroxylase gene in developing seeds of *Arabidopsis thaliana*. *Planta*, **217**, 507-516.

Somerville, C. and Browse, J. (1991) Plant Lipids: Metabolism, Mutants, and Membranes. *Science*, **252**, 80-87.

Somerville C, Browse J, Jaworski JG and JB, O. (2000) Lipids: chapter 10 In BB Buchanan, W Gruissem, RL Jones, eds, *Biochemistry and Molecular Biology of Plants. American Society of Plant Physiologists*, 456–527.

Stahl, U., Banas, A. and Stymne, S. (1995) Plant Microsomal Phospholipid Acyl Hydrolases Have Selectivities for Uncommon Fatty Acids. *Plant Physiol*, **107**, 953-962.

Stahl, U., Carlsson, A.S., Lenman, M., Dahlqvist, A., Huang, B., Banas, W., Banas, A. and Stymne, S. (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. *Plant Physiol*, **135**, 1324-1335.

Stobart, A.K. and Stymne, S. (1985) The interconversion of diacylglycerol and phosphatidylcholine during triacylglycerol production in microsomal preparations of developing cotyledons of safflower (*Carthamus tinctorius* L.). *Biochem J*, **232**, 217-221.

Stukey, J. and Carman, G.M. (1997) Identification of a novel phosphatase sequence motif. *Protein Sci*, **6**, 469-472.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*, **24**, 1596-1599.

van de Loo, F.J., Broun, P., Turner, S. and Somerville, C. (1995) An oleate 12-hydroxylase from *Ricinus communis* L. is a fatty acyl desaturase homolog. *Proc Natl Acad Sci U S A*, **92**, 6743-6747.

Vasudevan, P.T. and Briggs, M. (2008) Biodiesel production--current state of the art and

challenges. *J Ind Microbiol Biotechnol*, **35**, 421-430.

Vital, M.S., Milani, M., Meneses, C.H.S.G. and Bezerra, C. (2005) Comparação entre Protocolos para Extração de DNA Total de *Ricinus communis* L. *EMBRAPA: Comunicado Técnico*.

Voelker, T. and Kinney, A.J. (2001) Variations in the Biosynthesis of Seed-Storage Lipids. *Annu Rev Plant Physiol Plant Mol Biol*, **52**, 335-361.

Weier, D., Lühs, W.J., Dettendorfer, J. and Frentzen, M. (1998) sn-1-Acylglycerol-3-phosphate acyltransferase of *Escherichia coli* causes insertion of cis-11 eicosenoic acid into the sn-2 position of transgenic rapeseed oil. *Molecular Breeding*, **4**, 39-46.

Weselake, R.J., Taylor, D.C., Rahman, M.H., Shah, S., Laroche, A., McVetty, P.B. and Harwood, J.L. (2009) Increasing the flow of carbon into seed oil. *Biotechnol Adv.*

Williams, J.G. and McMaster, C.R. (1998) Scanning alanine mutagenesis of the CDP-alcohol phosphotransferase motif of *Saccharomyces cerevisiae* cholinephosphotransferase. *J Biol Chem*, **273**, 13482-13487.

Xu, C., Yu, B., Cornish, A.J., Froehlich, J.E. and Benning, C. (2006) Phosphatidylglycerol biosynthesis in chloroplasts of *Arabidopsis* mutants deficient in acyl-ACP glycerol-3- phosphate acyltransferase. *Plant J*, **47**, 296-309.

Xu, J., Francis, T., Mietkiewska, E., Giblin, E.M., Barton, D.L., Zhang, Y., Zhang, M.

and Taylor, D.C. (2008) Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (DGAT1) gene from *Tropaeolum majus*, and a study of the functional motifs of the DGAT protein using site-directed mutagenesis to modify enzyme activity and oil content. *Plant Biotechnol J*, **6**, 799-818.

Yu, B., Wakao, S., Fan, J. and Benning, C. (2004) Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in *Arabidopsis*. *Plant Cell Physiol*, **45**, 503-510.

Zhang, Q.X., Pilquill, C.S., Dewald, J., Berthiaume, L.G. and Brindley, D.N. (2000) Identification of structurally important domains of lipid phosphate phosphatase-1: implications for its sites of action. *Biochem J*, **345 Pt 2**, 181-184.

Zheng, P., Allen, W.B., Roesler, K., Williams, M.E., Zhang, S., Li, J., Glassman, K., Ranch, J., Nubel, D., Solawetz, W., Bhatramakki, D., Llaca, V., Deschamps, S., Zhong, G.Y., Tarczynski, M.C. and Shen, B. (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat Genet*, **40**, 367-372.

Zheng, Z., Xia, Q., Dauk, M., Shen, W., Selvaraj, G. and Zou, J. (2003) *Arabidopsis* AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell*, **15**, 1872-1887.

Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G. and Taylor, D.C. (1999) The *Arabidopsis thaliana* TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene.

Plant J, **19**, 645-653.

ANEXO I



www.journalexperts.com

American Journal Experts Editorial Certification

This document certifies that the manuscript titled "Identification, phylogenetic relationship and expression analysis of castor bean (*Ricinus communis*) genes encoding enzymes from triacylglycerol biosynthesis pathway." was edited for proper English language, grammar, punctuation, spelling, and overall style by one or more of the highly qualified native English speaking editors at American Journal Experts. Neither the research content nor the authors' intentions were altered in any way during the editing process.

Documents receiving this certification should be English-ready for publication - however, the author has the ability to accept or reject our suggestions and changes. To verify the final AJE edited version, please visit our verification page. If you have any questions or concerns over this edited document, please contact American Journal Experts at support@journalexperts.com

Manuscript title: Identification, phylogenetic relationship and expression analysis of castor bean (*Ricinus communis*) genes encoding enzymes from triacylglycerol biosynthesis pathway.

Authors: Alexandro Cagliari , Guilherme Loss, Matheus Etges, Rogério Margis and Márcia Margis-Pinheiro

Key: 2706-AA62-D1B5-00EF-00C4

This certificate may be verified at www.journalexperts.com/certificate

American Journal Experts is an association of Ph.Ds and Ph.D. graduate students from America's top 10 research universities. Our editors come from nearly every research field and possess the highest qualifications to edit research manuscripts written by non-native English speakers. We provide the quickest turnaround times at the lowest prices in the industry. For more information, please visit www.journalexperts.com, or for volume discounts for academic journals, please contact us by email at sales@journalexperts.com