UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

Clonagem, caracterização e expressão de genes envolvidos na síntese de compostos isoprenóides em *Eucalyptus grandis*

GENARO AZAMBUJA ATHAYDES

Dissertação submetida ao Programa de Pós-graduação em Biologia Celular e Molecular do Centro de Biotecnologia da UFRGS como requisito parcial para obtenção do grau de Mestre.

Orientador: Giancarlo Pasquali

Porto Alegre
Abril de 2010

Instituição e Fontes Financiadoras

O presente trabalho foi desenvolvido no Laboratório de Biologia Molecular Vegetal (LBMV) do Centro de Biotecnologia (CBiot) da Universidade Federal do Rio Grande do Sul (UFRGS).

Foram fontes financiadoras deste trabalho: Bolsa de Mestrado da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Ministério da Educação); Financiadora de Estudos e Projetos (FINEP) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Ministério da Ciência e Tecnologia) junto aos Projetos "Genolyptus – Rede Brasileira de Pesquisa do Genoma do *Eucalyptus*" e "WoodGenes - Gênese de Madeira em *Eucalyptus*: Genes, Funções, Regulação e Expressão Transgênica" por intermédio dos Fundos Setoriais Verde-Amarelo e de Biotecnologia do Ministério da Ciência e Tecnologia (MCT) e as empresas de celulose e papel Aracruz Celulose S.A., Grupo Raiz, Celulose Nipo-Brasileira S.A. – CENIBRA, International Paper do Brasil Ltda., Jarí Celulose S.A., Klabin S.A., Lwarcel Celulose e Papel Ltda., Veracel Celulose S.A., Votorantim Celulose e Papel S.A., Zanini Florestal S.A. e Suzano-Bahia Sul Papel.

Agradecimentos

Agradeço ao Prof. Giancarlo Pasquali por toda a atenção, orientação, ensinamentos e por ter sido o exemplo de professor que me inspirou a seguir minha Pós-Graduação na área de biologia molecular. Ao Prof. Jeverson Frazzon, agradeço pela confiança em meu potencial e pelas sugestões. Ao Prof. Hugo Verli, agradeço por ter disponibilizado um espaço para o desenvolvimento das análises de dinâmica molecular em seu laboratório e pelos ensinamentos nessa área.

Devo agradecer também às pessoas que, embora não estivessem comigo durante os dias de trabalho, estavam me apoiando sempre. Dentre essas pessoas, sou grato à minha esposa, Luana Cardoso, pela dedicação, amor, carinho que foram combustível indispensável para todos os meus dias. Aos meus pais, Adail e Lourdes Athaydes, que ainda não sabem em que eu trabalhava, mas que sempre me apoiaram de todas as formas em que estivera ao seu alcance. Ao meu irmão, Gabriel Athaydes pelo companheirismo de todos os dias.

Agradeço, ainda, aos meus queridos colegas do Laboratório de Biologia Molecular Vegetal (LBMV) que ajudaram a tornar a trajetória do Mestrado menos árdua e mais feliz. Especificamente, sou grato à Rochele Kirch, a primeira pessoa com quem conversei quando cheguei ao laboratório, o que não é de se espantar, e por ter se tornado uma amiga excepcional; à Michele Clàire, agradeço por ter me mostrado onde ficavam os meios de cultura e reagentes e pelas horas de descontração; à Luisa Abruzzi, pelas sugestões e incentivo em todos os momentos e pelo exemplo de caráter; à Marta Reis, por compartilhar seu modo perspicaz de avaliar as situações e de ver o mundo; ao Gustavo Riboldi, pelos conselhos nas análises de dinâmica molecular e exemplo de seriedade, ou não; ao Leonardo Pedrazza, pelos momentos recreacionais que nunca mais foram os mesmos depois que deixou o LBMV; ao Eduardo Mattos pelo exemplo de dedicação; ao Felipe Maraschin, pelas sugestões e fortes opiniões; à Juliana Borges, por sua agradabilíssima insolência; à Camila, por seus questionamentos que fazem todos pensarem no que estão fazendo. Sou grato ao Edson Vargas, pelo suporte e tempo dispensado nas análises de dinâmica molecular. De todos, levo um pouco do que mencionei, por isso farão parte de minha pessoa e os levarei sempre comigo em minhas ações. Por fim, agradeço às instituições financiadoras que tornaram o desenvolvimento desse trabalho possível.

Sumário

Lista de Abreviaturas, Símbolos e Unidades

µg – micrograma(s)

µL – microlitro(s)

Å – Angstron(s)

ADP – adenosina difosfato

Asp – aspartato

ATP – adenosina trifosfato

BLAST – *Basic Local Alignment Search Tool*

cDNA – DNA complementar

CDP-ME – 4-(difosfocitidil)-2-metil-D-eritritol

CDP-ME sintase – 4-(difosfocitidil)-2-metil-D-eritritol sintase

CDP-ME2P – 4-(difosfocitidil)-2-metil-D-eritritol-2-fosfato

cm – centímetro(s)

CMK – CDP-ME quinase

CMP – citidil monofosfato

CMS – CDP-ME sintase

CoA – Coenzima A

CTP – citosina trifosfato

DMAPP – dimetilalil difosfato

dNTPs – desoxirribonucleosídeo(s) trifosfato

DXP – 1-desoxixilulose-5-fosfato

DXR – 1-desoxixilulose-5-fosfato redutoisomerase

DXS – 1-desoxixilulose-5- fosfato sintase

EC – *enzyme commission*

EgIPPI1 – isopentenil difosfato isomerase 1 de *Eucalyptus grandis*

EgIPPI2 – isopentenil difosfato isomerase 2 de *Eucalyptus grandis*

ESTs – etiqueta(s) de sequência expressa

FMN – mononucleotídeo de flavina

FPP – farnesil difosfato

fs – fentosegundo(s)

g – grama(s)

GGPP – geranilgeranil difosfato

Glu – glutamato

Gly – glicina

GPP – geranil difosfato

GST – glutationa S-transferase

h – hora(s)

HDR – HMBPP redutase

HDS – HMBPP sintase

His – histidina

HMBPP – 1-hidroxi-2-C-metil-2-(E)-butenil-4-difosfato

HMG-CoA – éster de 3-hidroxi-3-metilglutaril coenzima A

HMGR – 3-hidroxi-3-metilglutaril reductase

IMP – isopentenil monofosfato

IPP – isopentenil difosfato

IPPI – isopentenil difosfato isomerase

IPPI1 – isopentenil difosfato isomerase 1

IPPI2 – isopentenil difosfato isomerase 2

IPTG – isopropil β-D-1-tiogalactosídeo

K – Kelvin

KEEG – *The Kyoto Encyclopedia of Genes and Genomes*

LB – meio Luria-Broth

MCS – MECP sintase

MDC – difosfomevalonato descarboxilase

ME-cPP – 2C-metil-D-eritritol-2,4-ciclodifosfato

MEP – 2-C-metil-D-eritritol-4-fosfato

Met – metionina

mg – miligrama(s)

min – minuto(s)

mL – mililitro(s)

mmol – milimol(ois)

mRNA – RNA Mitocondrial

MS – Murashige & Skoog

MVA – ácido mevalônico ou mevalonato

MVAP – mevalonato-5-fosfato

MVAPP – mevalonato-5-difosfato

NADH – nicotinamida adenina dinucleotídeo, forma reduzida

NADPH – nicotinamida adenina dinucleotídeo fosfato, forma reduzida

NCBI – *National Centre for Biotechnology Information*

ng – nanograma(s)

ns – nanosegundo(s)

ºC – graus Celsius

OD600 – densidade óptica a 600 nm

ORF – *open reading frame*

pmol – picomol(es)

ps – picosegundo(s)

PSI-BLAST – Position Specific Interactive BLAST

RNAse – ribonuclease

SDS – dodecil sulfato de sódio

SDS-PAGE – eletroforese em gel de poliacrilamida – SDS

TPP – tiamina difosfato

UV – radiação ultravioleta

v/v – volume por volume

ε – constante dielétrica

τ – constante de acoplamento

Lista de Figuras

Capítulo I Referencial Teórico

Capítulo II Cloning, characterization and expression of 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Eucalyptus grandis* (EgDXR).

Capítulo III Cloning, characterization, expression of two isopentenyl diphosphate isomerases from *Eucalyptus grandis* (EgIPPI1 and EgIPPI2), and molecular dynamics an IPPI1 protein.

Resumo

Os isoprenóides são compostos essenciais a todos os organismos e representam um grupo extremamente amplo estruturalmente e funcionalmente. Em plantas, mais de 30.000 compostos desta classe foram identificados até hoje. Todas as plantas produzem isoprenóides que desempenham papéis essenciais como os carotenóides, as clorofilas e as plastoquinonas (fotossíntese); ubiquinona (respiração); citocininas, brassinosteróides, giberelinas, ácido abscísico (regulação do crescimento e desenvolvimento). No entanto, a maior variedade de isoprenóides é representada pelos metabólitos secundários (óleos essenciais como eucaliptol, cineol, citronelal). Os isoprenóides possuem papel marcante nas relações da planta com o ambiente onde estão inseridas, já que medeiam as relações planta-inseto, planta-microorganismo e planta-planta.

Devido ao valor comercial dos compostos isoprenóides, há grande interesse em produzi-los por bioengenharia em bactérias ou em plantas e o entendimento do papel dos genes e proteínas codificadas na rota de síntese dessas unidades básicas de formação de terpenóides é extremamente importante. Nesse trabalho, nós recuperamos, a partir da seleção dos plasmídios de clonagem pSPORT1 (Invitrogen) purificados do Projeto "Genolyptus: Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*", os clones selecionados e potencialmente codificadores de 1-desoxi-D-xilulose-5-fosfato redutoisomerase (DXR) e de outras enzimas atuantes nas rotas de síntese de isoprenóides, como 1-desoxi-D-xilulose 5-fosfato sintase (DXS), mevalonato difosfato descarboxilase (MDC), isopentenil difosfato isomerase 1 (IPPI1) e isopentenil difosfato isomerase 2 (IPPI2).

Nos estoques de plasmídios do projeto Genolyptus, foi possível encontrar as sequências completas dos genes *dxr, ippi1* e *ippi2*. Os genes foram analisados *in silico* e a sequência de *ippi1* foi utilizada na modelagem molecular e dinâmica molecular para avaliação de características peculiares do *folding* protéico desse tipo de proteína eucariótica. Os genes foram clonados em vetores pGEX-4T (GE Healthcare) e heterologamente expressos em *Escherichia coli*. Foi realizada, também, uma análise transcricional comparativa dos genes selecionados pela técnica de microarranjos.

Abstract

Isoprenoids are essential to all organisms and are the most structurally and functionally diverse group of plant metabolites. In plants, more than 30,000 compounds of this class were identified to date. All plants produce isoprenoids that can play essential roles as carotenoids, chlorophylls and plastoquinone (photosynthesis); ubiquinone (respiration); regulation of growth and development (cytokinins, brassinosteroids, gibberellins, abscisic acid). However, the majority of isoprenoids is represented by secondary metabolites (essential oils like eucaliptol, cineol, citronelal). Isoprenoids have important roles in the relationships between plants and the environment, since they can mediate plant-insect, plant-microorganism and plant-plant interactions, as well as participate in abiotic stress responses.

Due to the high value of isoprenoid compounds, there is great interest in producing them by bioengineering in bacteria or plants. Understanding the role of genes and proteins related to the biosynthetic pathway of isoprenoids is extremely important. In this work, we retrieved, from the collection of cloning plasmids pSPORT1 (Invitrogen) generated in the "Genolyptus Project: The Brazilian Research Network on the *Eucalytpus* Genome", selected clones that potentially codified the 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) and other enzymes from the isoprenoid biosynthetic pathway including 1-deoxy-D-xylulose 5-phosphate synthase (DXS), mevalonate disphosphate decarboxilase (MDC), isopentenil disphosphate isomerase 1 (IPPI1) and isopentenil disphosphate isomerase 2 (IPPI2).

Complete sequences of the *dxr*, *ippi1* and *ippi2* genes were succesfully recovered from the Genolyptus stocks. The genes were analyzed *in silico* and the *ippi1* sequence was used in studies of molecular modeling and molecular dynamics in order to evaluate specific folding characteristics of this kind of eukaryotic IPPI. The genes were cloned into pGEX-4T vectors (GE Healthcare) and expressed in *Escherichia coli*. Also, the transcriptional analysis of the selected genes was performed by microarray analysis.

CAPÍTULO I

REFERENCIAL TEÓRICO

1.1 Projeto "Genolyptus – Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*"

Lançado oficialmente em 2002 pelo Ministério de Ciência e Tecnologia, com recursos do Fundo Setorial "Verde-Amarelo" e de empresas de papel e celulose, o Projeto "Genolyptus – Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*" contou com a participação de 14 empresas e dez centros de pesquisa. O principal objetivo do projeto foi a descoberta, o sequenciamento, o mapeamento e a determinação de genes de interesse econômico para diversas espécies de *Eucalyptus*, visando a incorporação de tecnologias de genômica em programas de melhoramento genético e produção florestal, buscando plantas mais resistentes a moléstias e a estresses bióticos e abióticos. Junto ao subprojeto "Sequenciamento do Transcriptoma de *Eucalyptus*" foram obtidas 96.493 marcas de sequências expressas (ESTs, do inglês, *Expressed Sequence Tags*) válidas (de Oliveira, 2007).

Em janeiro de 2005, teve início o projeto "*WoodGenes* - Gênese de Madeira em *Eucalyptus*: Genes, Funções, Regulação e Expressão Transgênica", o qual teve como objetivo explorar os clones de cDNA gerados no projeto Genolyptus, avaliar a função gênica das sequências isoladas e o potencial destas em proporcionar características de interesse econômico por meio da expressão transgênica. Junto a ambos os Projetos, foram executados e estão sendo analisados os resultados de experimentos de hibridização de microarranjos de DNA com sondas derivadas de 21.442 sequências únicas do transcriptoma de *Eucalyptus* (Bastolla, 2007; Jahns, 2007; Pasquali *et al.*, em preparação).

A partir dos resultados de anotação automática da potencial identidade funcional das sequências depositadas no banco de dados do Genolyptus (Laboratório de Genômica e Expressão Gênica, Universidade de Campinas, http://www.lge.ibi.unicamp.br/eucalyptus/; Laboratório de Bioinformática, Universidade Católica de Brasília, http://www.genoma.ucb.br/SistemaGenoma/), foram identificadas nesse estudo sequências únicas (*singlets*) de ESTs ou grupos de sequências (*clusters*) potencialmente codificadores de proteínas de *Eucalyptus* relacionadas à biossíntese dos isoprenóides. Após a reunião de todas estas sequências disponíveis no

banco de dados do Genolyptus, foram selecionados clones de cDNA potencialmente completos, isto é, caracterizados pela presença do códon de início de tradução, ATG. Considerando que os cDNAs foram gerados a partir de oligonucleotídeos com complementaridade às caudas de poli(A), subentende-se que todos os produtos de síntese, ou a sua maioria, possuem a região 3'-terminal e, por conseguinte, os códons de terminação da tradução. Finalmente, os clones de cDNA representantes de 1-desoxixilulose-5-fosfato redutoisomerase (DXR), isopentenil difosfato isomerase 1 (IPPI1) e 2 (IPPI2) foram investigados em maior profundidade, conforme descrito no presente estudo.

1.2 Metabolismo Secundário Vegetal

Estima-se que 15-25% dos genes das plantas são dedicados às rotas do metabolismo secundário vegetal (Pichersky e Gang, 2000). A maioria dos organismos é capaz de produzir compostos que não são responsáveis pela sobrevivência direta das células, mas desempenham um papel na interação do organismo com seu ambiente (Bennet e Bentley, 1989; Nugroho e Verpoorte, 2001). Essa definição é baseada em dois aspectos. Na natureza, as plantas produzem compostos para proteção contra predação por herbívoros, contra infecções por microrganismos ou para atração de insetos polinizadores via coloração e odor (Ersek e Kiraly, 1986). O segundo aspecto está relacionado com a restrição das rotas metabólicas secundárias a espécies e gêneros individuais e devem ser ativados apenas durante estágios específicos de desenvolvimento e de crescimento (Nugroho e Verpoorte, 2001).

As aproximadamente 400.000 espécies de plantas superiores existentes no mundo (Hostettmann e Terreaux, 2000) são fontes de numerosos metabólitos que podem ser utilizados para diferentes fins, como no desenvolvimento de fármacos, agroquímicos, flavorizantes, fragrâncias, biopesticidas e aditivos alimentares. Mais de 100.000 metabólitos foram identificados até hoje, o que representa apenas 10% do total atual estimado presente na natureza sendo que, desse, apenas metade das estruturas químicas foi completamente elucidada. A diversidade molecular é um fenômeno amplamente distribuído na natureza. Embora muitos metabólitos secundários de plantas sejam estruturalmente similares, eles podem possuir diferentes atividades biológicas. Essa enorme heterogeneidade de metabólitos secundários é, na maioria dos casos, derivada de modificações diferenciais de estruturas químicas primárias comuns através

de eventos pós-sintéticos como hidroxilação, glicosilação, metilação, acilação, prenilação, sulfação e benzoilação. Geralmente, a função de cada metabólito secundário é diferente. Os compostos terpenóides, representam um exemplo fascinante, já que estão presentes em todos os organismos e, em plantas, são especialmente abundantes com mais de 30.000 compostos descritos até hoje (Zhong e Yue, 2005).

O grande número e a diversidade de metabólitos secundários em plantas podem ser atribuídos, também, à ampla especificidade e aos múltiplos produtos de reações que são típicos das enzimas produtoras de metabólitos secundários. Essas reações enzimáticas aumentam a probabilidade de gerar diversidade química. A seleção e a retenção da diversidade química são fatores críticos para a adaptação dos organismos e é a razão primária para o vasto número de compostos químicos naturais (Sumner *et al.*, 2007). As plantas utilizam essa variedade química principalmente para a proteção contra desafios bióticos e abióticos (Dixon, 2001). Por isso, esses compostos são responsáveis por um aumento de vigor já que estão relacionados com atividades antimicrobianas, anti-herbivoria e atividade alelopática. Essas substâncias tóxicas combatem o potencial danoso de vírus, bactérias, fungos e herbívoros e/ou garantem uma vantagem competitiva frente às outras plantas (Sumner *et al.*, 2007). Apenas poucas plantas têm sido estudadas em detalhe para as diferentes rotas de metabolismo secundário. São exemplos, o tabaco (antocianinas/flavonóides, terpenóides e alcalóides), *Catharantus roseus* (alcalóides, esteróides, brassinolídeos, flavonóides e ácido 2,3-diidroxibenzóico) e *Cinchona sp* (antraquinonas e alcalóides; Verpoorte *et al.*, 2000).

As saponinas e as isoflavonas são exemplos de metabólitos que produzem respostas desfavoráveis a predadores de plantas, pois são, respectivamente, agente antinutricional e causador de infertilidade. Muitos produtos naturais também têm outras funções biológicas benéficas como garantir sabor, fragrância e cor atrativos, protetores contra radiação ultravioleta (UV), antioxidantes, compostos sinalizadores associados a interações ecológicas e modulações simbióticas (Sumner *et al.*, 2007).

A produção desses compostos nas plantas não é sempre satisfatória para extração com rendimento adequado e está restrita a algumas espécies e gêneros. Essa produção deve ser ativada apenas durante um estágio de desenvolvimento particular, sob determinadas condições sazonais, em situações de estresse ou em algumas condições de disponibilidade de nutrientes. Por isso, nos últimos anos, os pesquisadores têm se focado na regulação, direcionada pelas possibilidades da engenharia metabólica, da biossíntese dos compostos de interesse. Essa estratégia pode ser utilizada para aumentar

a produção de um composto desejado na própria planta, em plantas de espécies relacionadas ou até mesmo em microrganismos (Verpoorte *et al.*, 2000; Pasquali *et al.*, 2006).

Os maiores desafios associados ao desenvolvimento de hospedeiros microbianos para a biossíntese de isoprenóides incluem a engenharia de rotas de precursores dos isoprenóides além da descoberta e da expressão dos genes biossintéticos (Kirby *et al.*, 2009). Apesar dos enormes esforços comerciais através de várias décadas, apenas poucos compostos foram produzidos com custos menores do que aqueles necessários para produção pela extração direta das plantas ou por intermédio da síntese química. Como uma alternativa, as rotas metabólicas têm sido modificadas para aumentar a produtividade via transformação genética. Embora essa abordagem de engenharia pareça promissora, ela tem sido limitada pelo pouco conhecimento corrente do metabolismo secundário ao nível molecular (Liu, 2002).

1.2.1 Terpenos

Os terpenóides formam o grupo mais estrutural e funcionalmente diverso de todos os metabólitos secundários produzidos por todas as plantas e animais (Kirby *et al.*, 2009). Formados pela união de unidades isoprenóides pentacarbonadas (C5), os terpenos são classificados pelo número de unidades C5 em seu esqueleto básico (Fig. 1). São exemplos desses grupos os monoterpenos (C10), os sesquiterpenos (C15), os diterpenos (C20) e os triterpenos (C30; D´Auria *et al.*, 2005).

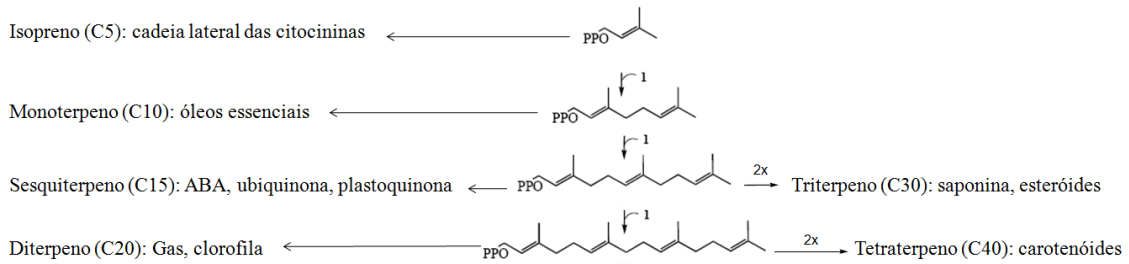


Figura 1. Síntese de várias classes de terpenóides em plantas. Adaptado de De Ruyck & Wouters (2008). 1, condensação de uma molécula de IPP; ABA, ácido abscísico; Gas, giberelinas.

As unidades isoprênicas podem ser unidas por ligações cabeça-cauda para formar cadeias lineares que podem ser arranjadas para formar anéis. Os terpenos/isoprenos podem existir como hidrocarbonetos ou possuir grupos funcionais oxigenados como hidroxila, carbonila, cetona ou aldeído. Após a modificação química dos terpenos, os compostos resultantes são referidos como terpenóides/isoprenóides (Paduch *et al.*, 2007).

Os terpenóides de plantas podem ser classificados como metabólitos primários, necessários para a manutenção da função celular, ou como metabólitos secundários, que não estão envolvidos seja com o crescimento seja com o desenvolvimento da planta. Os metabólitos primários incluem as giberelinas, os carotenóides e os esteróides que atuam na modulação do crescimento celular e elongação da planta, na captação de luz, na fotoproteção e no controle da fluidez e permeabilidade da membrana. Os metabólitos secundários terpenóides são, geralmente, atrativos comercialmente já que podem ser usados como flavorizantes, realçadores de cor, químicos utilizados na agricultura e medicamentos. Uma vasta gama de terpenóides tem se mostrado farmaceuticamente ativa contra doenças humanas como câncer (taxanos, incluindo paclitaxel proveniente de *Taxus* spp., alcalóides indol terpenóides, incluindo vincristina e vinblastina de *C. roseus*), malária (artemisina proveniente de *Artemisia annua*) e HIV (cumarinas, incluindo o calanolídeo A de *Calophyllum lanigerum*; Kirby *et al.*, 2009).

Como há um elevado interesse comercial nos compostos terpenóides, a aplicação de estratégias de engenharia metabólica é de grande importância social, já que pode tanto aumentar a produção de terpenóides de valor medicinal, quanto aumentar a resistência contra doenças e contra pragas ou, ainda, realçar a fragrância e aroma de frutos e de verduras.

1.2.1.1 Rotas Metabólicas Universais de Biossíntese de Isopentenil Difosfato e de Dimetilalil Difosfato

Apesar da enorme diversidade estrutural dos isoprenóides, suas rotas biossintéticas são organizadas de maneira regular e simples (Fig. 1). Todos os isoprenóides são sintetizados a partir das mesmas unidades básicas isoprênicas, isopentenil difosfato (IPP) e dimetilalil difosfato (DMAPP; Harada *et al.*, 2009).

Esses dois isômeros isoprênicos fosfatados, IPP e DMAPP, podem ser produzidos por duas rotas universais: a rota dependente de mevalonato (MVA) e a rota do metileritritol fosfato (MEP), também conhecida como "rota de Rhomer" ou, ainda, como a rota do 1-desoxi-D-xilulose-5-fosfato (DXP). A rota MVA é prevalente em eucariotos e em archaea, enquanto a rota MEP é amplamente distribuída em eubactérias. Entretanto, há muitas exceções a esse padrão de distribuição das rotas. Muitas eubactérias utilizam a rota MVA e algumas espécies possuem genes de ambas as rotas. As algas clorofitas, aparentemente, perderam a rota mevalônica e utilizam exclusivamente a rota MEP, bem como o protozoário, parasita humano causador da malária, *Plasmodium falciparum*, e outros apicomplexa sintetizam seus isoprenóides via rota MEP nos seus apicoplastos tipo-plastídios não-verdes (Kirby *et al.*, 2009).

Em plantas, ambas as rotas estão presentes, mas compartimentalizadas em locais diferentes dentro das células. As enzimas relacionadas com a rota MVA estão localizadas no citosol enquanto que as enzimas da rota MEP estão localizadas nos plastídios. A origem exata dos genes codificantes das enzimas atuantes na rota MEP não é clara, pois a maioria dos genes não ramifica junto de sua contraparte, a cianobactéria. Uma teoria propõe que esses genes, atualmente encontrados nas plantas, foram adquiridos após o surgimento dos plastídios e que a transferência lateral de genes da eubactéria tenha desempenhado um papel importante (Kirby *et al.*, 2009).

Já está bem estabelecido que, em cianobactérias e em plantas, IPP e DMAPP são sintetizados nos plastídios pela rota MEP e no citosol por meio da rota MVA (Bouvier *et al.*, 2005), como ilustrado na Fig. 2. Quantidades mínimas de metabólitos comuns a ambas as rotas podem ser trocadas nas duas direções pelas membranas plastídicas. Por isso, isoprenóides derivados de 1-desoxi-D-xilulose marcada podem ser desviados para o citoplasma onde podem se tornar parte de moléculas de esterol. Da mesma forma, uma pequena fração de grupamentos isoprenóides derivados da rota MVA consegue entrar

no compartimento plastídico e incorporar-se a mono e a diterpenos os quais são obtidos predominantemente via rota MEP (Eisenreich, 2004).

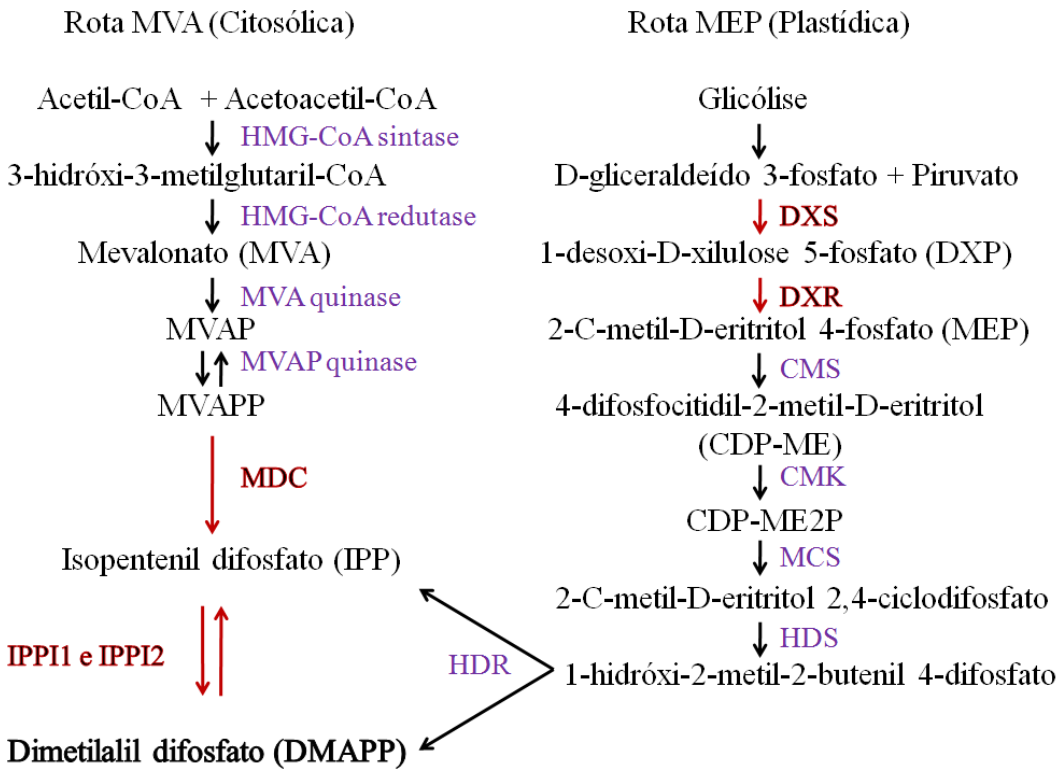
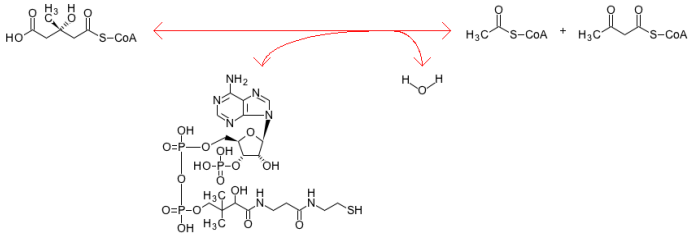
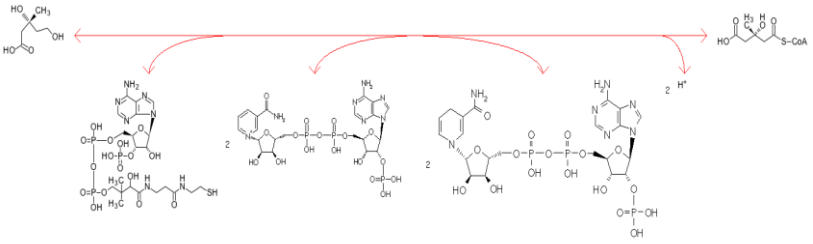
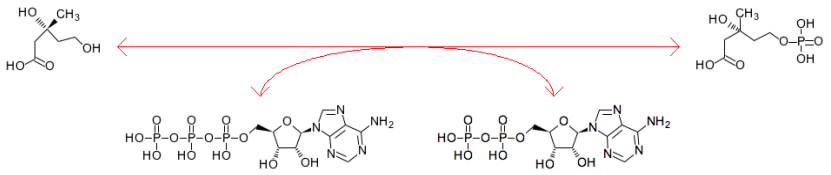
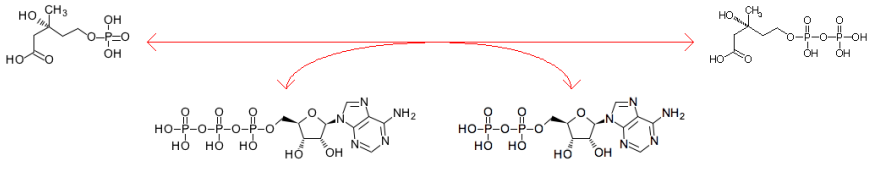
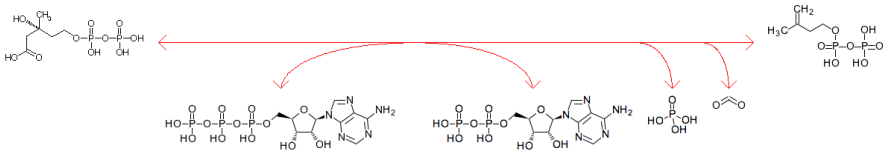


Figura 2. Rotas universais de biossíntese de IPP e de DMAPP. Em vermelho, as proteínas selecionadas neste trabalho para estudo detalhado. Em negrito, a forma isomerizada de IPP, DMAPP, capaz de condensar com IPP ou moléculas isoprenóides para formação de compostos mais complexos. MVA, ácido mevalônico; Val, valina; Leu, leucina; Ile, isoleucina; HMG-CoA sintase, 3-hidróxi-3-metilglutaril-coenzima A sintase; HMG-CoA redutase, 3-hidróxi-3-metilglutaril-coenzima A redutase; MVA quinase, mevalonato quinase; MVAP, mevalonato-5-fosfato; MVAP quinase, mevalonato-5-fosfato quinase; MVAPP, mevalonato-5-difosfato; MDC, mevalonato-5-difosfato descarboxilase; IPP, isopentenil difosfato; IPPI1, isopentenil difosfato isomerase 1; IPPI2, isopentenil difosfato isomerase 2; DMAPP, dimetilalil difosfato; MEP, 2C-metil-D-eritritol-4-fosfato; DXS, 1-desoxi-D-xilulose-5-fosfato sintase; DXP, 1-desoxi-D-xilulose-5-fosfato; DXR, 1-desoxi-D-xilulose-5-fosfato redutoisomerase; CMS, 4-difosfocitidil-2C-metil-D-eritritol sintase; CDP-ME, 4-difosfocitidil-2C-metil-D-eritritol;CMK, 4-difosfocitidil-2C-metil-D-eritritol quinase; CDP-ME2P, 4-difosfocitidil-2C-metil-D-eritritol-2-fosfato; MCS, 2C-metil-D-eritritol-2,4-ciclodifosfato sintase; HDS, 1-hidroxi-2C-metil-2-(E)-butenil-4-difosfato sintase; HDR, 1-hidroxi-2C-metil-2-(E)-butenil-4-difosfato redutoisomerase.

1.2.1.1.1 Rota citosólica dependente de mevalonato

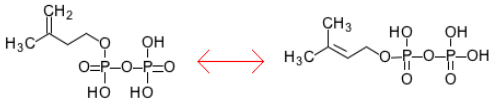
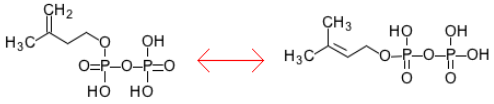
Por muito tempo, a rota dependente de MVA foi considerada a fonte universal dos precursores pentacarbonados, IPP e DMAPP. Em plantas superiores, a rota MVA está localizada no citoplasma e nas mitocôndrias sendo responsável pela formação de esteróis, triterpenos, ubiquinonas e sesquiterpenos (Fig. 3; Pang *et al*., 2006). As enzimas da rota MVA (Quadro 1) têm recebido muita atenção nos últimos 30 anos, e vários mecanismos regulatórios têm sido descobertos (D´Auria *et al.*, 2005). Nessa rota, ocorre a condensação de três unidades de acetil-CoA para formar o éster de 3-hidroxi-3-metilglutaril-CoA (HMG-CoA) pela atividade da enzima HMG-CoA sintase. Este produto é subsequentemente reduzido para produzir MVA pela ação da HMG-CoA redutase, uma enzima-chave reguladora dessa rota intensivamente estudada. Após, o MVA é fosforilado duas vezes em dois passos sequenciais. Primeiro, a enzima MVA quinase converte MVA a mevalonato-5-fosfato (MVAP) e, em seguida, a enzima MVAP quinase utiliza MVAP como substrato para produzir mevalonato-5-difosfato (MVAPP; Rodwel *et al*., 2000). O último passo da rota é catalisado pela MVAPP descarboxilase (MDC), convertendo MVAPP, uma molécula de 6 carbonos, em IPP, que contém 5 carbonos, às custas de uma molécula de adenosina trifosfato (ATP) e utilizando íons de magnésio (Mg^{+2}) como cofatores (Pang *et al*., 2006).

Quadro 1. Enzimas da rota MVA e suas respectivas reações de catálise. Disponível na base de dados da *Kyoto Encyclopedia of Genes and Genomes* (KEGG).

Enzima (EC)	Reação
HMG-CoA sintase (2.3.3.10)	 (S)-3-Hidroxi-3-metilglutaril-CoA + CoA <=> Acetil-CoA + H_2O + Acetoacetil-CoA
HMG-CoA redutase (1.1.1.34)	 (R)-Mevalonato + CoA + 2 $NADP^+$ <=> (S)-3-Hidroxi-3-metilglutaril-CoA + 2 NADPH + 2 H^+
MVA quinase (2.7.1.36)	 ATP + (R)-Mevalonato <=> ADP + (R)-5-fosfomevalonato
MVAP quinase (2.7.4.2)	 ATP + (R)-5-Fosfomevalonato <=> ADP + (R)-5-Difosfomevalonato
MVAPP descarboxilase, MDC (4.1.1.33)	 ATP + (R)-5-Difosfomevalonato <=> ADP + Ortofosfato + Isopentenil difosphato + CO_2

Um passo extremamente importante na biossíntese de isoprenóides é a isomerização do IPP para a produção de DMAPP. Esta conversão é catalisada pela IPP isomerase (IPPI; Quadro 2). A IPPI do tipo I foi primeiramente caracterizada em 1960 (Agranoff *et al.*, 1960). As IPPIs desse tipo são monômeros em sua forma ativa e são encontradas em eucariotos e bactérias. Esta enzima requer um cátion divalente e catalisa a isomerização de IPP a DMAPP por um mecanismo de protonação-desprotonação. As IPPIs do tipo II são proteínas homo-octaméricas (de Ruyck *et al.*, 2005) encontradas em bactérias e arqueobactérias e requerem a redução de uma flavina e um metal divalente como cofatores para sua atividade (Pang *et al.*, 2006). A ação das IPPIs mantém o equilíbrio entre IPP e DMAPP que tende a uma maior concentração celular do produto DMAPP. Um equivalente biológico das IPPIs, codificado pelo gene ortólogo *idi*, está presente em apenas algumas bactérias que utilizam a rota MEP (Cunningham *et al.*, 2000). Esse gene não é essencial em *E. coli* (Hahn *et al.*, 1999). Entretanto, as IPPIs podem desempenhar um papel mais importante em plantas do que em bactérias, já que o silenciamento de uma IPPI em *Nicotiniana benthamiana* resulta em um fenótipo com redução de 80% nos pigmentos (carotenóides e clorofila) comparado com controles (Page *et al.*, 2004).

Quadro 2. Enzimas IPPIs e suas respectivas reações de catálise. Disponível na base de dados da KEGG.

IPPIs, EC 5.3.3.2	Reação
Tipo 1 (IPPI1 e IPPI2)	 Isopentenil difosfato + NADPH <=> Dimetilalil difosfato + $NADP^+$ + H^+
Tipo 2	 Isopentenil difosfato + 2 NADPH + FMN <=> Dimetilalil difosfato + 2 $NADP^+$ + FMN

1.2.1.1.2 Rota do metileritritol fosfato

A rota MEP foi completamente elucidada apenas em 2002 e é bem menos compreendida em termos de requerimentos e regulação (Aharoni *et al.*, 2003). Rohmer *et al.* estudaram a incorporação de vários isotopômeros [$^{13}C_1$]glicose e isotopômeros [^{13}C]acetato em hopanóides de uma eubactéria e verificaram que a localização das moléculas marcadas não estava de acordo com o paradigma da existência de uma rota única de formação de IPP, a rota MVA. Um avanço decisivo para a elucidação da anteriormente desconhecida rota MEP foi a descoberta de que a 1-desoxi-D-xilulose é incorporada pelas células de *E. coli* nas cadeias laterais terpenóides da ubiquinona com uma eficácia extraordinária (Eisenreisch *et al.*, 2004).

Vários genes, enzimas e intermediários envolvidos na rota MEP foram primeiro descobertos em *E. coli* e extensivamente estudados, mas alguns genes ortólogos codificando várias enzimas da rota MEP também foram demonstrados em espécies de plantas (Tabela 1). A existência de uma rota MEP capaz de sintetizar uma variedade de terpenóides tem sido demonstrada em diversas espécies de plantas (Durbey *et al*., 2003).

Tabela 1. Enzimas da rota MEP clonadas e caracterizadas em várias espécies de plantas. Adaptado de Ganjewala *et al.* (2008).

Enzima	Planta	Código GB*	Tamanho**	Proteína/Gene	Referência
DXS (EC 4.1.3.37)	*Arabidopsis thaliana*	NP_001078570	565	DXPS3	Sato *et al.*, 2000
		NP_196699	700	DXPS3	
		NP_850620	629	DXPS1	
		BAB02345	604	DXS	
	Antirrhinum majus	AAW28999	733	DXPS	Dudareva *et al.*, 2005
	Capsicum annuum	CAA75778.1	719	*dxs* put.	Bouvier *et al.*, 1998
	Catharanthus roseus	CAA09804	716	DXS	Chahed *et al.*, 2000
	Chrysanthemum x morifoli	BAE79547	669	DXS	Kishimoto *et al.*, 2006
	Croton stellatopilosus	BAF75640	720	*dxs*	Wungsintaweekul *et al.*, 2008
	Elaeis guineensis	AAS99588	707	*dxs*	Khemvong *et al.*, 2005
	Ginkgo biloba	AAS89341	711		Gong *et al.*, 2006
	Hevea brasiliensis	AAS94123	720	DXS	Seetang-Nun *et al.*, 2008a
		ABF18929	711	DXS2	
	Lycopersicon esculentum	AAD38941	719	*dxs*	Lois *et al.*, 2000
	Morinda citrifolia	AAL32062	722	DXS	Han *et al.*, 2003
	Oryza sativa (Japonica)	NP_001055524	720	*dxs*	Ohyanagi *et al.*, 2006
	Oryza sativa (Indica)	AAB88295	594	CLA1	Campos *et al.*, 1997

Tabela 1. Enzimas da rota MEP clonadas e caracterizadas em várias espécies de plantas. Adaptado de Ganjewala *et al.* (2008). Continuação.

Enzima	*Planta*	Código GB*	Tamanho**	Proteína/Gene	Referência
		ABF18929		DXS1	
		AAD38941		DXS2	
		AAL32062		DXS3	
	Picea abies	ABS50520	746	DXS2B	Phillips *et al.*, 2007
		ABS50519	740	DXS2A	
		ABS50518	717	DXS1	
	Stevia rebaudiana	CAD22155	715	*dxs*	Totte *et al.*, 2003
	Tagetes erecta	AAG10432	725	*dxs*	Moehs *et al.*, 2001
DXR (EC 1.1.1.267)	*Arabidopsis thaliana*	CAB43344	406	*dxr*	Schwender *et al.*, 1999
	Antirrhinum majus	AAW28998	471	*dxr*	Dudareva *et al.*, 2005
	Camptotheca acuminata	ABC86579	472	Dxr	Yao *et al.*, 2008
	Catharanthus roseus	AAF65154	474	*dxr*	Veau *et al.*, 2000
	Chrysanthemum x morifolium	BAE79548	487	DXR	Kishimoto *et al.*, 2006
	Ginkgo biloba	AAR95700	477	Dxr	Gong *et al.*, 2005
	Hevea brasiliensis	AAS94121	471	DXR	Seetang-Nun *et al.*, 2008b
	Hordeum vulgare	CAE47438	484	*dxr*	Hans *et al.*, 2005
	Lycopersicon esculentum	AAK96063	475	DXR	Rodriguez-Concepcion *et al.*, 2001
	Mentha x piperita	AAD24768	470	DXR	Lange and Croteau, 1999ª
	Oryza sativa (Japonica)	NP_001041780	473	*dxr*	Ohyanagi *et al.*, 2006
	Oryza sativa (Indiaca)	EAY72208	473	*dxr*	Yu *et al.*, 2005
	Plectranthus barbatus	AAR99081	469	*dxr*	Engprasert *et al.*, 2005
	Pueraria montana var. lobata	AAQ84168	465	*dxr*	Sharkey *et al.*, 2005
	Rauvolfa verticillata	AAY87151	474	DXR	Liao *et al.*, 2007
	Salvia miltiorrhiza	ABJ80680	474	DXR	Wu *et al.*, 2009
	Stevia rebaudiana	CAD22156	473	*dxr*	Totte *et al.*, 2003
	Taxus cuspidata	AAT47184	477	*dxr*	Jennewein *et al.*, 2004
	Stevia rebaudiana	CAD22156	473	*dxr*	Totte *et al.*, 2003
	Taxus cuspidata	AAT47184	477	*dxr*	Jennewein *et al.*, 2004
CMS (EC 2.7.7.60)	*Arabidopsis thaliana*	NP_565286	302	*ISPD*	Seki *et al.*, 2002
		BAC42737	302	*ispD*	
	Ginkgo biloba	AAZ80386	327	*MECT*	Kim *et al.*, 2006
	Oryza sativa (Japonica)	BAD82130	297	*ispD put.*	Sasaki and Matsumoto, 2002
	Oryza sativa (Indica)	EAY76759	408	*ispD pat.*	Yu *et al.*, 2005
CMK (EC 2.7.1.148)	*Arabidopsis thaliana*	O81014	383	*ISPE*	Lin and Kaul, 1999
	Lycopersicon esculentum	AAF87717	401	*ISPE*	Rohdich *et al.*, 2000
	Mentha x piperita	P56848	405	*ISPE*	Lange and Croteau, 1999b
	Oryza sativa (Japonica)	NP_001044544	401	*ISPE*	Ohyanagi *et al.*, 2006

Tabela 1. Enzimas da rota MEP clonadas e caracterizadas em várias espécies de plantas. Adaptado de Ganjewala *et al.* (2008). Continuação.

Enzima	Planta	Código GB*	Tamanho**	Proteína/Gene	Referência
MCS (EC 4.6.1.12)	*Arabidopsis thaliana*	AAM62786	231	*MECDP_S*	Gao *et al.*, 2006
	Ginkgo biloba	AAY40863	239	*Mecps*	Alexandrov *et al.*, 2006
	Oryza sativa (Japonica)	EAZ24186	222	*MECDP_S*	Yu *et al.*, 2005
	Oryza sativa (Indica)	EAY87077	222	*MECDP_S*	Yu *et al.*, 2005
	Taxus x media	ABB88956	247	*mecs*	Jin *et al.*, 2006
HDS (EC, 1.17.4.3)	*Nicotiana benthamiana*	AAS75817	268	*gcpE/ispG*	Page *et al.*, 2004
	Oryza sativa (Japonica)	AAO72576	608	*gcpE*	Cooper *et al.*, 2003
HDR (EC 1.17.1.2)	*Arabidopsis thaliana*	AAW82381	468	*HDR/ISPH*	Guevara-Garcia *et al.*, 2005
	Nicotiana benthamiana	AAS75818	166	*ispH LytB*	Page *et al.*, 2004

* Código de acesso GenBank.
** Tamanho em aminoácidos.

A rota MEP consiste em sete passos que transformam uma molécula de gliceraldeído-3-fosfato e piruvato em IPP e DMAPP em uma razão de 6:1 (Quadro 3). No passo inicial da rota (Fig. 2), os precursores, piruvato e gliceraldeído-3-fosfato são condensados a DXP pela enzima 1-desoxi-D-xilulose-5-fosfato sintase (DXS; Rohmer *et al.*, 1996; Sprenger *et al.*, 1997; Lois *et al.*, 1998). A formação do esqueleto carbônico C12 a partir de precursores C4 e C5 é mecanisticamente reminiscente das reações catalisadas por transcetolases. Similaridades de sequência entre o gene *dxs* de *E. coli* e genes codificantes de transcetolases (EC 2.2.1.1) sugeriam que a enzima cognata poderia catalisar a formação de unidades C12 e dióxido carbônico. Essa hipótese foi confirmada pela expressão do gene em uma variedade de *E. coli* recombinante (Harada *et al.*, 2009). A DXS realiza essa descarboxilação do tipo transcetolase a partir de piruvato e gliceraldeído-3-fosfato em presença de tiamina difosfato (TPP) e um dos cátions divalentes Mg^{2+} ou Mn^{2+}, os quais funcionam como carreadores de um ânion acil bem como no mecanismo catalítico das transcetolases incluindo o complexo da piruvato desidrogenase (EC 1.2.4.1), da piruvato descarboxilase (EC 4.1.1.1) e da acetolactato sintase (EC 4.1.3.18). Os genes *dxs* de plantas apresentam sequências putativas de peptídeos-sinal em conformidade com a localização da rota MEP em plastídios de plantas. A atividade catalítica das DXS é relativamente elevada (300-500 $mmol.mg^{-1}$) em comparação com as outras enzimas da rota (Eisenreich *et al.*, 2004).

Diversos grupos têm independentemente isolado e clonado genes que codificam DXS de várias plantas (Tabela 1). Em *A. thaliana,* o gene *cla1* encontrado demonstrou alta similaridade com DXS de outras plantas estudadas e sua suposta localização em plastídios foi demonstrada (Araki *et al.,* 2000). Os genes que codificam essas proteínas de várias espécies de plantas fazem parte de seus nucleomas (Dubey *et al*., 2003). Estudos têm demonstrado que o esqueleto carbônico da 1-desoxi-D-xilulose pode ser incorporado em piridoxal (vitamina B6) e no anel tiazol da tiamina (vitamina B1). Portanto, a 1-desoxi-D-xilulose ou um derivado dela parece participar como intermediário na rota biossintética das vitaminas B1 e B6 (Eisenreich *et al.*, 2004).

No segundo passo, 1-desoxi-D-xilulose-5-fosfato (DXP) é convertido a MEP pelo rearranjo intramolecular e pela redução em uma reação catalisada pela enzima DXP redutoisomerase (DXR). Esta reação representa o primeiro passo comprometido da rota, o que explica a origem da denomeação de rota MEP utilizada por muitos pesquisadores. A partir deste passo, todas as reações estão exclusivamente comprometidas com a produção de IPP (Guevaro-Garcia *et al.*, 2005).

A DXR foi primeiramente isolada de *E. coli* (Takahashi *et al.*, 1998). A atividade catalítica das DXR é substancialmente menor do que a das DXS em uma faixa de 12 $mmol.mg^{-1}.min^{-1}$. A enzima pode utilizar Mg^{+2} ou Mn^{+2} como co-fator. A reação catalisada por esta enzima é iniciada por um rearranjo no esqueleto carbônico seguido de uma redução que requer a utilização da coenzima NADPH. A fosmidomicina é um potente inibidor desta enzima em concentrações nanomolares. A reação da DXR é reversível, porém em condições fisiológicas, a formação de MEP é favorecida. Cada subunidade do homodímero consiste de um domínio de ligação dinucleotídeo N-terminal, um domínio conector com o sítio catalítico e um domínio C-terminal helical (Eisenreich *et al.*, 2004). A estrutura do complexo com NADPH (Yajima *et al.*, 2002) confirmou o papel essencial dos resíduos Gly-14, Glu-231, His-153, His-209 e His-257 para o processo catalítico (Kuzuyama *et al.*, 2000). O grupo metila da metionina (Met-214) está localizado próximo ao anel nicotinamida do NADPH (Pellecchia *et al.*, 2002). Uma estrutura cristalográfica de DXR de *E. coli* complexada com Mn^{2+} e com fosmidomicina indica que este inibidor liga a enzima de maneira similar ao DXP (Steinbacher *et al.*, 2003). O conhecimento detalhado do modo de interação da fosmidomicina com a proteína pode contribuir para o desenvolvimento de novos potentes inibidores (Eisenreich *et al.*, 2004). Em plantas, diferentes pesquisadores clonaram genes que codificam DXR (Schwender *et al*., 1999) e expressaram estes em *E.*

coli. Carretero-Paulet *et al*. (2002) clonaram o cDNA de um gene de cópia única de *A. thaliana* que codifica DXR, e a análise da sequência de sua proteína prevê a presença de um peptídeo N-terminal para localização no plastídio, com um sítio de clivagem conservado e uma região conservada rica em prolinas na região N-terminal da proteína madura (Carretero-Paulet *et al*., 2002). Em plantas trangênicas de menta (*Menta piperita*), a superexpressão de DXR levou a um aumento em monoterpenos em folhas, comparado com plantas do tipo selvagem. De forma oposta, o silenciamento parcial do gene *dxr* levou à redução do acúmulo de óleos essenciais (Mahmoud & Croteau, 2001).

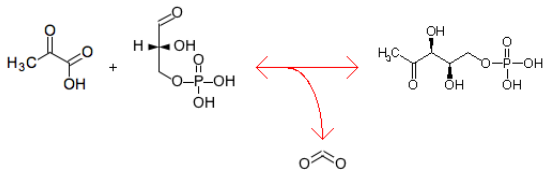
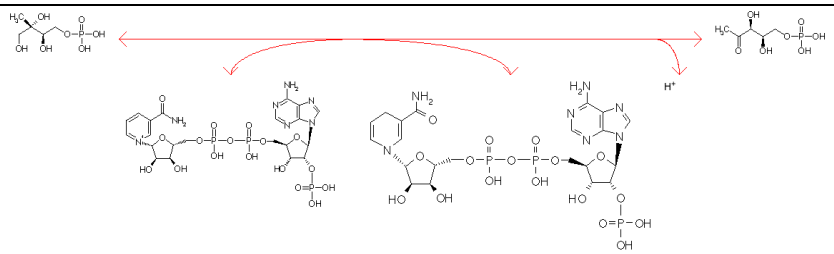
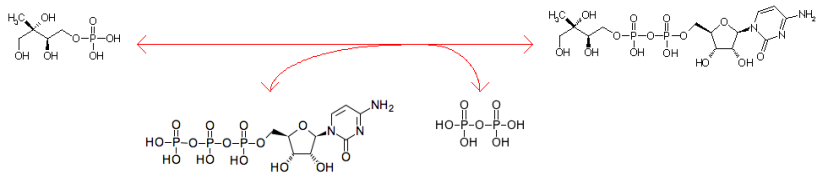
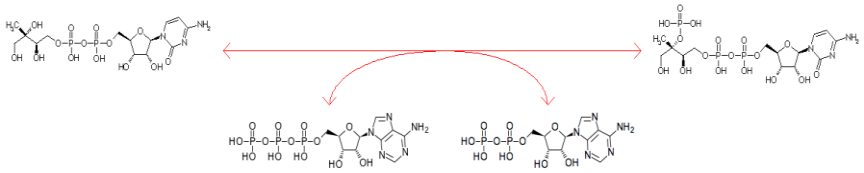
Seguindo na rota, MEP é convertido a 4-difosfocitidil-2C-metil-D-eritritol (CDP-ME) pela enzima CDP-ME sintase (CMS). Esta enzima catalisa a transferência de um grupamento fosfocitidil a MEP para formar CDP-ME e um grupo pirofosfato inorgânico (Ganjewala *et al.,* 2008). A CMS de *E. coli* pode utilizar tanto íons Mg^{2+} como Mn^{2+} e cobalto (Co^{2+}) como cofatores, enquanto, em plantas, a enzima é também catalítica em presença de íons de níquel (Ni^{2+}). A atividade catalítica de CMS é relativamente baixa (20–70 $mmol.min^{-1}.mg^{-1}$) comparada com DXS (Eisenreich *et al.*, 2004).

Em plantas, o gene que codifica a CMS foi primeiramente clonado a partir de *Arabidopsis*, e um fragmento que codifica a sequência-alvo potencial de plastídio da proteína foi expressa em *E. coli* recombinante. Estudos têm demonstrado que CDP-ME marcado radioativamente é incorporado eficientemente em carotenóides de pimenta vermelha (*C. annum*), demonstrando o papel essencial desta enzima como similar a CMS de *A. thaliana* (Herz *et al*., 2000). A enzima de *Arabidopsis* requer um cátion divalente, preferencialmente Mg^{+2} e a análise da sequência da proteína de *Arabidopsis* revelou uma suposta sequência de importação para o plastídio (Dubey *et al*., 2003).

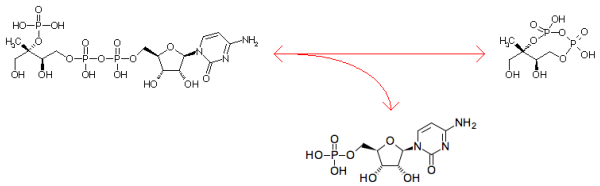
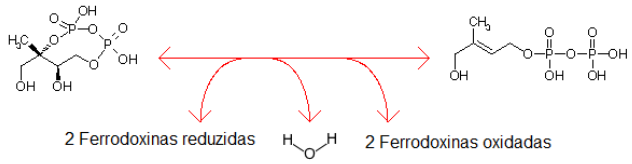
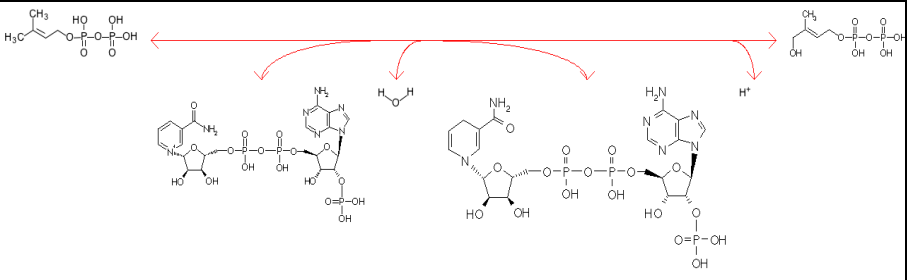
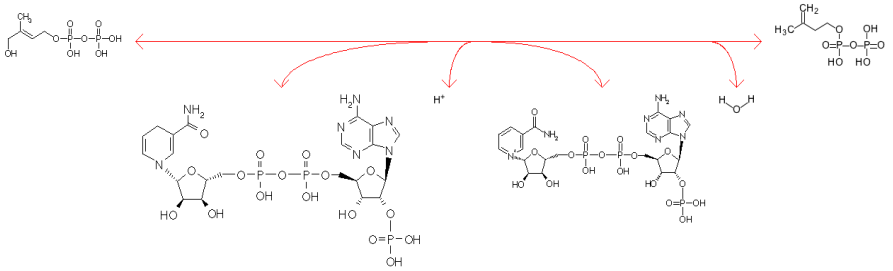
No quarto passo da rota MEP, a enzima CDP-ME quinase (CMK) fosforila CDP-ME para formar 4-difosfocitidil-2C-metil-D-eritritol-2-fosfato (CDP-ME2P). A proteína recombinante de *E. coli* atua fosforilando o grupamento 2-OH de CDP-ME2P. A enzima requer ATP como segundo substrato e Mg^{2+} como cofator (Eisenreich, *et al.*, 2004). Há poucos estudos em plantas que demonstram o possível papel da CMK na formação de CDP-ME2P, e a completa enzimologia e análise molecular desta enzima também não está disponível. Experimentos utilizando CDP-ME marcado com ^{14}C demonstraram que os átomos marcados foram eficientemente incorporados a carotenóides em *C. annuam*, sugerindo um possível papel dessa enzima (Rhodich *et al.*, 2000). Genes ortólogos de tomate e *M. piperita* carregam uma suposta sequência-alvo

para plastídio e um suposto sítio de ligação para ATP (Dubey *et al*., 2003). O CDP-ME2P, produzido no passo anterior, é utilizado como substrato pela 2C-metil-D-eritritol-2,4-ciclodifosfato (ME-cPP) sintase (MCS) para a produção de ME-cPP com liberação de uma molécula de citidil monofosfato (CMP) durante a catálise. A reação é reversível, mas tende à formação de ME-cPP em ambiente fisiológico. A enzima requer Mg^{2+} ou Mn^{2+} como co-fator. Estruturas cristalográficas da proteína homotrimérica indicaram que um íon Zn^{2+} coordenado por dois resíduos conservados de histidina e por um resíduo de Asp ajuda a posicionar o substrato no sítio ativo e facilita o ataque nucleofílico do grupo 2-fosfato. Após, ocorre a conversão de ME-cPP para IPP e DMAPP por meio de um intermediário 1-hidroxi-2C-metil-2-(E)-butenil-4-difosfato (HMBPP), envolvendo passos subsequentes catalisados por duas enzimas, a HMBPP sintase (HDS) e a HMBPP reductase (HDR). A análise da sequência de aminoácidos revelou um gene ortólogo codificando HDS em *A. thaliana* que é altamente homólogo ao gene *isp*G de *E. coli*, o qual catalisa a reação de conversão de ME-cPP para HMBPP por ação da HDS (Adam *et al*., 2002). Estudos demonstraram que a baixa atividade catalítica de cromoplastos isolados de pimenta vermelha (*C. annum*) foi aumentada pela adição da proteína HDR purificada a partir de *E. coli* (Eisenreich et al., 2001; 2004). Essa enzima é capaz de utilizar o HMBPP produzido no passo anterior e convertê-lo a uma mistura de IPP e DMAPP em uma proporção de 6:1 (Eisenriech *et al.*, 2004).

Quadro 3. Enzimas da rota MEP e suas respectivas reações de catálise. Disponível na base da KEGG.

Enzima (EC)	Reação
DXS (2.2.1.7)	 Piruvato + D-Gliceraldeído-3-fosfato <=> 1-Desoxi-D-xilulose 5-fosfato + CO_2
DXR (1.1.1.267)	 2-C-Metil-D-eritritol-4-fosfato + $NADP^+$ <=> 1-Desoxi-D-xilulose 5-fosfato + NADPH + H^+
CMS (2.7.7.60)	 2-C-Metil-D-eritritol 4-fosfato + CTP <=> 4-(Citidina-5'-difosfo)-2-C-metil-D-eritritol + Difosfato
CMK (2.7.1.148)	 4-(Citidina-5'-difosfo)-2-C-metil-D-eritritol + ATP <=> 2-Fosfo-4-(citidina-5'-difosfo)-2-C-metil-D-eritritol + ADP

Quadro 3. Enzimas da rota MEP e suas respectivas reações de catálise. Disponível na base da KEGG. Continuação.

MCS (4.6.1.12)	 2-Fosfo-4-(citidina-5'-difosfo)-2-C-metil-D-eritritol <=> 2-C-Metil-D-eritritol 2,4-ciclodifosfato + CMP
HDS (1.17.7.1)	 2-C-Metil-D-eritritol 2,4-ciclodifosfato + 2 Ferredoxinas reduzidas <=> 1-Hidróxi-2-metil-2-butenil 4-difosfato + H_2O + 2 Ferredoxinas oxidadas
HDR (1.17.1.2)	 Dimetilalil difosfato + $NADP^+$ + H_2O <=> 1-Hidróxi-2-metil-2-butenil-4-difosfato + NADPH + H^+  1-Hidróxi-2-metil-2-butenil 4-difosfato + NADPH + H^+ <=> Isopentenil difosfato + $NADP^+$ + H_2O

A jusante de ambas as rotas, a atividade das preniltranferases é responsável pela conversão de IPP e DMAPP em cadeias de precursores isoprenóides para a elaboração de terpenóides mais complexos. Alguns exemplos são:

- geranil difosfato (GPP), o precursor C10 dos monoterpenos (os óleos essenciais são exemplos de monoterpenos);
- farnesil difosfato (FPP), o precursor C15 dos sesquiterpenos e triterpenos;
- geranilgeranil difosfato (GGPP), o precursor C20 dos diterpenos.

Em plantas, como já mencionado, a síntese dos terpenos é compartimentalizada de maneira que os monoterpenos e diterpenos (carotenóides e clorofilas, por exemplo) são produzidos via rota MEP nos plastídios, enquanto os sesquiterpenos e triterpenos são produzidos no citosol via rota MVA (Fig. 3). No entanto, existem evidências de que ocorra expressão da geranilgeranil difosfato sintase no citosol e de que existam níveis significantes de FPP nos plastídios (Kirby *et al.*, 2009).

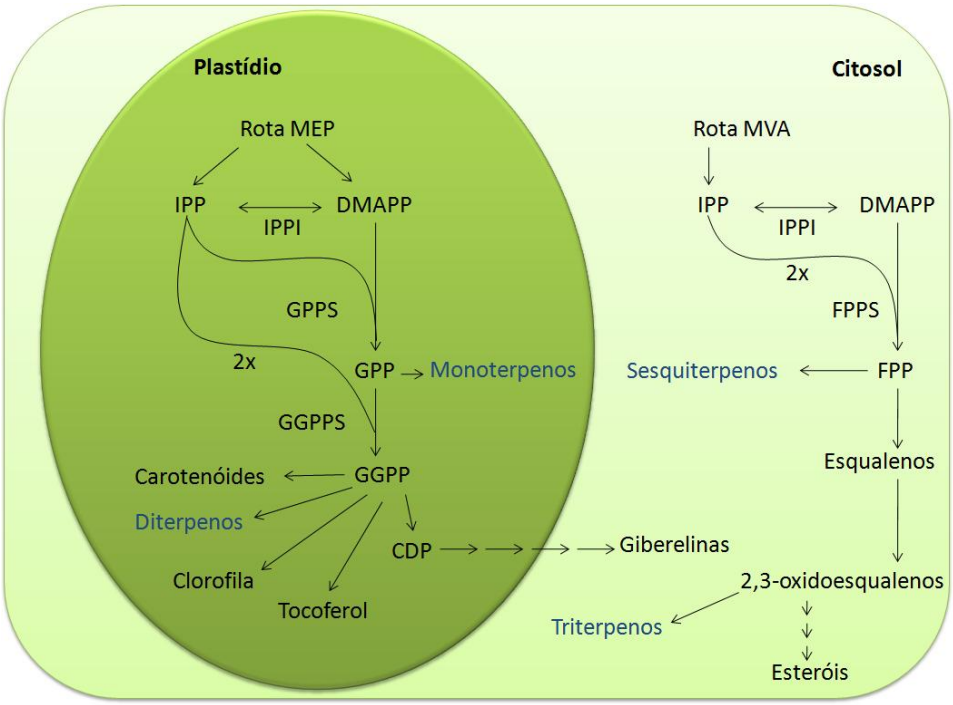


Figura 3. Esquema generalizado da produção de isoprenóides em plantas. As reações mostradas formam as rotas de produção de prenilfosfato mais comumente aceitas. CDP, citidil difosfato; DMAPP, dimetilalil difosfato; DXP, 1-desoxi-D-xilulose-5-fosfato; FPP, farnesil difosfato; FPPS, FPP sintase; GPP, geranil difosfato; GGPP, geranilgeranil difosfato; GPPS, GPP sintase; GGPPS, GGPP sintase; IPP, isopentenil difosfato; IPPI, IPP isomerase (modificado de Kirby *et al.*, 2009).

Muitos estudos permitiram evidenciar uma troca de precursores prenil difosfato entre plastídios e citosol. Entretanto, a extensão dessa troca é limitada, sendo incapaz de compensar uma perda de fluxo por uma das rotas, como foi demonstrado em estudos que utilizaram *A. thaliana* mutante para genes da rota MEP (Bouvier *et al.*, 2005).

Um detalhado conhecimento sobre os genes, enzimas e intermediários das reações da rota independente de MVA pode conduzir ao desenvolvimento de novas drogas derivadas de plantas de origem terpenóide e produção de vários terpenos importantes comercialmente em plantas por meio da biotecnologia moderna (Dubey *et al.*, 2003). Enzimas da rota MEP representam alvos em potencial para o desenvolvimento de novos inibidores. Pode-se antecipar o desenvolvimento de novos herbicidas contra plantas e algas assim como substâncias antimicrobianas contra bactérias patogênicas que possuem a rota MEP (Lichtenthaler, 1999). Adicionalmente, o rápido aumento da resistência de patógenos humanos às drogas disponíveis gera urgência na demanda por novas substâncias terapêuticas. Exclusivamente, a rota MEP é empregada por muitos microrganismos patogênicos, assim como pelo *Plasmodium falciparum* e pelo *Mycobaterium tuberculosis*. Esta rota não ocorre em humanos e animais superiores. Isto torna a rota MEP de biossíntese de isoprenóides um alvo ideal para o desenvolvimento de novos antibióticos e drogas para o tratamento da malária e da tuberculose. Em plantas, a inibição da rota DXP pode ser a base para o desenvolvimento de novos herbicidas. O conhecimento detalhado do mecanismo e da regulação dessa rota poderá beneficiar a produção biotecnológica de isoprenóides de interesse comercial como óleos essenciais, carotenóides, clorofila e hormônios vegetais (Eisenreich *et al.*, 2001).

1.3 Óleos Essenciais de Eucalipto

Os óleos essenciais de eucalipto são compostos por uma complexa mistura de componentes orgânicos, frequentemente envolvendo de 50 a 100 ou até mais componentes isolados (Zini, 2002), apresentando grupos químicos como hidrocarbonetos, álcoois aldeídos, cetonas, ácidos e ésteres (Vitti e Brito, 1999). No caso dos eucaliptos, as referências são de que as ocorrências de óleos essenciais estariam relacionadas com a defesa da planta contra insetos, resistência ao frio quando no estágio de plântula, ao efeito alelopático e à redução da perda de água, resultados estes que dependem ainda de mais estudos comprobatórios (Doran, 1991).

Existem vários fatores que têm sido citados por exercerem influência na obtenção de óleos essenciais de eucalipto, os mais típicos são: a variabilidade genética, a idade das folhas, as condições ambientais, o tipo de manejo florestal, os métodos utilizados para amostragem das folhas, os processos de extração e de análise do óleo (Vitti e Brito, 2003).

Tabela 2. Principais espécies de eucalipto utilizadas para a produção de óleos essenciais, componentes principais, teores e rendimento em óleos extrativos (adaptado de Vitti e Brito, 2003).

Espécies	**Componente Principal**		**Rendimento**
	Nome	**Teor (%)**	**(%)***
Óleos medicinais			
E. camaldulensis	Cineol	80-90	0,3-2,8
E. cneorifolia	Cineol	40-90	2,0
*E. dives (*var. *cineol*)	Cineol	60-75	3,0-6,0
E. dumosa	Cineol	33-70	1,0-2,0
E. alaeophara	Cineol	60-80	1,5-2,5
E. globulus	Cineol	60-85	0,7-2,4
E. lecoxylon	Cineol	65-75	0,8-2,5
E. oleosa	Cineol	45-52	1,0-2,1
E. polybractea	Cineol	60-93	0,7-5,0
E. radiata subesp. *Radiata* (var. *cineol*)	Cineol	65-75	2,5-3,5
E. sideroxylon	Cineol	60-75	0,5-2,5
E. smithii	Cineol	70-80	1,0-2,2
E. tereticornis	Cineol	45	0,9-1,0
E. viridis	Cineol	70-80	1,0-1,5
Óleos industriais			
E. dives (var. *felandreno*)	Felandreno	60-80	1,5-5,0
E. dives (var. *piperitona*)	Piperitona	40-56	3,0-6,5
E. elata (var. *piperitona*)	Piperitona	40-55	2,5-5,0
E. radiata subesp. *Radiata* (var. *felandreno*)	Felandreno	35-40	3,0-4,5
Óleos para perfumaria			
E. citriodora (var. *citronelal*)	Citronelal	65-80	0,5-2,0
E. macarthurii	Acido de geranil	60-70	0,2-1,0
E. staigerana	Citral (a+b)	16-40	1,2-1,5

*Rendimento base de peso de folha fresca

Óleos essenciais com quantidade mínima de 70% de cineol são classificados como medicinais (Tabela 2) e podem ser destinados à fabricação de produtos farmacêuticos como inalantes, estimulantes de secreção nasal, produtos de higiene bucal ou podem ser utilizados para dar sabor e aroma aos medicamentos. *E. globulus* é a

principal espécie produtora deste tipo de óleo no Brasil, havendo também algumas referências de extração a partir do *E. smithii*. Os óleos essenciais compostos principalmente por felandreno são utilizados na indústria como solventes e como matéria prima na produção de desinfetantes e desodorizantes. Os óleos ricos em piperitona são utilizados para a fabricação do timol, um preservativo para gomas, pastas e colas. Finalmente, os óleos que predominantemente possuem mentol são utilizados como aromatizantes de produtos medicinais. Na indústria de perfumaria, são utilizados aqueles óleos ricos em citronelal, sendo o *E. citriodora* a principal espécie explorada para esse fim no Brasil. Este óleo é utilizado na composição de perfumes para diversos fins, sendo mais usado nos produtos de limpeza, como sabões e desinfetantes. *E. staigeriana,* cultivado em pequena escala no Brasil, também é utilizado para extrações de óleos essenciais para perfumaria, sendo que o principal componente químico de seu óleo essencial é o citral (Vitti e Brito, 2003).

Os principais consumidores de óleo essencial são os países da Comunidade Européia, sendo a França, a Alemanha e a Inglaterra os maiores compradores. Em 1990, estes países importaram a quantia de 1.840 toneladas de óleo de eucalipto. Os Estados Unidos da América são também considerado um grande consumidor, tendo importado no ano de 1990 cerca de 378 toneladas de óleo (Braga, 2001).

1.4 Objetivos

1.4.1 Objetivo Geral

O presente trabalho foi desenvolvido com vistas a identificar genes potencialmente relacionados com a síntese de isoprenóides (óleos essenciais) na coleção de ESTs do Projeto Genolyptus e estudá-los para a comprovação de suas funções.

1.4.2 Objetivos Específicos

Os objetivos específicos que nortearam o desenvolvimento dos trabalhos que resultaram na composição da presente Dissertação de Mestrado foram:

- A recuperação, a partir da seleção dos plasmídeos de clonagem pSPORT1 (Invitrogen) purificados do projeto Genolyptus, do(s) clone(s) selecionado(s)

representantes dos genes *dxs, dxr, mdc*, *ippi1 e ippi2.* Além da obtenção e a caracterização da sequência completa do(s) mesmo(s);

- A análise *in silico* da estrutura dos genes *dxr*, *ippi1* e *ippi2;*
- A expressão heteróloga dos genes selecionados (*dxr*, *ippi1* e *ippi2*) que codificam 1-desoxi-D-xilulose-5-fosfato redutoisomerase e as isopentenil difosfato isomerases 1 e 2, respectivamente, fusionados ao gene codificador da glutationa-S-transferase (GST) em *E.coli*;
- A modelagem molecular da proteína codificada pelo gene *dxr* e o estudo da dinâmica molecular de uma proteína IPPI1 para a avaliação do papel estrutural da extremidade N-terminal peculiar às IPPIs eucarióticas.

CAPÍTULO II

CLONING, CHARACTERIZATION AND EXPRESSION OF 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE FROM *Eucalyptus grandis*.

Manuscrito de artigo científico redigido conforme normas do periódico
Journal of Plant Physiology

Running title: Cloning, characterization and expression of 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Eucalyptus grandis*.

Corresponding author: Pasquali G.

Mailing address: Av. Bento Gonçalves, 9500 - Building 43431 – Lab 212 Campus do Vale/UFRGS 91501-970 - Porto Alegre, RS - Brazil – Postal-box: 15005

Fone: 55-51-3308.9410 Fax: 55-51-3308.7309

E-*mail address*: pasquali@cbiot.ufrgs.br

Cloning, characterization and expression of 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Eucalyptus grandis*.

Genaro Azambuja Athaydes[a], Sinara Artico[a] and Giancarlo Pasquali[a]*

[a]Biotechnology Center, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil.

Summary

Terpenoids represent the most diverse class of chemical compounds in nature. Such complexity is reached through the ample specificity enzimatic activity characteristic of the soncondary metabolism enzymes which, some times, can lead to the formation of more than one product. The carbon skeleton used in these reactions has origins in the condensation of two five-carbon basic units, isopentenyl diphosphate and dimethylallyl diphosphate. Plants synthesise these basic building blocks from two different pathways, the citosolic mevalonate pathway, and the plastidic 2C-methyl-D-erythritol-4-phosphate (MEP) pathway. Via the MEP pathway, mono and diterpenes that are components of essential oils and chlorophyl are synthesised. Although a strong economical appeal for valuable terpenoids exists in the market, the yield of these compounds is unfavorable in certain species including *Eucalyptus grandis*. In order to start circunventing such limitations, better understanding of the metabolic pathway is necessary. The reaction catalised by 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR, EC: 1.1.1.267) is the first commited step of the MEP pathway. Although there are few reports on the gene cloning and characterization of DXR from plants, no work has described the isolation and characterization of *dxr* from *Eucalyptus* species. In this work we describe the cloning, sequence characterization, heterologous expression of the *E. grandis dxr* (EgDXR) gene, performing a comparative expression analysis in vascular and leaf tissues.

Keywords: 1-deoxy-D-xylulose-5-phosphate reductoisomerase, DXR, essential oils, *Eucalyptus grandis*, MEP pathway

Abreviations: DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; EST, *expressed sequence tag*; GST, glutathione S-transferase; MEP, 2C-methyl-D-erythritol-4-phosphate; MVA, mevalonate; ORF, *open reading frame*; *rnp*, ribonucleoprotein L23A gene

Introduction

Globally, broadleaves make up 40% of the planted forest area with *Eucalyptus* species being the main fast-growing, short-rotation crop, accounting for 17.8 million ha (Tournier *et al.*, 2003). About half of the forest plantation estate is for industrial end-use. *Eucalyptus grandis* is one of the most economically important species for the paper industries in Brazil and several countries around the world, since it fits well to the production of high-quality and high-opacity papers (Eldridge *et al.*, 1994). Besides cellulose pulp and paper produced from timber, essential oils extracted from leaves and young stems are the second most valuable product derived from *Eucalyptus* trees. The essential oils from *Eucalyptus* are composed by a complex mix of organic molecules, frequently ranging from 50 to 100 or more isolated compounds. These essential oils can be responsible for plant defense against insects, cold resistance, allelophatic effect, and reduction of water loss (Doran *et al.*, 1991). The main species of *Eucalyptus* used for the extraction of essential oils in Brazil are *E. citriodora*, *E. globulus* and *E. staigeriana* (Vitti & Brito, 2003). The *E. grandis* leaf biomass produced in Brazil surpasses that derived from these three species together by many millions of tons. *E. grandis* leaf biomass is normally incorporated into soils and, if a higher yield in essential oils would be achieved, it would be possibly directed at least in part to the production of these valuable compounds. Countries from the European Community such as France, Germany and the UK are major importers of these oils (Braga *et al.*, 2001).

A strategy that could improve the yield of essential oils in *E. grandis* would be the molecular engineering of the isoprenoid pathway. Isoprenoids are a large class of natural compounds found in all living organisms. Many isoprenoids have essential functions in cells, e.g. plastoquinone and ubiquinone act as electron transporters in photosynthesis and respiration, the phytol chain tail on chlorophyll molecules, sterols as structural components of cytosolic membranes, and abscisic acid, cytokinins, and gibberellins as plant growth regulators. Some isoprenoids have pharmaceutical and biotechnological applications, such as taxol and ginkgolides (Sacchettini and Poulter, 1997).

All isoprenoids are synthesized through consecutive condensations of two five-carbon precursors, isopentenyl diphosphate (IPP) and its allyl isomer dimethylallyl diphosphate (DMAPP). In plants, the biosynthesis of IPP and DMAPP proceeds by two independent biosynthetic pathways: the mevalonic acid (MVA) and the 2C-methyl-D-erythritol 4-phosphate (MEP) pathways (Lichtenthaler *et al.*, 1997). The MEP pathway involves seven enzymatic steps from pyruvate and glyceraldehyde 3-phosphate, in which the homodimer 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) is the second enzyme and it is known to catalise the first committed step in the formation of IPP and DMAPP, converting deoxyxylulose 5-phosphate (DXP) to MEP (Proteau, 2004). It was previously reported that the overexpression of DXR regulated the metabolic flux through the MEP pathway, resulting in increased levels of terpenoids in different types of plants (Mahmoud and Croteau, 2003; Hasunuma *et al.*, 2008). In order to better understand the role played by DXR in essential oil production in *Eucalyptus*, here we report the isolation, cloning, expression profile and heterologous expression of the *dxr* gene from *E. grandis*, named EgDXR.

Material and methods

cDNA Library construction and sequencing

Seeds of *E. grandis* (supplied by Aracruz Cellulose S.A., Guaíba, RS, Brazil) were surface sterilized by sequentially soaking in 70% ethanol for 2 min, 1% (v/v) active chlorine solution for 15 min and five times in sterilized water. Seeds were then placed in culture flasks (5 cm in diameter, 8 cm high) containing the Murashige & Skoog (MS, Invitrogen) complete medium solidified with 0.7% (w/v) Phytoagar (Duchefa) and left to germinate in the dark for two days at 26±2 °C. After germination, flasks were transferred to a culture room with a 16 h photoperiod at the same temperature. Total RNA was extracted from 60-days old *E. grandis* seedlings using the PureLink Plant RNA Purification Reagent (Invitrogen) according to the manufacturer's instructions for multi RNA minipreps. Messenger RNA was extracted by the employment of the Oligotex mRNA Purification System (Qiagen). A total amount of 5 μg mRNA was finally used for cDNA library construction by the Superscript Plasmid System with Gateway Technology for cDNA Synthesis and Cloning Kit (Invitrogen).

Plasmid DNA preparation was carried out in 96-well microplates using standard methods based on alkaline lysis and filtration in Millipore filter plates. Plasmid samples were sequenced using the automatic sequencer ABI-PRISM 3100 Genetic Analyzer armed with 50 cm capillaries and POP6 polymer (Applied Biosystems). DNA templates (30 to 45 ng) were labeled with 3.2 pmol of primer GlyptsRev1 (5'-ATAGGGAAAGCTGGTACGC-3') or M13 -40 forward (5'-GTTTTCCCAGTCACGACGTTGTA-3', Amersham Biosciences) and 2 μL of BigDye Terminator v3.1 Cycle Sequencing RR-100 (Applied Biosystems) in a final volume of 10 μL. Labeling reactions were performed in a GeneAmp PCR System 9700 (Applied Biosystems) termocycler with a initial denaturing step of 96 ºC for 3 min followed by 25 cycles of 96 ºC for 10 sec, 55 ºC for 5 sec and 60 ºC for 4 min. Labeled samples were purified by isopropanol precipitation followed by 70% ethanol rinsing. Precipitated products were suspended in 10 μL formamide, denatured at 95 ºC for 5 min, ice-cooled for 5 min and electroinjected in the automatic sequencer. Sequencing data were collected using the software Data Collection v1.0.1 (Applied Biosystems) Data generated were processed by a suite of programs available at the Universidade

Católica de Brasília (http://www.ucb.br/genolyptus) prior to assembly with the PHRED (Ewing, 1998a; b) and PHRAP (http://phrap.org/) algorithms. Accepted ESTs were selected based on a minimal length of 250 bp with every base having a quality of PHRED higher than 20.

Heterologous expression of *EgDXR*

For expression in *Escherichia coli* BL21 (DE3), *EgDXR* gene was transferred from pSPORT1-*dxr* into pSK^+ Bluescript (Stratagene) by *Bam*HI and *Eco*RI digestion. *EgDXR* was then cloned into pGEX-4T-1 (GE Healthcare) by *Sal*I and *Not*I digestion. Recombinant *E. coli* cells harboring pGEX-4T-1-*dxr* were cultivated at 37 ^oC overnight and inoculated into 500 mL erlenmeyer flasks containing 50 ml Luria-Bertani (LB) medium supplemented with 100 g/mL ampicillin. When an OD_{600} of about 0.45 was reached, the expression was induced by adding isopropyl β-D-thiogalactopyranoside (IPTG) at 100 µg/mL. Cultures were kept for addicional 3, 6 or 14 h at 30 °C. For each culture period, a non-induced control culture was kept under the same conditions. Recombinant and control *E. coli* cells were harvested by centrifugation and then disrupted by ultrasonication (Feliu *et al.*, 1998). A soluble protein fraction was finally obtained after centrifugation at 12,000 ×g for 20 min for SDS-PAGE.

Protein electrophoresis

SDS-PAGE was performed essentially as described by Laemmli (1970). SDS-PAGE was performed on a Mini-PROTEIN^{TM} electrophoresis system (Bio-Rad) with 12% polyacrylamide gels (30% acrylamide/0.8% bis-acrylamide) under denaturing conditions.

Bioinformatic analysis

Comparative and bioinformatic analysis were carried out with online tools available at the *National Center for Biotechnology Information* (NCBI; http://www.ncbi.nlm.nih.gov), Expasy (http://cn.expasy.org) and Center for Biological Sequence Analysis (http://www.cbs.dtu.dk/services/ChloroP/). Nucleotide sequence, deduced amino acid sequence and *open reading frame* (ORF) were analyzed, and

sequence comparisons were conducted by database search using the BLAST program (NCBI). The phylogenetic analysis of *EgDXR* and DXRs from other plant species was performed with CLUSTAL W version 2.0 (Larkin *et al.*, 2007) using default parameters. A phylogenetic tree was constructed using MEGA version 4.0 (Kumar *et al.*, 2001) from the resulting CLUSTAL W alignments. The neighbor-joining method (Saitou and Nei, 1987) was used to construct the tree. Homology-based three-dimensional (3D) structural modeling of EgDXR was accomplished by Swiss-Modeling (Schwede *et al.*, 2003) and WebLab ViewerLite was used for 3D structure displaying (homology-based modeling by Swiss-Model).

Microarray Analysis

To investigate the *EgDXR* expression pattern in leaves and xylem, 20 µg of each RNA sample from leaves and xylem of adult *E. grandis* trees were dried down in speed-vac and sent to NimbleGen Systems Inc. (Reykjavik, Iceland). Nine 50-mer probes were designed and synthesized in duplicates from each of the 21.442 ESTs from the Genolyptus Project at NimbleGen. Synthesis of Cy3-labeled cDNAs, hybridizations, washings, scanning and preliminary analysis were carried out by NimbleGen, following a standard expression design. Normalized gene expression values for each feature were yielded taking into account the hybridization values for all nine probes per gene, in duplicates. The final analysis was performed using MultiExperiment Viewer (MEV), developed at The Institute for Genomic Research (TIGR). Data normalization was performed using the $2^{-\Delta\Delta Ct}$ method (Livak *et al.*, 2001) adapted to microarray results and a Student´s *t-test* was performed with the Statistical Package for the Social Sciences Software (SPSS 12.0 for Windows; Leech *et al.*, 2005). Relative expression of the genes in leaves and xylem was calibrated with an internal reference-gene encoding the ribonucleoprotein L23A. Detailed technical descriptions of the procedure are referred by Bastolla (2007), Jahns (2007) and Pasquali *et al*. (in preparation).

Results and Discussion

Cloning of the full-length cDNA of *EgDXR*

DXR genes have been cloned and characterized from bacteria (Takahashi *et al.*, 1998; Harker *et al.*, 1999; Altincicek *et al.*, 2000; Grolle *et al.*, 2000; Miller *et al.*, 2000; Cane *et al.*, 2001; Koppisch *et al.*, 2002; Hoeffler *et al.* 2002; Radykewicz *et al.*, 2000; the protozan *Plasmodium falciparum* (Jomaa *et al.*, 1999), and plant species (Ganjewala *et al.* 2009), However, there is no report up to date on the cloning of *dxr* genes from any *Eucalyptus* species.

Based on the conserved region of plant *dxr* sequences, we were able to locate a pSPORT1 plasmid harboring the potentially complete *dxr* gene sequence in the plasmid library of the Genolyptus Project. One single contig was found in the transcriptome database. This contig is composed by 3 reads, all belonging to *E. grandis* libraries. Since we did not find any other kind of cDNA from *dxr* gene in the libraries, it is possible that *dxr* is present in one copy in the genome, which is common for the majority of plants except for *Hevea brasiliensis* which was proposed to have two copies (Seetang-Nun *et al.*, 2008). A DNA fragment of approximately 1,500 bp was subcloned into pSK^+ Bluescript (Stratagene). Resequencing of the clone confirmed its identity (Fig. 1). The sequence obtained exhibited 1,671 bp and spans a 1,422 bp ORF. The 473 amino acids EgDXR predicted peptide has a calculated molecular mass of 51.9 kDa and an isoelectric point of 6.36 (http://cn.expasy.org/tools/protparam.html; Fig. 1).The deduced amino acid sequence of *EgDXR* was submitted to the NCBI for PSI-BLAST searching (http://www.ncbi.nlm.nih.gov/BLAST/) and the result showed that EgDXR has high homology with DXR sequences from other plant species, especially with the DXR sequences from *P. trichocarpa* and *H. brasilienses*, sharing 88% identity. EgDXR exhibited quite high identities to other plant DXRs including *S. lycopersicum* (86%), *R. verticillata* (86%), *C. roseus* (85%), *A. majus* (84%), *P. kurrooa* (83%), *S. miltiorrhiza* (84%), *N. tabacum* (85%), *C. stellatopilosus* (87%), *O. sativa* (85%), *Z. mays* (85%), *C. acuminata* (84%), and *G. biloba* (78%), suggesting that EgDXR belongs, indeed, to the DXR family with a very high degree of sequence conservation.

The amino acid sequences of 14 plant DXRs were aligned according to the algorithm of CLUSTAL W (Higgins *et al.*, 1994). The result, as shown in Fig. 2, allowed us to conclude that DXRs have indeed a very high aminoacid similarity

throughout the entire coding region. Hence, the three common domains found in all plant DXR sequences were clearly located in the EgDXR sequence (Carretero-Paulet *et al.*, 2002; Rane and Calvo, 1997; Lange and Croteau, 1999; Schwender *et al.*, 1999; Kuzuyama *et al.*, 2000).

The first domain is a transit peptide located in the N-terminal region of EgDXR, which possibly directs the enzyme to plastids where the MEP pathway is known to operate in plants (Eisenreich *et al.*, 1998). It was found that the cleavage site of this transit peptide is located in the conserved $(H/Q)^{49}\downarrow(C/S)^{50}(S/L)$ motif (Fig. 1 and 2). The transit sequence of EgDXR and its hypothetical cleavage site were predicted with the use of the ChloroP program (Emanuelsson *et al.*, 1999). The second domain is characterized by a proline-rich motif described as P(P/Q)PAWPG(R/T), typical of the plant DXR protein family and located immediately downstream of the plastid signal cleavage site. Both motifs are not present in prokaryotic DXRs (Carretero-Paulet *et al.*, 2002). EgDXR clearly exhibits this proline-rich motif at its N-terminus, as shown in Fig. 1 and 2. An exception in plant DXRs, however, was reported for *Vanda* Mimi Palmer, which does not have the transit sequence and, therefore, it was suggested that the protein is localized in the citosol (Chan *et al.*, 2009). More effort is clearly needed to elucidate the role of DXR from *Vanda* Mimi Palmer and the biological function of the proline-rich domain in all plant DXRs. The PSI-BLAST also allowed us to identify a third highly conserved domain of DXRs, a NADPH binding motif characterized by GSTGS(I/V)GT (Rane and Calvo, 1997; Lange and Croteau, 1999; Schwender *et al.*, 1999; Kuzuyama *et al.*, 2000). In the multiple alignment analysis, it was revealed that EgDXR exhibits almost the same aminoacid composition of this substrate binding motif as all other fourteen plant DXRs (Fig. 2).

Since *EgDXR* was the first *dxr* gene cloned from an *Eucalyptus* species, it would be interesting to investigate its evolutionary position among the phylogenetic tree of various DXRs. A phylogenetic tree was constructed using the program MEGA and the results from CLUSTAL W alignments based on the deduced amino acid sequences of *EgDXR* and other DXRs (Fig. 3). The result revealed that DXRs were derived from an ancestor gene that, among sequences tested, exhibited closer resemblance to the actual *M. tuberculosis* DXR. According to the phylogenetic tree, *EgDXR* is among the most distant plant DXRs in comparison to the bacterial group, having the closest relationship with DXR from *L. esculentum*. The DXR protein from *G. biloba* is more ancient than other DXRs in the plant group.

The homology-based 3D structural modeling of EgDXR was analyzed by SWISS-MODEL (Schwede *et al.*, 2003; http://swissmodel.expasy.org/) using the *E. coli* DXR crystal structure (PDB ID 2EGH) as template. The overall folding of EgDXR confirms that this protein is a typical DXR built from β-sheets connected by turns and loops, creating a very tight structural scaffold (Fig. 4). EgDXR was found to have many characters commonly possessed by the DXR protein family especially in domains involved in NADPH binding. Thus, the structural modeling of EgDXR allowed us to show that this protein is very similar to other DXRs from bacteria whose biochemical function has been proved (Grolle *et al., 2000;* Kuzuyama *et al., 2000;* Henriksson *et al.*, 2006), strongly indicating that EgDXR possess its supposed metabolic function.

Expression of the *E. grandis dxr* gene in *E. coli*

To identify and characterize the DXR enzyme from *E. grandis*, the corresponding *EgDXR* gene was cloned into the expression vector pGEX-4T-1 (GE Healthcare), which allows the addition of a glutathione S-transferase (GST) tag to the N-terminal region of the protein of interest. Initially, only soluble fractions of crude cell extracts were analysed by SDS-PAGE. A significant amount of protein was observed after 14 h when the cultures were incubated with 0.1 mM IPTG (Fig. 5, lane 9). The molecular mass of the GST-EgDXR protein was approximately 75 kDa, as a result of the fusion of 23 kDa from GST and 51.9 kDa expected from the deduced amino acid composition of *EgDXR*. The calculated molecular mass of EgDXR was in good agreement with that determined by SDS-PAGE.

Microarray Gene Expression Analysis

In order to have a first clue about the pattern of *EgDXR* gene expression in mature leaf and vascular tissues of *E. grandis* adult trees, previous results of microarray analysis (Pasquali *et al.*, in preparation) were evaluated specifically for the gene of interest. To do so, two genes known to be constitutively expressed in xylem and leaves of *Eucalyptus* (Bastolla, 2007; Jahns, 2007), ribonucleoprotein L23A (*rnp*) and histone H2B were employed as house-keeping, reference genes. The microarray signal intensities of all samples were normalized with the *rnp* gene signal acquired from cDNAs derived from leaf tissues. As shown in Fig. 6, this analysis revelated that

EgDXR is significantly more expressed in leaves than xylem tissues. Our microarray results confirmed the importance of this gene to the metabolism of isoprenoids. Leaf is in fact the tissue where monoterpenes and diterpenes (i.e. essential oils, carotenoids, and chlorophyls) are produced through the MEP pathway in the plastids, as previously reported for the *A. thaliana dxr* gene (Carretero-Paulet *et al.*, 2002; Kirby *et al.*, 2009). Although validation of these results needs to be achieved by quantitative real-time RT-PCR, for instance, it is expected that key genes of the MEP pathway may be more expressed in leaves of an adult plant than in non-photosynthetic, vascular tissues as previously reported (Yan *et al.*, 2009).

In this work we cloned and characterized the first *dxr* gene from an *Eucalyptus* species, putatively related to isoprenoid biosynthesis. Our results strongly suggest that the *EgDXR* gene encodes a fully functional protein based on amino acid sequence homology and structural modeling of the protein with biochemically characterized DXRs. The presence of a typical plastid signal peptide and a proline-rich motif indicate that the EgDXR protein is localized in plastids where the MEP pathway takes place. Additionally, expression analysis revealed a much stronger mRNA accumulation in leaves than xylem tissues of *E. grandis* adult trees. We hope that the availability of this sequence could help future experiments focused on understanding the full role played by *EgDXR* gene in the biosynthesis of monoterpenes and sesquiterpenes compounds such as those constituents of essential oils in *Eucalyptus*.

Aknowledgments

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) MCT, Brasília, Brazil; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Ministério da Educação), MEC, Brasília, Brazil; Financiadora de Estudos e Projetos (FINEP) along with the projects "Genolyptus – Rede Brasileira de Pesquisa do Genoma do *Eucalyptus*" and "WoodGenes - Gênese de Madeira em *Eucalyptus*: Genes, Funções, Regulação e Expressão Transgênica" through the Fundos Setoriais Verde-Amarelo e de Biotecnologia do Ministério da Ciência e Tecnologia (MCT) and paper and cellulose companies including Aracruz Celulose S.A., Grupo Raiz, Celulose Nipo-Brasileira S.A. – CENIBRA, International Paper do Brasil Ltda., Jarí Celulose S.A., Klabin S.A., Lwarcel Celulose e Papel Ltda., Veracel Celulose S.A., Votorantim Celulose e Papel S.A., Zanini Florestal S.A. e Suzano-Bahia Sul Papel. The authors are also grateful to the Programa de Pós-Graduação em Biologia Celular e Molecular/UFRGS (Brazil).

References

Altincicek B, Hintz M, Sanderbrand S, Wiesner J, Beck E, Jomaa H. Tools for discovery of inhibitors of the 1-deoxy-D-xylulose 5-phosphate (DXP) synthase and DXP reductoisomerase: an approach with enzymes from the pathogenic bacterium *Pseudomonas aeruginosa.* FEMS Microbiol Lett 2000;190:329-33.

Bastolla FM. Seleção e avaliação de genes de referência para estudos de expressão gênica em *Eucalyptus*. Porto Alegre: UFRGS, 2007, 97 p. Dissertation (M.Sc.) – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Braga NC. Os óleos essenciais no Brasil: estudo econômico. Rio de Janeiro: Instituto de óleos. 2001; p 158-61.

Cane DE, Chow C, Lillo A, Kang I. Molecular cloning, expression and characterization of the first three genes in the mevalonate-independent isoprenoid pathway in *Streptomyces coelicolor.* Bioorganic Med Chem 2001;9:1467–77.

Carretero-Paulet L, Ahumada I, Cunillera N, Rodríguez-Concepción M, Ferrer A, Boronat A, Campos N. Expression and molecular analysis of the *Arabidopsis* DXR gene encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. Plant Physiol 2002;129:1581–91.

Chan W-S, Abdullah OJ, Namasivayam P, Mahmood M. Molecular characterization of a new 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) transcript from Vanda Mimi Palmer. Sci Horticul 2009;121:378-82.

Doran JC. Commercial sources, uses, formation, and biology. In: Boland DJ, Brophy JJ, House APN. *Eucalyptus* leaf oils: use, chemistry, distillation and marketing. Melbourne: Inkata, 1991. p 11-28.

Eldridge K, Davidson J, Hardwiid C, Van Wyk G. Eucalypt domestication and breeding. Oxford: Clarendon, 1994.

Eisenreich W, Schwarz M, Cartayrade A, Arigoni D, Zenk MH, Bacher A. The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. Chem Biol 1998;5:R221–33.

Emanuelsson O, Nielsen H, Von Heijne G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 1999;8:978-84.

Ewing B, Green P. Base-calling of automated sequencer traces using phred, II: Error probabilities. Genome Res 1998;8:186-94a.

Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred, I: Accuracy assessment. Genome Res 1998;8:175-85b.

Feliu JX, Cubarsi R, Villaverde A. Optimized Release of Recombinant Proteins by Ultrasonication of E. coli Cells. Biotechnol Bioeng 1998;58:536-40.

Ganjewala D, Kumar S, Luthra R. An Account of Cloned Genes of Methyl-erythritol-4-phosphate Pathway of Isoprenoid Biosynthesis in Plants. Curr Issues Mol Biol 2008;11:i35-45.

Grolle S, Bringer-Meyer S, Sahm H. Isolation of the *dxr* gene of *Zymomonas mobilis* and characterization of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase. FEMS Microbiol Lett 2000;191:131-37.

Harker M, Bramley PM. Expression of prokaryotic 1-deoxy-D-xylulose-5-phosphatases in *Escherichia coli* increases carotenoid and ubiquinone biosynthesis. FEBS Lett 1999;448:115-19.

Hasunuma T, Takeno S, Hayashi S, Sendai M, Bamba T, Yoshimura S, Tomizawa K, Fukusaki E, Miyake C. Overexpression of 1-Deoxy-*D*-Xylulose-5-Phosphate

Reductoisomerase Gene in Chloroplast Contributes to Increment of Isoprenoid Production. J Biosc Bioen 2008;15:518-26.

Henriksson LM, Björkelid C, Mowbray SL, Unge T. The 1.9 Å Resolution Structure of Mycobacterium Tuberculosis 1-Deoxy-D-Xylulose 5-Phosphate Reductoisomerase, a Potential Drug Target. Acta Crystallogr Sect D Biol Crystallogr 2006;62:807-13.

Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673-80.

Hoeffler JF, Tritsch D, Grosdemange-Billiard C, Rohmer M. Isoprenoid biosynthesis via the methylerythritol phosphate pathway. Mechanistic investigations of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase. Eur J Biochem 2002;269:4446-57.

Jahns, MT. Avaliação da expressão gênica diferencial entre folhas e tecidos vasculares de *Eucalyptus grandis*. Porto Alegre: UFRGS, 2007, 88 p. Dissertation (M. Sc.) – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Jomaa H, Wiesner J, Sanderbrand S, Altincicek B, Weidemeyer C, Hintz M, Türbachova I, Eberl M, Zeidler J, Lichtenthaler HK, Soldati D, Beck E. Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. Science 1999;285:1573-76.

Kirby J, Keasling JD. Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering. Annu Rev Plant Biol 2009;60:335-55.

Koppisch AT, Fox DT, Blagg BS, Poulter CD. E. coli MEP synthase: steady-state kinetic analysis and substrate binding. Biochem 2002;41:236-43.

Kumar S, Tamura K, Jakobsen IB, Nei M. MEGA2: molecular evolutionary genetics analysis software. Bioinform 2001;12:1244-45.

Kuzuyama T, Takahashi S, Takagi M, Seto H. Characterization of 1-deoxy-D-xylulose 5-phosphate reductoisomerase, an enzyme involved in isopentenyl diphosphate biosynthesis, and identification of its catalytic amino acid residues. J Biol Chem 2000;275:19928-32.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;227:680-85.

Lange BM, Croteau R. Isoprenoid biosynthesis via the mevalonate-independent pathway in plants: cloning and heterologous expression of 1-deoxy-D-xylulose 5-phosphate reductoisomerase from peppermint. Arch Biochem Biophys 1999; 365:170–4.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinform, 2007; 23:2947-48.

Leech NL, Barrett KC, Morgan GA. SPSS for intermediate statistics. Use and interpretation, 2° ed. Erlbaum, London, 2005.

Lichtenthaler HK, Schwender J, Disch A, Rohmer M. Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway. FEBS Lett 1997;400:271-74.

Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C}{}_{T}$ Method. Methods 2001;25:402-08.

Mahmoud SS, Croteau RB. Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. Proc Natl Acad Sci USA 2003;98:8915-20.

Miller B, Heuser T, Zimmer WA. *Synechococcus leopoliensis* SAUG 1402-1 operon harboring the 1-deoxyxylulose 5-phosphate synthase gene and two additional open reading frames is functionally involved in the dimethylallyl diphosphate synthesis. FEBS Lett 1999;460:485–90.

Proteau PJ. 1-Deoxy-d-xylulose 5-phosphate reductoisomerase: an overview. Bioorg Chem 2004;32:483–93.

Radykewicz T, Rodich F, Wungsintaweekul J, Herz S, Kis K, Eisenreich W, Bacher A, Zenk MH, Arigoni D. Biosynthesis of terpenoids: 1-deoxy-D-xylulose 5-phosphate reductoisomerase from E. coli is a class B dehydrogenase. FEBS Lett 2000;465:157–60.

Rane MJ, Calvo KC. Reversal of the nucleotide specificity of ketol acid reductoisomerase by site-directed mutagenesis identifies the NADPH binding site. Arch Biochem Biophys 1997;338:83–89.

Sacchettini JC, Poulter CD. Creating isoprenoid diversity. Science **1997;**277:1788–89.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–25.

Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 2003;31:3381-85.

Schwender J, Müller C, Zeidler J, Lichtenthaler HK. Cloning and heterologous expression of a cDNA encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase of *Arabidopsis thaliana*. FEBS Lett 1999;455:140–44.

Seetang-Nun Y, Sharkey TD, Suvachittanont W. Molecular cloning and characterization of two cDNAs encoding 1-deoxy-D-Xylulose 5-phosphate reductoisomerase from *Hevea brasiliensis*. J Plant Physiol 2008;165:991-02

Takahashi S, Kuzuyama T, Watanabe H, Seto H. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-Cmethyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. Proc Natl Acad Sci USA 1998;95:9879–84.

Tournier V, Grat S, Marque C, El Kayal W, Penchel R, De Andrade G, Boudet A-M, Teulières C. An efficient procedure to stably introduce genes into an economically important pulp tree (*Eucalyptus grandis*×*Eucalyptus urophylla*). Transg Res 2003;12: 403-12.

Vitti AMS, Brito JO. Óleo essencial de eucalipto. Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz. Documentos Florestais, n. 17, 2003.

Yan X, Zhang L, Liao WPJ, Zhang Y, Zhang R, Kai G. Molecular characterization and expression of 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) gene from Salvia miltiorrhiza. Acta Physiol Plant 2009;31:1015-22.

Legends of Figures

Figure 1. Full-length cDNA and deduced amino acid sequences of the *E. grandis* 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) gene (*EgDXR*). The start codon (ATG) is underlined in bold and the stop codon (TGA) is underlined in italics and bold. The hypothetical transit sequence cleavage site is indicated by an arrow. The NADPH binding site (doted line) and the proline-rich region (dashed line) are indicated.

Figure 2. Multi-alignment of amino acid sequences of EgDXR and other plant DXRs. Identical amino acids are marked by black background and conserved amino acids are represented in gray background. The aligned DXRs were from *Rauvolfia verticillata* (RvDXR, GenBank accession AAY87151), *Catharanthus roseus* (CaDXR, AAF65154), *Antirrhinum majus* (AmDXR, AAW28998), *Picrorhiza kurrooa* (PkDXR, ABC74566), *Salvia miltiorrhiza* (SmDXR, ACK57535), *Solanum lycopersicum* (SlDXR, AAK96063), *Nicotiana tabacum* (NtDXR, ABH08964), *Hevea brasiliensis* (HbDXR, ABD92702), *Croton stellatopilosus* (CsDXR, ABO38177), *Oryza sativa* (OsDXR, EAY72208), *Zea mays* (ZmDXR, ACG33012), *Camptotheca acuminata* (CaDXR, ABC86579) and *Ginkgo biloba* (GbDXR, AY494186). The main domains are underlined. The transit peptide cleavage site is underlined with a the continuous line, the proline-rich domain is underlined with dashed line, and the NADPH binding motif is double underlined with dashed lines.

Figure 3. Phylogenetic tree analysis of DXRs from plants and bacteria by MEGA version 4.0 from CLUSTAL W alignments. The neighbor-joining method was used to construct the tree. Bootstraps values ≥ 80 after 1000 repetitions are marked with dots. GenBank accessions to plant DXR sequences are indicated in the legend of Fig. 2. Additionally to those, DXR sequences from algae and bacteria were included in the analysis, like *Chlamydomonas reinhardtii* (GenBank accession EDP02894), *Mantoniella squamata* (ACI45953), *Ostreococcus lucimarinus* (ABO95578), *Crocosphaera watsonii* (EAM51359), *Synechococcus elongatus* (ABB57543), *Nostoc punctiforme* (ACC84261), *Porphyra yezoensis* (ACI45960), *Zymomonas mobilis* (AAD29659), *Pseudomonas aeruginosa* (AAF97241), *Vibrio cholerae* (EEO02855), *E. coli* (ACA79101), *Haemophilus influenza* (EEW77116), *Mycobacterium tuberculosis* (CAA98375).

Figure 4. Cartoon representation of a preliminary 3D structure of EgDXR. The model was built with the SWISS-MODEL using the crystal structure of *E. coli* DXR (PDB ID 2EGH) as template, and displayed with the PyMol viewer program (DeLano Scientific LLC.). N-terminal, central catalytic, and C-terminal domains are represented in blue, cyan, and orange colors, respectively. An overlap of EgDXR (orange), *E. coli* (green), *M. tuberculosis* (ruby), and *Z. mobilis* (blue) DXRs structures is displayed at the botton .

Figure 5. SDS-PAGE analysis of protein extracts derived from control and recombinant *E. coli* cultures expressing the GST-*EgDXR* gene fusion. The gel was satined with Comassie Blue. The protein soluble extracts from *E. coli* BL21 (DE3) were run on a 12% (w/v) polyacrylamide gel (30% acrylamide/0.8% bis-acrylamide) under denaturing conditions. Cells were cultured in rotatory shakers (180 rpm) under 30 °C during the different time periods indicated (h). No induction (-) or IPTG induction (+) was applied to cultures as indicated. Cells harboring the pGEX-4T-1 parental vector or the recombinant pGEX-4T-1::*EgDXR* are identified by the plasmid name. The Broad Range Protein Molecular Weight Marker (Promega) was employed in the first lane for protein weight estimation in kDa. The black arrow indicates the band relative to the GST tag and the white arrow indicates the band relative to the GST-EgDXR fusion protein.

Figure 6. Expression analysis of the *EgDXR* gene using microarray DNA hybridization. The signal intensities were normalized by signals obtained for the reference *rnp* gene in leaves. Student´s *t*-test was applied for statistical treatment of the data. Results are expressed as means ± standard errors of signal intensity, considering two biological replicates for leaves and four biological replicates for xylem, in addition to two technical replicates for each replicate (*** $p<0,001$).

Figures

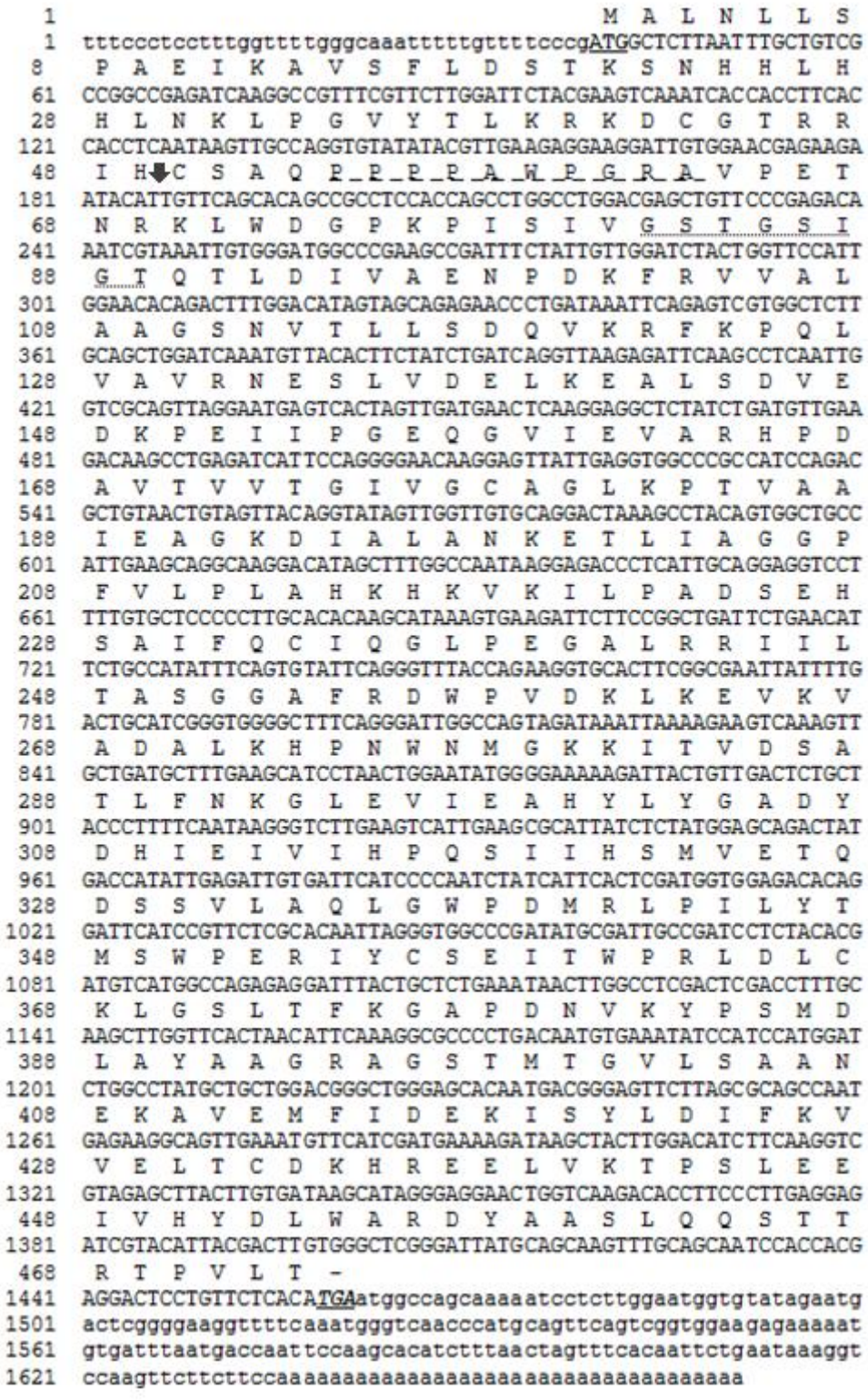


Figure 1.

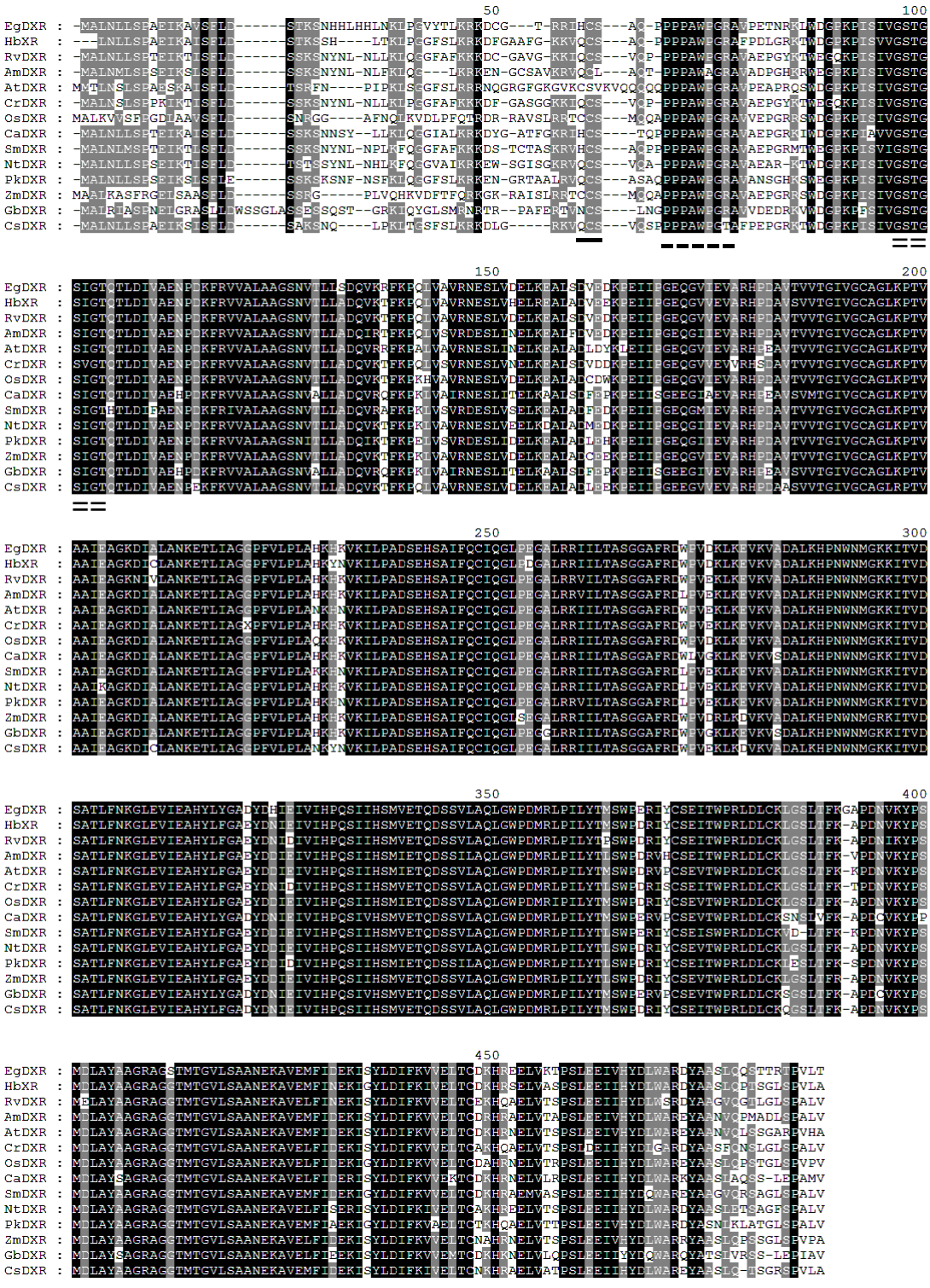

EgDXR
HbXR
RvDXR
AmDXR
AtDXR
CrDXR
OsDXR
CaDXR
SmDXR
NtDXR
PkDXR
ZmDXR
GbDXR
CsDXR

Figure 2.

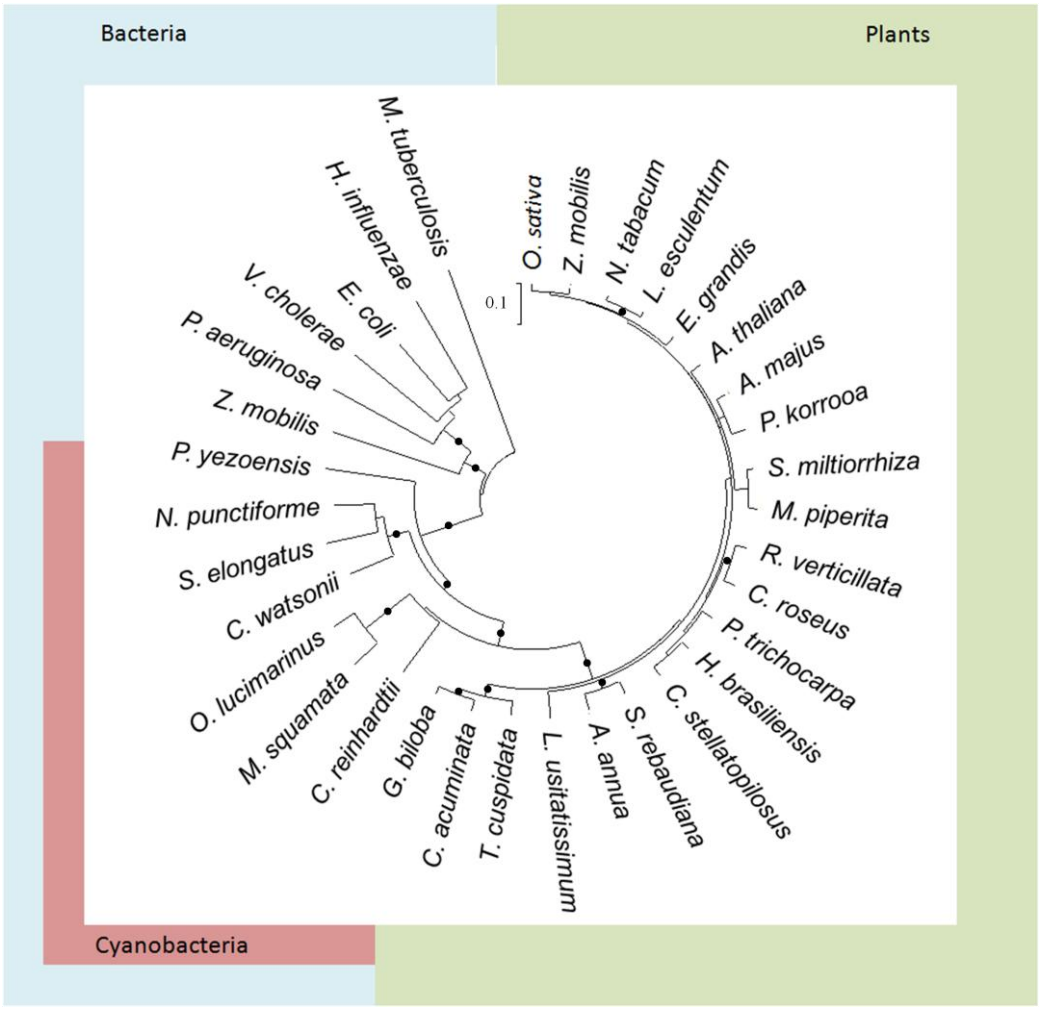


Figure 3.

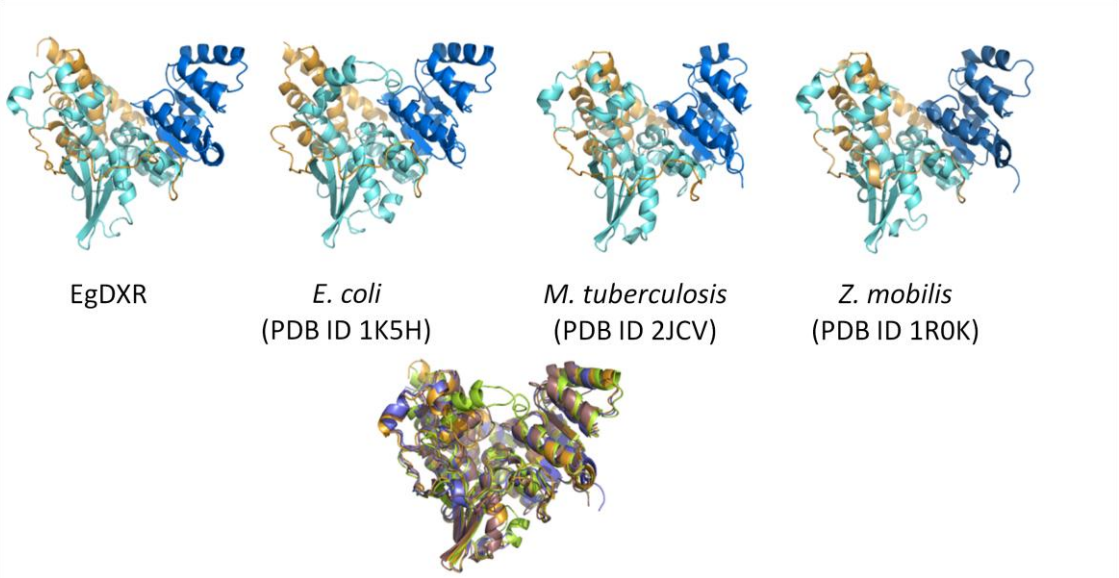


Figure 4.

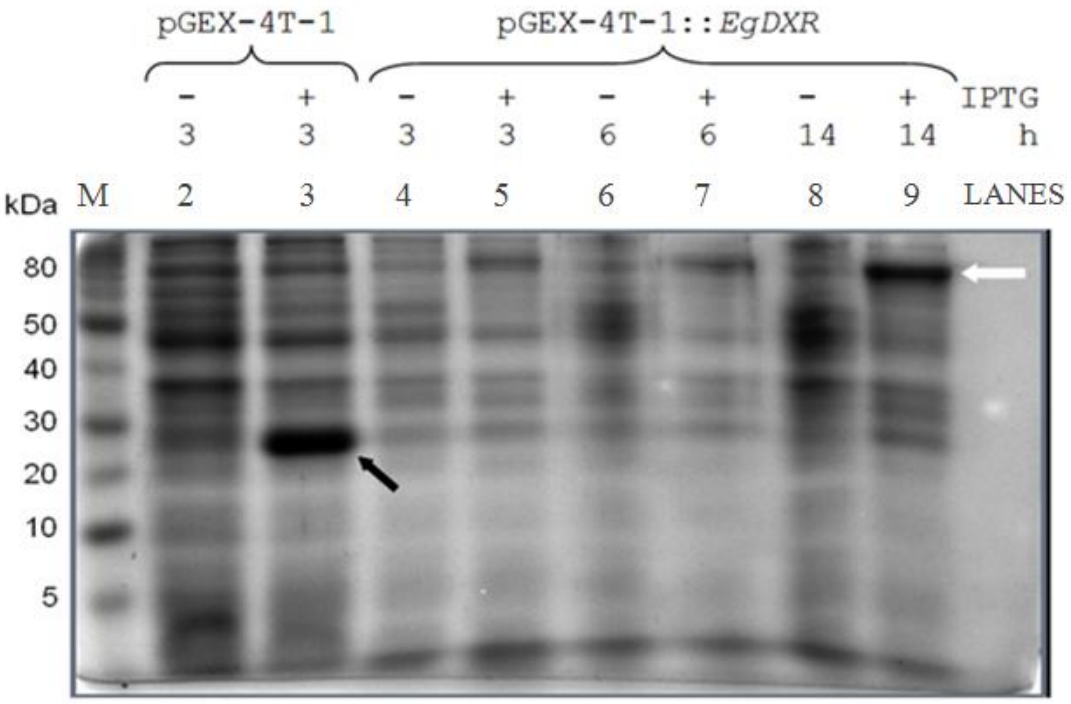


Figure 5.

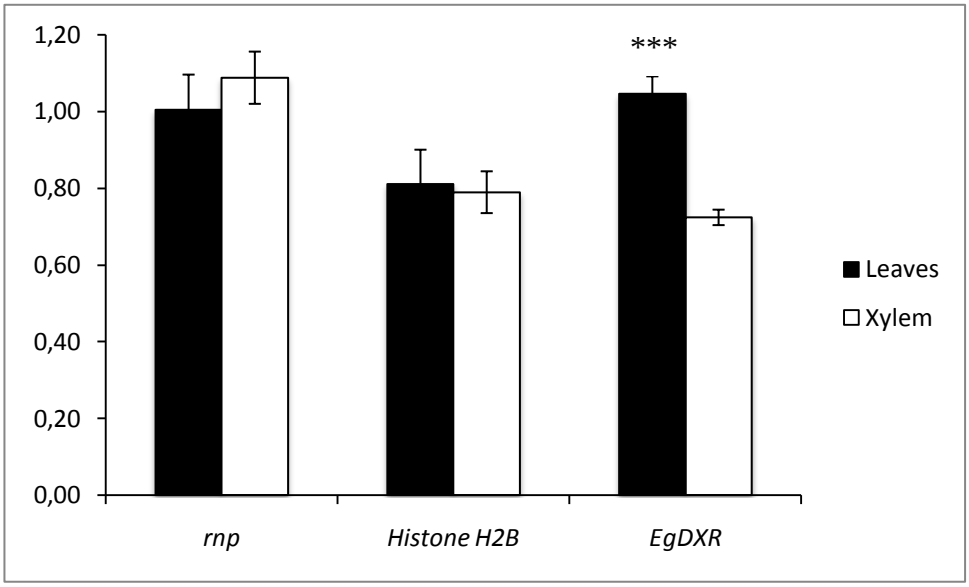


Figure 6.

CAPÍTULO III

CLONING, CHARACTERIZATION AND HETEROLOGOUS EXPRESSION OF ISOPENTENYL DIPHOSPHATE ISOMERASES FROM *Eucalyptus grandis* (EgIPPI1 AND EgIPPI2), AND MOLECULAR DYNAMICS OF AN IPPI1 PROTEIN.

Manuscrito de artigo científico redigido conforme normas do periódico
Journal of Plant Physiology

Running title: Cloning, characterization, expression of two isopentenyl diphosphate isomerases from *Eucalyptus grandis* and molecular dynamics an isopentenyl diphosphate isomerase 1 protein.

Corresponding author: Pasquali G.

Mailing address: Av. Bento Gonçalves, 9500 - Building 43431 – Lab 212 Campus do Vale/UFRGS 91501-970 - Porto Alegre, RS - Brazil – Postal-box: 15005

Fone: 55-51-3308.9410 Fax: 55-51-3308.7309

E-*mail address*: pasquali@cbiot.ufrgs.br

Cloning, characterization, expression of two isopentenyl diphosphate isomerases from *Eucalyptus grandis* (EgIPPI1 and EgIPPI2), and molecular dynamics an IPPI1 protein.

Genaro Azambuja Athaydes[a], Edson Fauth Vargas[a], Hugo Verli[a] and Giancarlo Pasquali[a]*

* Corresponding author.

[a]Biotechnology Center, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil.

Summary

Isopentenyl diphosphate (IPP) and its isomer, dimethylallyl diphosphate (DMAPP), are the precursors of all isoprenoids, the broadest and most chemically complex class of natural chemicals. Since isoprenoids are used as drugs, flavors, fragrances, cosmetics, pigments, insecticides and others, they have many quite important social impacts. An important step in the isoprenoid production is the isomerization of IPP to DMAPP, catalyzed by isopentenyl diphosphate isomerase (IPPI, EC: 5.3.3.2). The two main routes for the production of IPP and DMAPP are well known. The classic mevalonate-(MVA-)dependent pathway has been known for 6 decades, but the plastidic 2C-methyl-D-erythritol-4-phosphate (MEP) pathway was discovered in the 1990´s. The MVA pathway gene encoding hydroximethyl glutaryl-coenzyme A reductase was used as target for developing lowering cholesterol drugs like statins. Likewise, MEP pathway genes are considerated good targets for developing antibiotics and other drugs. Indeed, the metabolic engineering of isoprenoid biosynthetic genes can yield compounds of commercial interest both in bacteria and plants. Specifically for *Eucalyptus grandis*, these genes can allow the improvement of essential oil production. As a key step toward that goal, we cloned two IPPI genes from *E. grandis* (*EgIPPI1* and *EgIPPI2*) in this work. The proteins EgIPPI1 and 2 operate in a subsequent step after production of IPP by MVA and MEP pathways. Here we describe their molecular characterization, the heterologous expression of both proteins, and a comparative mRNA expression analysis. Finally, we developed the first simulation study of the molecular dynamics of an IPPI1 enzyme to clarify specific structural properties concerning the N-terminal protein region, only present in eukaryotic IPPIs.

Keywords: Isopentenyl diphosphate isomerase, IPPI1, IPPI2, molecular dynamics, *Eucalyptus grandis*.

Abbreviations: DXS, 1-deoxy-D-xylulose-5-phosphate sintase; EST, *expressed sequence tag;* HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA reductase; IPPI, isopentenyl diphosphate isomerase; MDC, mevalonate-5-diphosphate decarboxylase; MEP, 2C-methyl-D-erythritol-4-phosphate; MVA, mevalonate; ORF, *open reading frame*; *rnp*, ribonucleoprotein L23A gene; RMSD, root mean square deviation; RMSF, root mean square fluctuation.

Introduction

Isoprenoids, the largest class of natural products with unmatched chemical complexity, play essential roles in basic microbial and plant processes, including photosynthesis, growth, development, and reproduction. However, the broad diversity of isoprenoid compounds is generated on plant secondary metabolic pathways related to defense and adaptation to environmental conditions (Rodríguez-Concepción and Boronat, 2002; Philips *et al.*, 2008). Isoprenoids became an obligatory part of the human life since they are employed as drugs, flavors, fragrances, cosmetics, dyes, insecticides, and others (Dixon, 2001; Harborne, 2001; Verpoorte *et al.*, 2002)

Plants generally store low concentrations of secondary metabolites, thus the extraction of metabolites from them is problematic and the purification of desired compounds is quite hard to achieve because it requires the separation from many other similar molecules. Furthermore, the yield is often directly related to local and climate-determining factors (Chemler *et al.*, 2008). The chemical synthesis of isoprenoids is also generally difficult to achieve and has inherent high costs given their high molecular complexity (Harada *et al.*, 2009). Since isoprenoids have high commercial value, there are great efforts to produce them by bioengineering both in bacteria and plants (Xiao *et al.*, 2009).

All isoprenoids are synthesized via the condensation of the five-carbon universal isoprenoid precursors, isopentenyl diphosphate (IPP) and its allylic isomer dimethylallyl diphosphate (DMAPP; McGarvey and Croteau, 1995). In higher plants, two independent pathways localized in separate intracellular compartments are involved in the biosynthesis of IPP and DMAPP (Rodriguez-Concepcion *et al.*, 2006). The classic route, mevalonate- (MVA-) dependent pathway, is citosolic and the more recently discovered 2C-methyl-D-erythritol 4-phosphate (MEP) pathway is located in plastids. The MVA pathway is composed of five catalytic steps, starting with the condensation of acetyl-CoA and finishing with the formation of IPP (Pang *et al.*, 2006). In the MEP pathway, glyceraldehyde 3-phosphate and pyruvate are converted into IPP and DMAPP through seven catalytic steps in a 6:1 rate (Rohmer *et al.*, 1996; Sprenger *et al.*, 1997; Lois *et al.*, 1998).

These pathways are distributed throughout nature, and generally the MVA pathway is prevalent in eukaryotes and archaea, whereas the MEP pathway is widespread in eubacteria. However, there are many exceptions to this pattern; for

example, several eubacteria are known to utilize the MVA pathway instead of the MEP pathway, with some species carrying genes from both pathways; on the other hand, the protozoan parasite *Plasmodium falciparum* relies only on the MEP pathway (Eisenreich *et al.*, 2004).

IPP isomerase (IPPI) reversibly catalyzes the isomerization of IPP and DMAPP maintaining the equilibrium between these isomers. Two types of IPPIs are known: IPPIs of type 1 are compact globular proteins that belong to the class of α/β proteins found in eukarya and bacteria (de Ruyck *et al.*, 2008). This class of enzymes requires a divalent metal ion for catalysis through an antarafacial protonation-deprotonation mechanism in which a highly reactive carbocationic intermediate is formed (Thibodeaux *et al.*, 2008); the type 2-IPPI monomer consists of a regular TIM barrel (α8β8 barrel) domain and requires flavin mononucleotide (FMN), a reduced nicotinamide adenine dinucleotide cofactor (either NADH or NADPH), and divalent metal ions (Mg^{2+}) for catalysis (Kaneda *et al.*, 2001; de Ruyck *et al.*, 2005).

The mechanism of substrate entrance into the active poket related to human IPPI-1 (hIPPI-1) and the conformational behavior of the N-terminus extension present only in eucariotic IPPIs during this process is not clear to date (Zhang *et al*., 2007; Zheng *et al*., 2007). In one hypothesis, the N-terminal helix of hIPPI1, which is not present in EcIPPI1, may be fully unfolded until the formation of the substrate complex (Zhang *et al*., 2007). However, in another work this unfolded state was not observed and a different was suggested as the region of substrate entrance in hIPPI1 when compared with EcIPPI1 (Zheng *et al*., 2007).

In this work, we describe the cloning of complete cDNA sequences for two *Eucalyptus grandis* IPPI, named *EgIPPI1* and *EgIPPI2*. The cDNA structure of both genes and main protein properties assumed from the deduced aminoacid sequences are described as well as the patterns of gene expression in mature leaves and vascular tissues of adult *E. grandis* trees. Both proteins were heterologously produced in *Escherichia coli* cells. Also, we performed structural bioinformatics' analysis with an IPPI1 protein, the *Homo sapiens* IPPI1 (hIPPI1). To the best of our knowledge, this is the first work of molecular dynamics (MD) analysis of an IPPI enzyme.

Material and methods

cDNA Library construction and sequencing

Seeds of *E. grandis* (supplied by Aracruz Cellulose S.A., Guaíba, RS, Brazil) were surface sterilized by sequentially soaking in 70% ethanol for 2 min, 1% (v/v) active chlorine solution for 15 min and five times in sterilized water. Seeds were then placed in culture flasks (5 cm in diameter, 8 cm high) containing the Murashige & Skoog (MS, Invitrogen) complete medium solidified by 0.7% (w/v) Phytoagar (Duchefa) and left to germinate in the dark for two days at 26±2 °C. After germination, flasks were transferred to a culture room with a 16 h photoperiod at the same temperature. Total RNA was extracted from 60-days old *E. grandis* seedlings using the PureLink Plant RNA Purification Reagent (Invitrogen) according to the manufacturer's instructions for multi RNA minipreps. Messenger RNA was extracted by the employment of the Oligotex mRNA Purification System (Qiagen). A total amount of 5 μg mRNA was finally used for cDNA library construction by the Superscript Plasmid System with Gateway Technology for cDNA Synthesis and Cloning Kit (Invitrogen).

Plasmid DNA preparation was carried out in 96-well microplates using standard methods based on alkaline lysis and filtration in Millipore filter plates. Plasmid samples were sequenced using the automatic sequencer ABI-PRISM 3100 Genetic Analyzer armed with 50 cm capillaries and POP6 polymer (Applied Biosystems). DNA templates (30 to 45 ng) were labeled with 3.2 pmol of primer GlyptsRev1 (5'-ATAGGGAAAGCTGGTACGC-3') or M13 -40 forward (5'-GTTTTCCCAGTCACGACGTTGTA-3', Amersham Biosciences) and 2 μL of BigDye Terminator v3.1 Cycle Sequencing RR-100 (Applied Biosystems) in a final volume of 10 μL. Labeling reactions were performed in a GeneAmp PCR System 9700 (Applied Biosystems) termocycler with a initial denaturing step of 96 ºC for 3 min followed by 25 cycles of 96 ºC for 10 sec, 55 ºC for 5 sec and 60 ºC for 4 min. Labeled samples were purified by isopropanol precipitation followed by 70% ethanol rinsing. Precipitated products were suspended in 10 μL formamide, denatured at 95 ºC for 5 min, ice-cooled for 5 min and electroinjected in the automatic sequencer. Sequencing data were collected using the software Data Collection v1.0.1 (Applied Biosystems) Data generated were processed by a suite of programs available at the Universidade

Católica de Brasília (http://www.ucb.br/genolyptus) prior to assembly with the PHRED (Ewing, 1998a; b) and PHRAP (http://phrap.org/) algorithms. Accepted ESTs were selected based on a minimal length of 250 bp with every base having a quality of PHRED higher than 20.

Heterologous expression *EgIPPI1* and *EgIPPI2*

For expression in *Escherichia coli* BL21(DE3), *EgIPPI1* and *EgIPPI2* cDNAs were transfered from pSPORT1 into pSK^+ Bluescript (Stratagene) by *Bam*HI and *Eco*RI digestion and resequenced. EgIPPI1 and 2 cDNAs were then cloned into pGEX-4T-3 and pGEX-4T-2 (GE Health care) respectively by *Sal*I and *Not*I digestion. Recombinant *E. coli* cells harboring pGEX-4T-3-*EgIPPI1* and pGEX-4T-2-*EgIPPI2* were cultivated at 37 °C overnight and inoculated into 500-mL erlenmeyer flasks containing a 50ml of Luria–Bertani (LB) medium supplemented with 100 µg/mL of ampicillin. When an OD600 of about 0.45 was reached, the expression was induced by adding isopropyl β-D-thiogalactopyranoside (IPTG) at 100 µg/mL. Cultures were kept for addicional 14h at 30 °C. For each culture, a non-induced control culture was kept under the same condictions. Recombinant and control *E. coli* cells were harvested by centrifugation and then disrupted by ultrasonication (Feliu *et al.*, 1998). A soluble protein fraction was finally obtained after centrifugation at 12,000×g for 20min for SDS-PAGE assay.

Protein electrophoresis

SDS-PAGE were performed essentially as described by Laemmli (1970). SDS-PAGE was performed on a Mini-PROTEIN^{TM} electrophoresis system (Bio-Rad) with 12% polyacrylamide gels (30% acrylamide/0.8% bis-acrylamide) under denaturing conditions.

Bioinformatics' analysis of DNA and protein sequences

Comparative and bioinformatic analysis of EgIPPI1 and EgIPPI2 were carried out with online tools available at the *National Center for Biotechnology Information* (NCBI; http://www.ncbi.nlm.nih.gov), Expasy (http://cn.expasy.org), and Center for Biological Sequence Analysis (http://www.cbs.dtu.dk/services/ChloroP/, and

http://www.cbs.dtu.dk/services/TargetP/). Nucleotide sequences, deduced amino acid sequences and *open reading frames* (ORF) were analyzed, and sequences comparisons were carried out through by database search using the BLAST program (NCBI). The phylogenetic analysis of EgIPPI1 and 2 and IPPIs from other plant species was performed with CLUSTAL W version 2.0 (Larkin *et al.*, 2007) using default parameters. Phylogenetic trees were constructed using MEGA version 4.0 (Kumar *et al.*, 2001) from the resulting CLUSTAL W alignments. The neighbor-joining method (Saitou and Nei, 1987) was used to construct the trees.

Microarray Analysis

To investigate the expression pattern of important genes from MEP and MVA pathways and the EgIPPI1 and 2 genes in leaves and xylem, 20 µg of each RNA sample from leaves and xylem of *E. grandis* adult trees were dried down in speed-vac and sent to NimbleGen Systems Inc. (Reykjavik, Iceland). Nine 50-mer probes were designed and synthesized in duplicates from the 21.442 ESTs of the GENOLYPTUS Project at NimbleGen. Synthesis of Cy3-labeled cDNAs, hybridizations, washings, scanning and preliminary analysis were carried out by NimbleGen, following a standard expression design. Normalized gene expression values for each feature were obtained taking into account the hybridization values for all 9 probes per gene, in duplicates. The final analysis were performed using MultiExperiment Viewer (MEV), developed at The Institute for Genomic Research (TIGR). Data normalization was performed using the $2^{-\Delta\Delta Ct}$ method (Livak *et. al.*, 2001) adapted to microarray data and a Student´s *t-test* was performed with the Statistical Package for the Social Sciences Software (SPSS 12.0 for Windows; Leech *et al.*, 2005). Relative expression of the genes in leaves and xylem was calibrated with an internal reference-gene encoding the ribonucleoprotein L23A. Detailed technical descriptions of the procedure are referred by Bastolla (2007), Jahns (2007) and Pasquali *et al.* (in preparation).

Molecular dynamics simulations

The recommendations and symbols of nomenclature employed in the MD analysis were those proposed by the International Union of Pure and Applied Chemistry (IUPAC-IUB, 1970). MD calculations and analyses were performed using the GROMACS

package (van der Spoel *et al*., 2008) and GROMOS96 43a1 force field (van Gunsteren *et al*., 1996). The manipulation of structures was carried out in the VMD environment (Humphrey *et al*., 1996) and Pymol (Delano *et al*., 2002), while comparative modelling of proteins was done with the Swiss PDB viewer (Guex and Peitsch, 1997). The monitoring of secondary structure content was performed with PROCHECK (Laskowski *et al*., 1993).

In this experiment, we started from 2I6K crystal. We removed the ligands Mn^{2+} and DMAPP to let the protein free during 100 ns of MD simulation. The human IPPI1 (PDB ID 2I6K) structure was obtained from the Protein Data Bank. The oligomeric state was divided for simulations of the monomeric forms, with the A chains being employed in the calculations. The protein was solvated in a rectangular box using periodic boundary conditions and the SPC water model (Berendsen *et al*., 1981). Counterions (Na^+ and Cl^-) were added to neutralize the systems. The employed MD protocol was based on previous studies, as previously described (de Groot and Grubmüller, 2001; Verli and Guimarães, 2005). The Lincs method (Hess *et al*., 1997) was applied to constrain covalent bond lengths, allowing an integration step of 2.0 fs after an initial energy minimization using the Steepest Descent algorithm. Electrostatic interactions were calculated with the Particle Mesh Ewald method (Darden *et al*., 1993). Temperature and pressure were kept constant by coupling proteins, ions, and solvent to external temperature and pressure baths with coupling constants of τ = 0.1 and 0.5 ps (Berendsen *et al*., 1984), respectively. The dielectric constant was treated as ε = 1, and the reference temperature adjusted to 310 K. The system was slowly heated from 50 to 310 K, in steps of 5 ps, each one increasing the reference temperature by 50 K, allowing a progressive thermalization of the molecular system. The simulation was then performed to 100 ns, without any restraint (van der Spoel *et al*., 2005).

Results and Discussion

Cloning of *EgIPPI1* and *EgIPPI2* full-length cDNAs

Based on the conserved regions of known plant IPPI1 and IPPI2 sequences, we were able to identify pSPORT1 plasmid clones harboring the potentially complete IPPI1 and IPPI2 cDNAs from *E. grandis* at the transcriptome of the Genolyptus Project. One single contig was found for each *EgIPPI* in the transcriptome database. *EgIPPI1* and *EgIPPI2* had contigs composed respectively by 11 and 10 reads all from *E. grandis* libraries. Since we did not find any other kind of cDNA from *ippi*s genes in the libraries, it is possible that these genes are present in one copy in the *E. grandis* genome. The cDNA inserts with approximately 1,200 and 1,000 bp respectively were subcloned into SK^+ pBluescript (Stratagene) phagemids and resequenced. Final sequences exhibited 1,137 bp for *EgIPPI1* (Fig. 1) and 999 bp for *EgIPPI2* (Fig. 2). Using the ORF Finding tool available at the NCBI and the ProtParam tool from the Swiss Institute of Bioinformatics, we could observe that *EgIPPI1* contains an 882 bp ORF encoding a peptide of 293 amino acids (Fig. 1), with a calculated molecular mass of 33.5 kDa, and an isoelectric point of 6.16. The same analysis was performed with *EgIPPI2* revealing an ORF of 723 bp encoding 232 amino acids (Fig. 2), with a calculated molecular mass of 26.5 kDa, and an isoeletric point of 5.21. Both sequences exhibit typical poly(A) signals and poly(A) tails at conserved regions in the cDNA (Nag *et al.*, 2006), underlined in Figs. 1 and 2.

The deduced amino acid sequences of EgIPPI1 and EgIPPI2 were submitted to the NCBI for PSI-BLAST searching and the result allowed us to conclude that EgIPPI1 and EgIPPI2 have high homologies with known IPPI sequences from other plant species. EgIPPI1 is highly similar to IPPI1 from *Nicotiana tabacum*, *Catharanthus roseus* and *Zea mays*, with 82% identity in amino acid composition, and also with *Melaleuca alternifolia* (77%) and *Arabidopsis thaliana* (76%). High similarity was also found between EgIPPI2 and IPPI2 proteins from other plants, with 94% identity with the *N. tabacum* IPPI2, 88% identity with the protein from *M. alternifolia*, and 81% with *Z. mays*.

The amino acid sequences of nine IPPIs available at the GenBank were aligned with EgIPP1 and EgIPP2 according to the algorithm of CLUSTAL W (Higgins *et al.*, 1994; Fig. 3). The results showed that all 11 IPPIs have high similarity throughout the

entire coding region. The Nudix domain, a highly conserved 23-residue block present in all IPPI sequences was also found in both EgIPPI1 and EgIPPI2 (Fig. 3). In the multiple alignment analysis, it was clear that EgIPPIs exhibit almost the same amino acid composition in the substrate binding motifs when compared to the other nine plant IPPIs. These results are indicative that EgIPPIs are members of the Nudix hydrolase superfamily that includes the IPP-binding protein family.

Philips *et al.* (2008) reported that the two IPPIs from *Arabidopsis* (AtIPPI1 and AtIPPI2) are differentially compartimentalized. In their work, it was not only suggested that AtIPPI1 is targeted to plastids and AtIPPI2 to mitochondria, but also that additional IPPI activities encoded by both genes are also present in the cytoplasm. The authors concluded that each IPPI plays distinct roles in isoprenoid metabolism (Philips *et al.*, 2008). Differently from *Arabidopsis* IPPIs, when analysed with ChloroP and TargetP prediction programs (Emanuelsson *et al.*, 1999), no transit peptide cleavage site could be clearly found in any of the IPPI sequences from *E. grandis.* These results suggest that the plastid EgIPPI1 protein is possibly active with the N-terminus extension. The EgIPPI2 does not have the signal-peptide, so it seems to be directed to the cytosol. It was previously suggested that the source of cytosolic and plastid forms of IPPIs are possibly derived from the alternative translation from two or more of the AUG start codons well-conserved at the 5'-terminal region of the mRNAs (see Figs. 1, 2 and 3, square-marked methionines) that precede the portion of the polypeptide beyond residue 75 in the alignment (Cunningham *et al.*, 2000; Fig. 3). The use of alternative Met for the initiation of translation is one of several strategies employed by plants to produce polypeptides needed in more than one compartment (Small *et al.*, 1998). However, differently of *Eg*IPPI1, the *EgIPPI2* sequence does not have an alternative Met for translation and it seems to be directed to the cytosol.

Since there are at least two IPPI isoforms in plants (Cunningham *et al.*, 2000) and given that *EgIPPI1* and *EgIPPI2* are the first IPPI genes cloned from a *Eucalyptus* species, it would be interesting to investigate their evolutionary position among the phylogenetic tree of various IPPIs. A phylogenetic tree was constructed using the program MEGA and CLUSTAL W alignment results based on the deduced amino acid sequences of *EgIPPIs* and other IPPIs from different organisms including plants, mammals, fungi and bacteria. The resulting tree made a clear distinction among the IPPIs from plants, fungi, mammals and bacteria.

As it can be observed in Fig. 4, EgIPPI1 is positioned among sequences of the plant IPPI group, with closest relationship with IPPI1 from *M. alternifolia* (Fig. 4). EgIPPI2 exhibited closest relationship with IPPI2 from *N. tabacum* as can be observed in Fig. 4. The relatively high DNA and amino acid sequence similarity and lack of any obvious dichotomy in grouping do not support the concept of distinct ancestral genes or an ancient gene duplication giving rise to contemporary genes encoding plastid, cytosolic, and other IPPI isoforms in plants. Therefore, these genes may share the same ancestral gene origin or may be fruit of recent gene duplication.

Expression of *EgIPPI1* and *EgIPPI2* genes in *Escherichia coli*

In order to start the biochemical characterization of the enzymes encoded by *EgIPPI1* and *EgIPPI2*, the corresponding genes were cloned into the pGEX-4T-3 and pGEX-4T-2 (GE Healthcare) expression vectors, respectively, and transformed into *E. coli* BL21 (DE3). The enzymes, coupled N-terminally to a gluthatione S-transferase (GST) tag, were mostly produced after IPTG induction in cultures maintained for 14 h at 30 °C. The soluble fractions of crude cell extracts were used in SDS-PAGE assays, as shown in Fig. 5. The molecular mass of the GST-EgIPPI1 protein was estimated from the deduced amino acid composition as being of approximately 59 kDa (23 kDa). A value of 55 kDa was estimated for the GST-EgIPPI2 fusion protein. The calculated molecular masses of both EgIPPIs were in good agreement with those observed by SDS-PAGE (33.5 kDa for EgIPPI1 and 26.5 kDa for EgIPPI2), as shown in Fig. 5.

Microarray Gene Expression Analysis

In order to investigate the gene expression profiles in leaf and vascular tissues of *E. grandis*, previous results from microarray analysis (Pasquali *et al.*, in preparation) were evaluated for important genes in the MEP and MVA pathways and also for the genes of interest, *EgIPPI1* and *EgIPPI2*. The genes included the 1-desoxy-D-xylulose 5-phospohate synthase (DXS, EC: 4.1.3.37) gene from MEP pathway, and the gene mevalonate-5-diphosphate decarboxylase (MDC, 4.1.1.33) from MVA pathway. To do so, two genes known to be constitutively expressed in xylem and leaves of *Eucalyptus* (Bastolla, 2007; Jahns, 2007), ribonucleoprotein L23A (*rnp*) and histone H2B were employed as house-keeping reference genes. The microarray signal intensities of all

samples were normalized with the *rnp* gene signal acquired from cDNAs derived from leaf tissues. As shown in Fig. 6, this analysis revelated that the gene EgIPPI1 is significantly more expressed in leaves than xylem of *E. grandis* (Fig. 4). Philips *et al.* (2008) reported a differential expression of IPPIs genes from *Arabdopisis thaliana* (*ATIPPI1* and *ATIPPI2*). *ATIPPI1* and *ATIPPI2* were mainly expressed in flowers and roots respectively, revealing that these genes are not redundant and that each one may play different roles in the production of IPP and DMAPP pools in plant cells. Leaf is the tissue where monoterpenes and diterpenes, including essential oils, are produced through the MEP pathway in *E. grandis*. Thus, our microarray results suggest that these genes may play an important role to the metabolism of these isoprenoids in this tissue. In the MVA pathway, the key regulatory enzyme 3-hydroxy-3-metylglutaryl-CoA reductase (HMG-CoA reductase, EC: 2.3.3.10) was extensively studied, but other enzymes less studied are equally important to this route (de Ruyck *et al.*, 2008), including MDC. Here, we report a high expression level of MDC gene in xylem tissue when compared to leaf tissue. This result thus may indicate an important influence of MDC gene in the metabolism of isoprenoids in the vascular tissue.

Molecular dinamics' analysis of an IPPI1

Different proposals to the mechanism of substrate entrance appear to be related to differences on hIPPI1 crystals (PDB IDs 2DHO, 2I6K, 2ICJ and 2ICK). In one of them (PDB ID 2DHO), the protein was bound to Mn^{2+} cofactor but unbound to any of the substrates. In this crystal, the N-terminal helix of the enzyme, a region not present on eucariotic IPPIs, seems to be unfolded. Therefore, it was suggested that this region act as a flap, in which the N-terminus is unfolded when the enzyme lies in its substrate non-bounded form, and so giving access to the catalytic pocket (Zhang *et al.*, 2007). However, in another work a crystal of hIPPI1 also bound to its cofactor Mn^{2+} was solved in its substrate bound form (PDB ID 2ICK). In this work, the authors suggested that the substrate entrance is blocked by sidechains of Gln14 and Leu17 (Gln 13 and Leu16 in the 2I6K crystal) and it was also proposed a relative rigidity in the N-terminus helix which lies just on top of the cavity (Zheng *et al.*, 2007; Fig. 7 B). Analysis of potential electrostatic distribution and position of aminoacid lateral chains lead the authors to propose that the region for substrate entrance is slightly shifted when compared to *E. coli* IPPI1 (EcIPPI1) substrate entrance (Zheng *et al.*, 2007). With the

purpose of getting further insights on the role played by the N-terminal region of eucaryotic IPPIs, we performed a molecular dynamic (MD) simulation of the enzyme in its substrate non-bound state.

In order to monitor the progress of structural conformational changes in hIPPI1 MD, and check flexibility and potential conformational changes, we evaluated the root mean square deviation (RMSD) of the simulated protein. Considering the entire protein, the hIPPI1 molecule achived a plateau around 0.37 nm (Fig. 8 A), suggesting that the protein is relatively stable with no expressive conformational changes. The multi-alignment of secondary and terciary structures obtained in 10 ns intervals reinforces this stability in the N-terminus helix (Fig. 7 B; Fig. 8 C).

While the data presented in RMSD plot gives a global perspective of the protein conformational modifications over the MD trajectory, it lacks resolution at a residue level. So, in order to gain further insights into the conformational behavior of the simulated proteins, we employed a strategy that describes the structural fluctuation of a protein structure as a function of both time and residue number (Verli *et al*., 2005). Such root mean square fluctuation (RMSF) analysis confirmed a relative stability of the protein throughout the trajectory of 100 ns including the N-terminal region (Fig. 8 B).

Our results show that the N-terminal region was stable during the entire MD simulation, thus reinforcing the theory proposed by Zheng *et al*. (2008). It was previously suggested that the residues Lys38 and Ile76 of 2ICK crystal (Lys37 and Ile75 in 2I6K crystal) would be in an appropriate position to substrate entrance and also, that the region around these residues was positively charged and so facilitating substrate entrance (Zheng *et al.*, 2008). Although a quantitative analysis must be employed for confirmation, the position of lateral chains of Lys37 and Ile75 seemed to oscilate through the 100 ns of MD, but visually no impediment was detected (Fig. 8 B).

In summary, two new isopentenyl diphosphate isomerase genes (*EgIPPI1* and *2*) were cloned from *Eucalyptus grandis*. These EgIPPIs resembles other IPPIs of plant origin and has the Cys149-Glu212 active site residues as well the nudix domain characteristic of all IPPIs. The EgIPPI genes encode two enzymes that could be used to forward the metabolic flux to the biosynthesis of secondary metabolites through metabolic engineering of plants or bacteria. The cloning and characterization of *EgIPPI1* and *2* will be helpful to understand the biosynthesis and metabolic engineering of valuable isoprenoids in *E. grandis*.

The structural diversity of isoprenoids is immense. Terpenoids play important roles in all living organisms. In light of the many diverse functions of terpenes, it is not surprising that their biosynthesis has been subject of intense investigation. IPPIs have been studied for five decades, and the mechanism of the catalysis is now well understood. However, the structural function played by typical eucariotic N-terminal extension region of IPPIs is not yet well understood. In this work, we performed MD simulations for evaluation of the protein N-terminal region in solution. Since the N-terminus of the protein remained stable during the 100 ns of MD, our results corroborate the mechanism proposed by Zheng *et al.* (2008) in which the region of substrate entrance is slightly shifted when compared to the *E. coli* IPPI1 enzyme.

Aknowledgments

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) MCT, Brasília, Brazil; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Ministério da Educação), MEC, Brasília, Brazil; Financiadora de Estudos e Projetos (FINEP) along with the projects "Genolyptus – Rede Brasileira de Pesquisa do Genoma do *Eucalyptus*" and "WoodGenes - Gênese de Madeira em *Eucalyptus*: Genes, Funções, Regulação e Expressão Transgênica" through the Fundos Setoriais Verde-Amarelo e de Biotecnologia do Ministério da Ciência e Tecnologia (MCT) and paper and cellulose companies including Aracruz Celulose S.A., Grupo Raiz, Celulose Nipo-Brasileira S.A. – CENIBRA, International Paper do Brasil Ltda., Jarí Celulose S.A., Klabin S.A., Lwarcel Celulose e Papel Ltda., Veracel Celulose S.A., Votorantim Celulose e Papel S.A., Zanini Florestal S.A. e Suzano-Bahia Sul Papel. The authors are also grateful to the Programa de Pós-Graduação em Biologia Celular e Molecular/UFRGS (Brazil).

References

Anita N, Kazim N, Amir K, Harold GM. The conserved AAUAAA hexamer of the poly(A) signal can act alone to trigger a stable decrease in RNA polymerase II transcription velocity. RNA 2006;12:1534-44.

Bastolla FM. Seleção e avaliação de genes de referência para estudos de expressão gênica em *Eucalyptus*. Porto Alegre: UFRGS, 2007, 97 p. Dissertation (M.Sc.) – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Becker CF, Guimarães JA, Verli H. Molecular dynamics and atomic charge calculations in the study of heparin conformation in aqueous solution. Carbohydr Res 2005;340:1499-07.

Berendsen HJC, Postman JPM.; Van Gunsteren W, Hermans J. Interaction models for water in relation to protein hydration, in: B. Pullman (Ed.), Intermolecular forces, Reidel, Dordrecht, The Netherlands, 1981. p. 331-42.

Berendsen HJC, Postman JPM, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984; 81: 3684-90.

Chemler JA, Koffas MAG. Metabolic engineering for plant natural product biosynthesis in microbes. Curr Opin Biot 2008;19:597-05.

Darden T, York D, Pedersen L. Particle Mesh Ewald – an N log(N) method for Ewald sums in large systems. J Chem Phys, 1993;98:10089-92.

De Groot BL, Grubmüller H. Water Permeation Across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and GlpF. Science, 2001;294:2353-57.

Delano WL. The Pymol Molecular Graphics System, Delano Scientific, Sao Carlos, CA, USA, http://www.pymol.org, 2002.

De Ruyck J, Pouyez J, Rothman SC, Poulter D, Wouters J. Crystal Structure of Type 2 Isopentenyl Diphosphate Isomerase from *Thermus thermophilus* in Complex with Inorganic Pyrophosphate. Biochem 2008;47:9051-53.

De Ruyck J; Rothman SC, Poulter CD, Wouters J. Structure of Thermus thermophilus type 2 isopentenyl diphosphate isomerase inferred from crystallography and molecular dynamics. Biochem Biophys Res Comm 2005;338:1515-18.

Dixon RA. Natural products and disease resistance. Nature 2001;411:843-47.

Eisenreich W, Bacher A, Arigoni D, Rohdich F. Biosynthesis of isoprenoids via the non-mevalonate pathway. Cell Mol Life Sci 2004;61:1401-26.

Emanuelsson O, Nielsen H, Von Heijne G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 1999;8978-84.

Ewing B, Green P. Base-calling of automated sequencer traces using phred, II: Error probabilities. Genome Res 1998;8:186-94a.

Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred, I: Accuracy assessment. Genome Res 1998;8:175-85b.

Feliu JX, Cubarsi R, Villaverde A. Optimized Release of Recombinant Proteins by Ultrasonication of E. coli Cells. Biotechnol Bioeng 1998;58:536-40.

Jahns, MT. Avaliação da expressão gênica diferencial entre folhas e tecidos vasculares de *Eucalyptus grandis*. Porto Alegre: UFRGS, 2007, 88 p. Dissertation (M. Sc.) – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Cunningham FX JR, Gantt E. Identification of Multi-Gene Families Encoding Isopentenyl Diphosphate Isomerase in Plants by Heterologous Complementation in *Escherichia coli*. Plant Cell Physiol 2000;41:119-23.

Guex N, Peitsch MC. Swiss-model and the Swis Pdb Viewer: and environment for comparative protein modeling. Electrophoresis 1997;18:2714-23.

Harada H, Misawa N. Novel approaches and achievements in biosynthesis of functional isoprenoids in Escherichia coli. App Microb Biotechnol 2009;84:1021-31.

Harborne JB. Twenty-five years of chemical ecology. Nat Prod Rep 2001;18:361-79.

Hess B, Bekker H, Berendsen JC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463-72.

Hirts JD, Brooks III CL. Helicity, circular dichroism and molecular dynamics of protein, J Mol Biol,1994;243:173-78.

Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics, J Mol Graph Model 1996;14:33-38.

IUPAC-IUB Commission on biochemical nomenclature, abbreviations and symbols for description of conformation of polypeptide chains. J Mol Biol 1970;52:1-17.

Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopol 1983;12:2577-637.

Kaneda K, Kuzuyama T, Takagi M, Hayakawa Y, Seto H. An unusual isopentenyl diphosphate isomerase found in the mevalonate pathway gene cluster from *Streptomyces* sp. strain CL190. Proc Natl Acad Sci USA 2001;98:932-37.

Kumar S, Tamura K, Jakobsen IB, Nei M. MEGA2: molecular evolutionary genetics analysis software. Bioinform 2001;12:1244–45.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;227:680-85.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinform 2007;23:2947-48.

Laskowski A, Macarthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the steochemical quality of protein structures. J Appl Cryst 1993;26:283-91.

Lee J-K, Oh D-K, Kim S-Y. Cloning and characterization of the *dxs* gene, encoding 1-deoxy-d-xylulose 5-phosphate synthase from *Agrobacterium tumefaciens*, and its overexpression in *Agrobacterium tumefaciens*. J Biotechnol 2007;128:555–66.

Leech NL, Barrett KC, Morgan GA. SPSS for intermediate statistics. Use and interpretation, 2° ed. Erlbaum, London, 2005.

Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C}{}_{T}$ Method. Methods 2001;25:402–08.

McGarvey DJ, Croteau R. Terpenoid metabolism. Plant Cell 1995;7:1015-26.

Pang Y, Shen G, Berge T, Wu CW, Sun X, Tang K. Molecular cloning, characterization and heterologous expression in *Saccharomyces cerevisiae* of a mevalonato diphosphate decarboxylase cDNA from Ginkgo biloba. Physiol Plant 2006;127:19-27.

Phillips MA, D'auria JC, Gershenzon J, Pichersky E. The Arabidopsis thaliana Type I Isopentenyl Diphosphate Isomerases Are Targeted to Multiple Subcellular Compartments and Have Overlapping Functions in Isoprenoid Biosynthesis. Plant Cell 2008;20:677–96.

Rodríguez-Concepción M. Early steps in isoprenoid biosynthesis: Multilevel regulation of the supply of common precursors in plant cells. Phytochem Rev 2006;5:1–15.

Rodríguez-Concepción M, Boronat A. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids: a metabolic milestone achieved through genomics. Plant Physiol 2002;130:1079-89.

Rohmer M, Seemann M, Horbach S, Bringer-Meyer S, Sahm H. Glyceraldehyde 3-phosphate and pyruvate as precursors of isoprenic units in an alternative pathway for terpenoid biosynthesis. J Am Chem Soc 1996;118:2564–66.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–25.

Small I, Wintz H, Akashi K, Mireau H. Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol Biol 1998;38:265-77.

Sprenger GA, Schörken U, Wiegert T, Grolle S, De Graaf AA, Taylor SV, Begley TP, Bringer-Meyer S, Sahm H. Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin and pyridoxol. Proc Natl Acad Sci USA 1997;94**:** 12857-62.

Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 2003;31:3381–85.

Lois LM, Campos N, Putra SR, Danielsen K, Rohmer M, Boronat A. Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase.like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. Proc Natl Acad Sci USA 1998;95:2105–10.

Thibodeaux CJ, Mansoorabadi SO, Kittleman W, Chang W-C, Liu H-W. Biochem,2008;47:2547-58.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJCJ. GROMACS: fast, flexible, and free. Comput Chem 2005;26:1701-18.

Van Der Spoel D, Lindahl E, Hess B, Van Buuren AR, Apol E, Meulenhoff PJ, Tieleman DP, Sijbers ALTM, Feenstra KA, Van Drunen R, Berendsen HJC. GROMACS user manual version 4.0, Department of Biophysical Chemistry, University of Groningen, Groningen, The Netherlands, 2008.

Van Gunsteren WF, Billeter SR, Eising AA, Huenenberger PH, Krueger P, Mark AE, Scott WRP, Tironi IG. Biomolecular Simulation: *The GROMOS96* Manual and User Guide; Vdf Hochschulverlag, AG Zurich: Switzerland, 1996.

Verli H, Guimarães JA. Insights into the induced fit mechanism in antithrombin-heparin interactions using molecular dynamiscs silulations. J Mol Graph Model 2005;24:203-12.

Verpoorte R, Memelink J. Engineering secondary metabolite production in plants. Curr Opin Biotechnol 2002;13:181–87.

Xiao Y, Zahariou G, Sanakis Y, Liu P. IspG Enzyme Activity in the Deoxyxylulose Phosphate Pathway: Roles of the Iron-Sulfur Cluster. Biochem, 2009;48:10483-85.

Zhang C, Liu L, Xu H, Wei Z, Wang Y, Lin Y, Gong W. Crystal Structures of Human IPP Isomerase: New Insights into the Catalytic Mechanism. J Mol Biol 2007;366:1437-46.

Zheng W, Sun F, Bartlam M, Li X, Li R, Rao Z. The Crystal Structure of Human Isopentenyl Diphosphate Isomerase at 1.7 Å Resolution Reveals its Catalytic Mechanism in Isoprenoid Biosynthesis. J Mol Biol 2007;366:1447-58.

Legends of Figures

Figure 1. *EgIPPI1* full-length cDNA and deduced amino acid sequences. Start (ATG) and stop (TAA) codons are underlined in bold. Methionine residues that may represent alternative translational initiation are marked in squares. The classic poly(A) signal (AATAAA) is underlined.

Figure 2. *EgIPPI2* full-length cDNA and deduced amino acid sequences. Start (ATG) and stop (TAA) codons are underlined in bold. A putative poly(A) signal (AATAAG) is underlined.

Figure 3. Multi-alignment of amino acid sequences of EgIPPI1, EgIPPI2 and other plant IPPIs. Identical amino acids are shown in white with black background. Conserved amino acid residues among sequences are represented in grey background. Residues putatively involved in the catalysis are marked with black circles. Metal binding residues are indicated by black diamonds and the Nudix domain is marked with white diamonds. Methionine residues that may represent alternative translational initiation are marked in squares. The aligned IPPIs were from *Melaleuca alternifolia* (MaIPPI1 and MaIPPI2, GenBank accession no.: AF483191 and AF483190, respectively), *Arabidopsis thaliana* (AtIPPI1 and AtIPPI2, BAB09611 and AAF29979, respectively), *Zea mays* (ZmIPPI1 and ZmIPPI2, ACR38423 and ACG24849, respectively), *Catharanthus roseus* (CaIPPI1, ABW98669), and *Nicotiana tabacum* (NtIPPI1 and NtIPPI2, BAB40973 and BAB40974, respectively).

Figure 4. Phylogenetic tree analysis of IPPI protein sequences from different organisms by MEGA and CLUSTAL W alignments. The neighbor-joining method was used to construct the tree. IPPIs without distinction between isoforms 1 or 2 have letters A or B added to its name. Bootstraps values ≥ 80 after 1000 repetitions are marked with dots. GenBank accessions to plant IPPIs sequences are indicated in the legend of Fig. 3. Additionally to those, IPPI sequences from other plants, bacteria, fungi and mammals were included in the analysis, like *Tagetes erecta* IPPIA and IPPIB (AAG10423 and AAF29977), *Adonis aestivalis* IPPIA and IPPIB (AAF29973 and AAF29974), *Bupleurum chinense* IPPIA and IPPIB (ACV74320 and ABY90138), *Hevea brasiliensis* IPPIA and IPPIB (BAF98286 and BAF98287), *Populus trichocarpa* IPPIA and IPPIB

(EEE99850 and ACD70403), *Candida albicans* IPPI (EEQ45250), *Neosartorya fischeri* IPPI (EAW16336), *Aspergillus clavatus* IPPI (EAW06787), *Coccidioides posadasii* IPPI (EER24566), *Ajellomyces capsulatus* IPPI (EEH06877), *Mus musculus* IPPI1 and 2 (EDL32280 and AAH99586, respectively), *Rattus norvegicus* IPPI1 and 2 (EDL86437 and XP_225502, respectively), *Homo sapiens* IPPI1 and 2 (AAI07894 and AAK49436), *Pan troglodytes* IPPI1 and 2 (XP_507622 and XP_507621), *Bos taurus* IPPI1 (AAI16134), *Escherichia coli* IPPI (ZP_05949076), and *Salmonella enterica* IPPI (GenBank accession no.: CBG26007).

Figure 5. SDS-PAGE analysis of protein extracts derived from control and recombinant *E. coli* cultures expressing the GST-*EgIPPI1* and GST-*EgIPPI2* gene fusions. The protein soluble extracts from *E. coli* BL21 (DE3) were run on a 12% (w/v) polyacrylamide gel (30% acrylamide/0.8% bis-acrylamide stock) under denaturing conditions. Cells were cultured in rotatory shakers (180 rpm) under 30 °C during 14 h. No induction (-) or IPTG induction (+) was applied to cultures as indicated. Cells harboring the pGEX-4T-3 and pGEX-4T-2 parental vectors or the recombinant pGEX-4T-3::*EgIPPI1 and* pGEX-4T-2::*EgIPPI2* are identified by the plasmid name. The Broad Range Protein Molecular Weight Marker (Promega) was employed in the first and fifth lanes for protein weight estimation in kDa. Black arrows indicate the bands relative to the GST tag and the white arrows indicate the bands relative to the GST-EgIPPI1 and GST-IPPI2 fusion proteins.

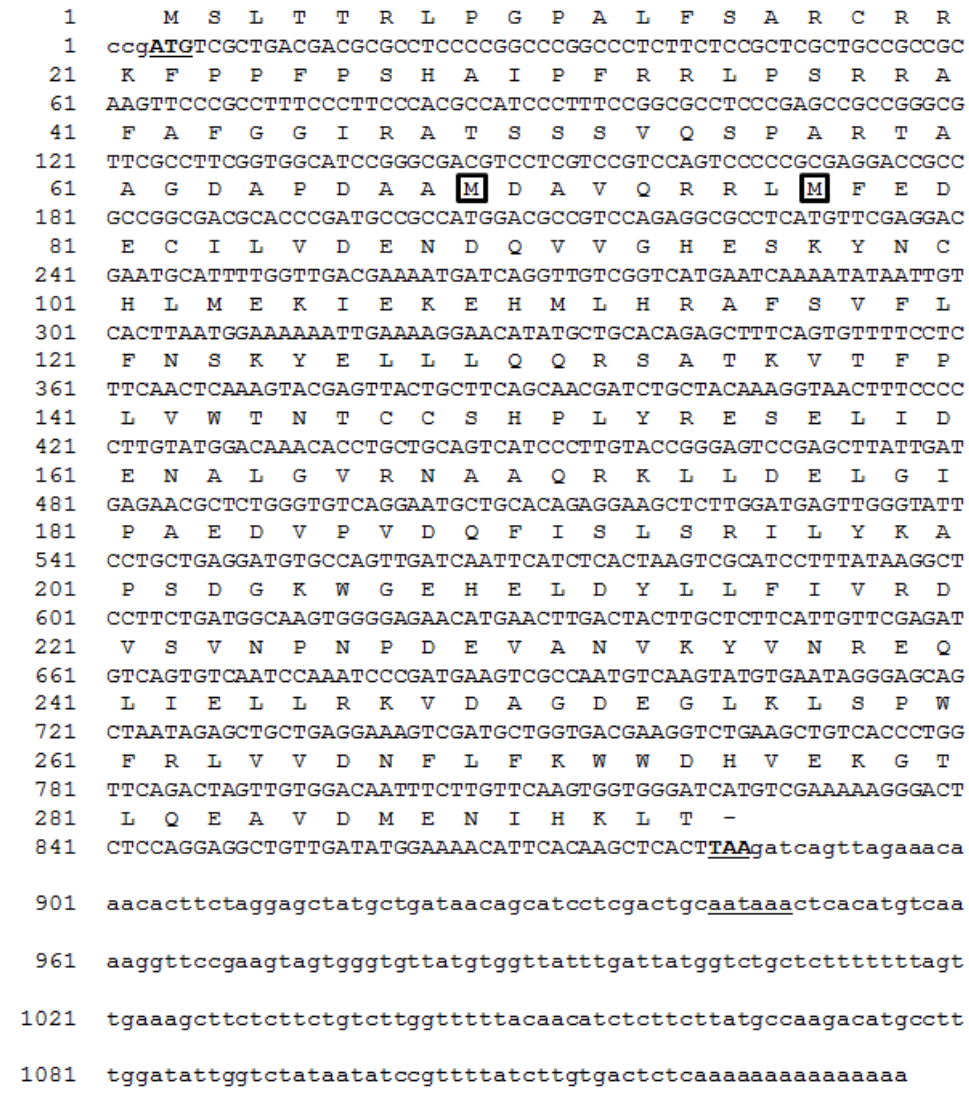
Figure 6. Expression analysis of the *EgIPPI1* and *EgIPPI2*, and other genes important to metabolism of isoprenoid in plants using microarray DNA hybridization. The signal intensities were normalized by signals obtained for the reference *rnp* gene in leaves. from leaves. Student´s *t*-test was applied for statistical treatment of the data. Results are expressed as means ± standard errors of signal intensity, considering two biological replicates for leaves and four biological replicates for xylem, in addition to two technical replicates for each replicate (*** $p<0,001$).

Figure 7. N-terminal unfolded hIPPI1 (PDB ID 2DHO) crystal regenerated, the N-terminal and the C-terminal regions are coulored in red and blue, respectively (A). Superimposition of crystal hIPPI1 (PDB ID 2ICK; green) and the structure obtained at 100 ns (light orange) of MD (B). Lateral chains of residues Lys37 and Ile75 from

Crystal and from structures obtained in 10 ns intervals are coulored in blue and red, respectively. The Mg^{2+} cofactor is coulored in purple. Molecular surface of Crystal hIPPI1 with the active pocket covered (C). N-terminus helix from Crystal and MD structures are coulored in green and red respectively. Molecular surface of EcIPPI1 (PDB ID 1HZT) in green superimposed to N-terminus helix from hIPPI1 crystal and structure at 100 ns coulored in light orange and red, respectively (D).

Figure 8. Root Mean Square Deviation (RMSD) analysis for the entire protein (A). Root Mean Square Fuctuation (RMSF) (B). Multi-alignment of the secondary structures obtained throughout the MD simulation (C). The red boxes show the N-terminus helix region. A scale from white (accessible) to blue (buried) was used in the exposure to solvent prediction.

Figures


1 M S L T T R L P G P A L F S A R C R R
1 ccgATGTCGCTGACGACGCGCCTCCCCGGCCCGGCCCTCTTCTCCGCTCGCTGCCGCCGC
21 K F P P F P S H A I P F R R L P S R R A
61 AAGTTCCCGCCTTTCCCTTCCCACGCCATCCCTTTCCGGCGCCTCCCGAGCCGCCGGGCG
41 F A F G G I R A T S S S V Q S P A R T A
121 TTCGCCTTCGGTGGCATCCGGGCGACGTCCTCGTCCGTCCAGTCCCCCGCGAGGACCGCC
61 A G D A P D A A M D A V Q R R L M F E D
181 GCCGGCGACGCACCCGATGCCGCCATGGACGCCGTCCAGAGGCGCCTCATGTTCGAGGAC
81 E C I L V D E N D Q V V G H E S K Y N C
241 GAATGCATTTTGGTTGACGAAAATGATCAGGTTGTCGGTCATGAATCAAAATATAATTGT
101 H L M E K I E K E H M L H R A F S V F L
301 CACTTAATGGAAAAAATTGAAAAGGAACATATGCTGCACAGAGCTTTCAGTGTTTTCCTC
121 F N S K Y E L L L Q Q R S A T K V T F P
361 TTCAACTCAAAGTACGAGTTACTGCTTCAGCAACGATCTGCTACAAAGGTAACTTTCCCC
141 L V W T N T C C S H P L Y R E S E L I D
421 CTTGTATGGACAAACACCTGCTGCAGTCATCCCTTGTACCGGGAGTCCGAGCTTATTGAT
161 E N A L G V R N A A Q R K L L D E L G I
481 GAGAACGCTCTGGGTGTCAGGAATGCTGCACAGAGGAAGCTCTTGGATGAGTTGGGTATT
181 P A E D V P V D Q F I S L S R I L Y K A
541 CCTGCTGAGGATGTGCCAGTTGATCAATTCATCTCACTAAGTCGCATCCTTTATAAGGCT
201 P S D G K W G E H E L D Y L L F I V R D
601 CCTTCTGATGGCAAGTGGGGAGAACATGAACTTGACTACTTGCTCTTCATTGTTCGAGAT
221 V S V N P N P D E V A N V K Y V N R E Q
661 GTCAGTGTCAATCCAAATCCCGATGAAGTCGCCAATGTCAAGTATGTGAATAGGGAGCAG
241 L I E L L R K V D A G D E G L K L S P W
721 CTAATAGAGCTGCTGAGGAAAGTCGATGCTGGTGACGAAGGTCTGAAGCTGTCACCCTGG
261 F R L V V D N F L F K W W D H V E K G T
781 TTCAGACTAGTTGTGGACAATTTCTTGTTCAAGTGGTGGGATCATGTCGAAAAAGGGACT
281 L Q E A V D M E N I H K L T -
841 CTCCAGGAGGCTGTTGATATGGAAAACATTCACAAGCTCACTTAAgatcagttagaaaca

901 aacacttctaggagctatgctgataacagcatcctcgactgcaataaactcacatgtcaa

961 aaggttccgaagtagtgggtgttatgtggttatttgattatggtctgctttttttagt

1021 tgaaagcttctcttctgtcttggtttttacaacatctcttcttatgccaagacatgcctt

1081 tggatattggtctataatatccgttttatcttgtgactctcaaaaaaaaaaaaaaa

Figure 1.

```
  1  ctcctcctccgctgtccttcccctctccctcctcgtccctctctctctctctgtcgatcg
  1        M  A  D  G  A  D  A  G  M  A  D  G  A  D  A  G  M  D
 61  accgccATGGCCGACGGTGCCGACGCCGGCATGGCCGACGGTGCCGACGCCGGCATGGAC
 21   A  V  Q  R  R  L  M  F  E  D  E  C  I  L  V  D  E  N  D  N
121  GCCGTCCAGCGGCGCCTCATGTTCGAGGACGAATGCATCTTGGTGGATGAGAACGACAAT
 41   V  V  G  H  E  S  K  Y  N  C  H  L  M  E  K  I  E  S  L  N
181  GTCGTCGGTCATGAGTCGAAGTATAACTGTCATTTGATGGAGAAAATCGAGTCCTTGAAT
 61   L  L  H  R  A  F  S  V  F  L  F  N  S  K  Y  E  L  L  L  Q
241  CTGTTGCATAGGGCATTCAGTGTGTTCTTGTTCAACTCAAAGTATGAGCTACTGCTTCAG
 81   Q  R  S  A  T  K  V  T  F  P  L  V  W  T  N  T  C  C  S  H
301  CAACGCTCTGCCACAAAGGTAACGTTCCCCCTTGTGTGGACAAACACCTGCTGCAGCCAT
101   P  L  Y  R  E  S  E  L  I  A  E  N  A  L  G  A  R  N  A  A
361  CCGTTGTACCGGGAGTCCGAGCTCATTGCTGAGAACGCCCTTGGGGCAAGGAATGCTGCA
121   Q  R  K  L  L  D  E  L  G  I  P  A  E  D  V  P  V  D  E  F
421  CAAAGGAAGCTCTTGGACGAACTGGGCATTCCTGCTGAAGATGTGCCAGTTGATGAATTC
141   L  P  L  G  R  M  L  Y  K  A  P  S  D  G  K  W  G  E  H  E
481  CTTCCTCTGGGTCGAATGCTGTATAAGGCACCTTCTGATGGAAAGTGGGGGGGAGCACGAA
161   I  D  Y  L  L  F  I  V  E  D  V  K  V  N  P  D  E  V  P  D
541  ATTGATTACTTGCTGTTTATTGTCGAAGATGTGAAGGTCAATCCCGATGAGGTCCCGGAT
181   I  R  Y  V  N  Q  E  Q  L  K  E  T  L  R  K  P  D  A  G  Q
601  ATCAGGTATGTAAACCAGGAGCAGTTGAAGGAAACTTTGAGGAAACCCGATGCGGGGCAA
201   G  G  L  K  A  F  P  W  F  Q  I  M  G  A  N  F  L  F  K  W
661  GGGGGCCTGAAGGCTTTCCCCTGGTTTCAGATTATGGGTGCCAATTTCTTGTTCAAGTGG
221   V  D  H  V  K  R  G  P  L  K  E  A  A  D  M  K  T  I  H  K
721  GTGGATCATGTCAAGAGGGGGCCACTGAAGGAAGCCGCGGATATGAAAACCATTCACAAG
241   G  P  -
781  GGGCCTTGAgattttgtgtggagtagcagtcaaataaggcgggaaaatttggacttcaat
841  tgatgtttttaggttgttcaatccaaattatgtgggccttaaatttcaaaaggtaagggt
901  aaaactttttttcccccaggccctttttctttttgggggccccctttgtttttttttg
961  gcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

Figure 2.

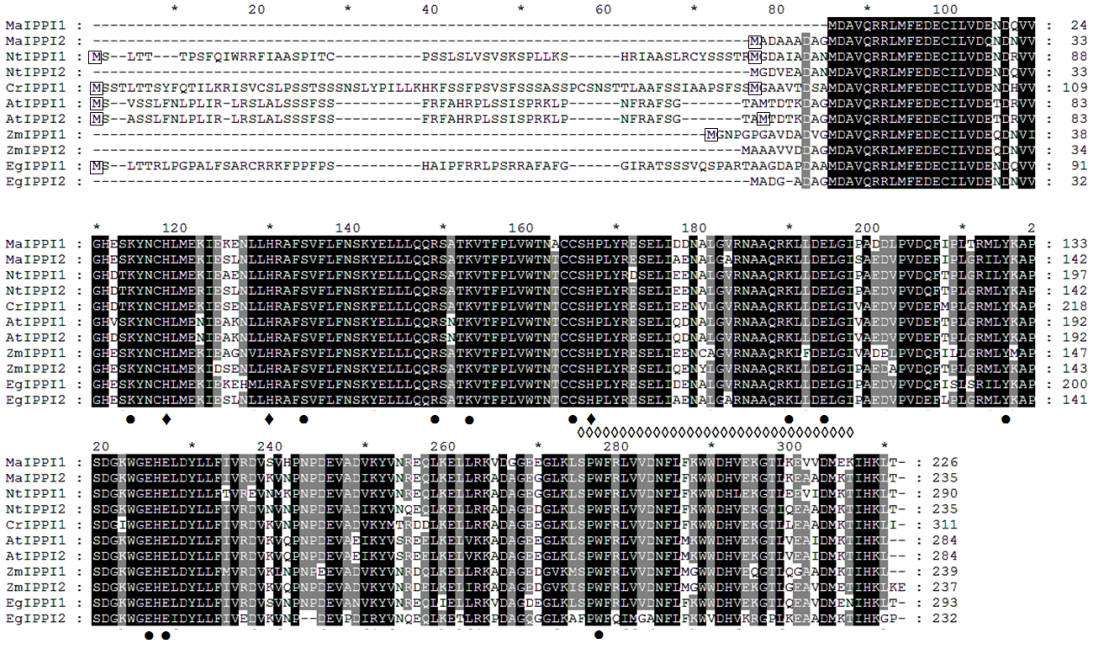


Figure 3.

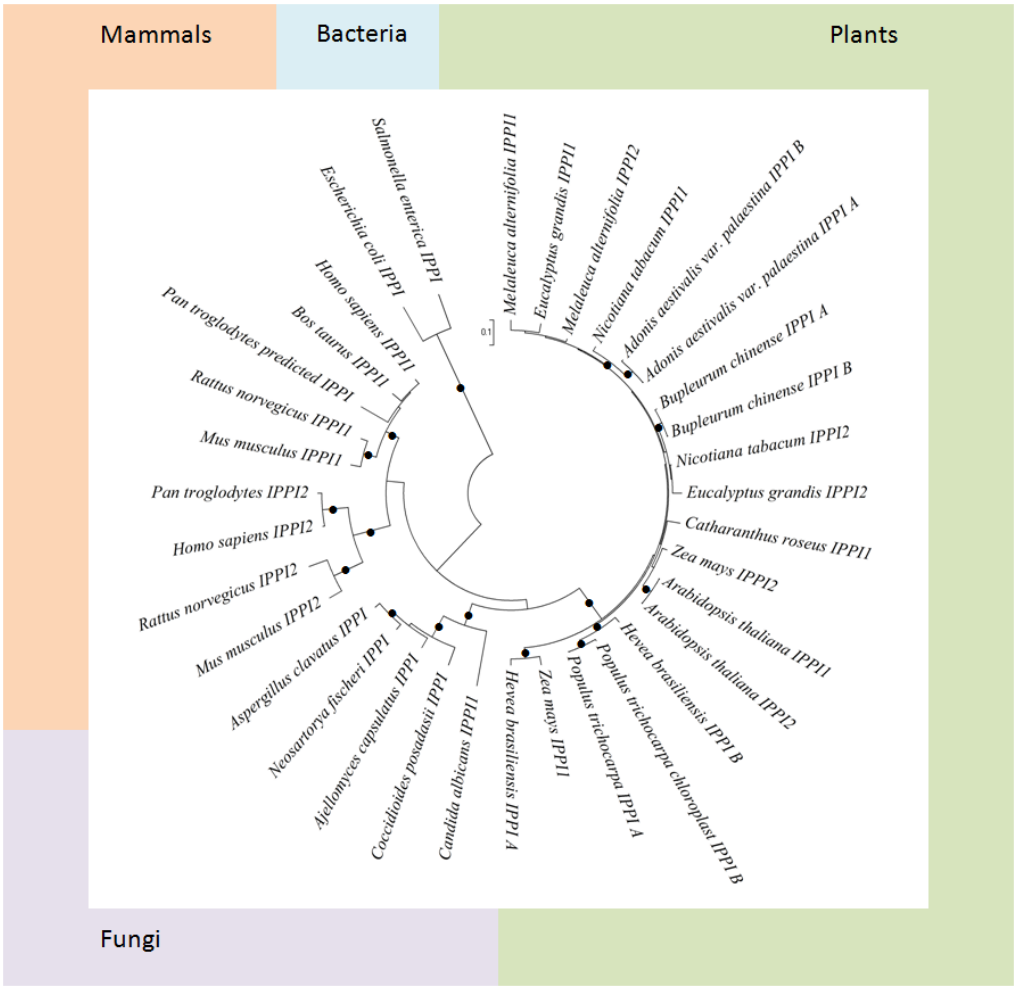


Figure 4.

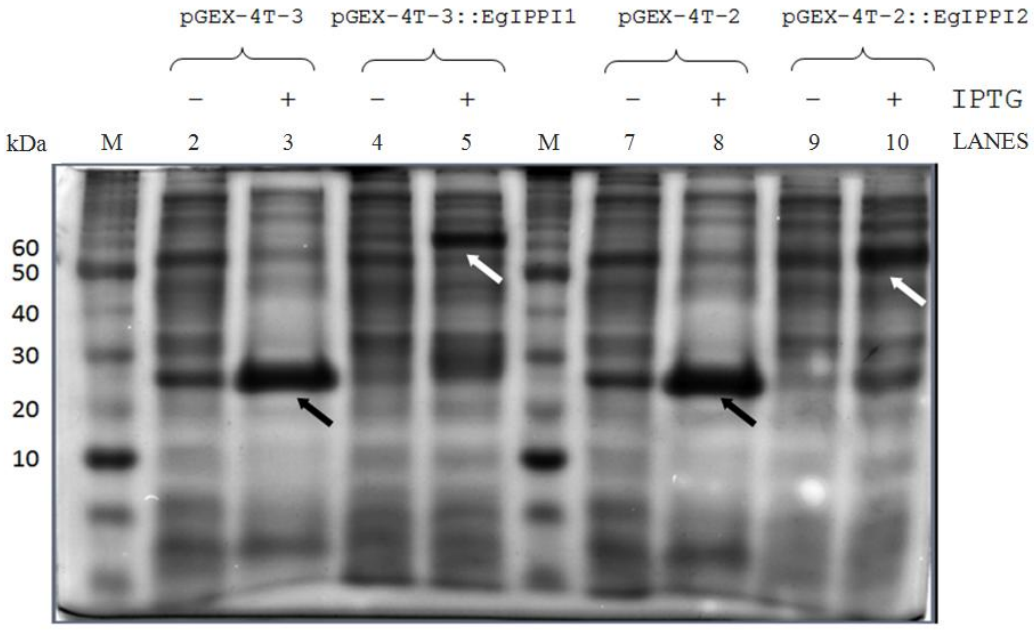


Figure 5.

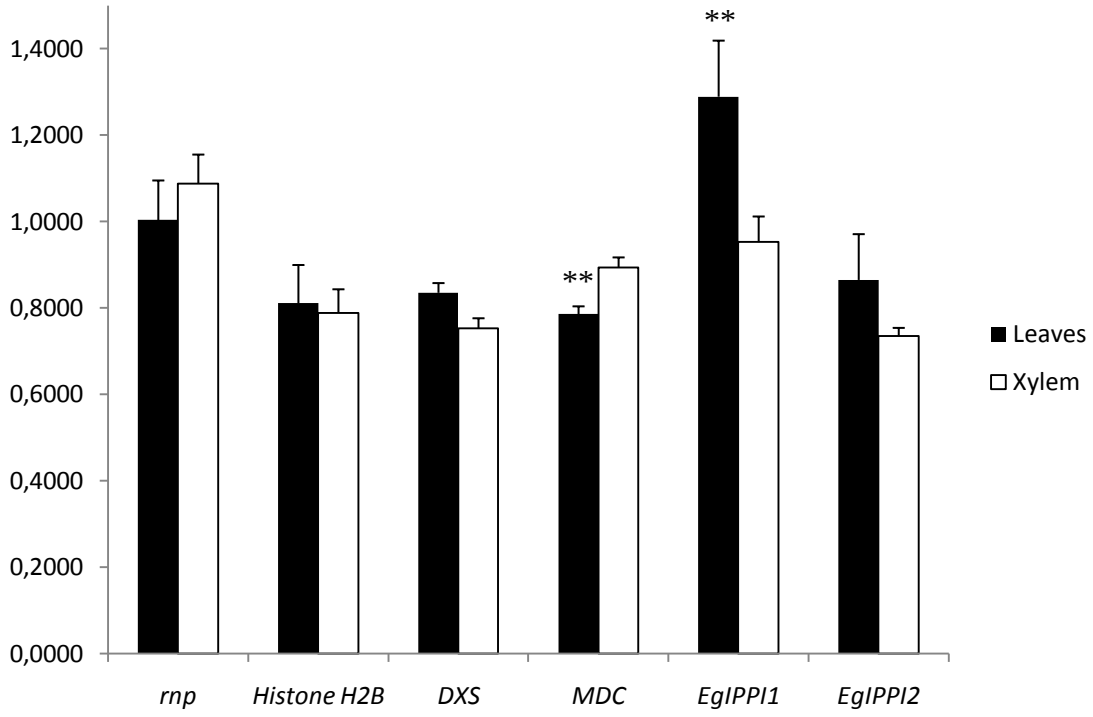


Figure 6.

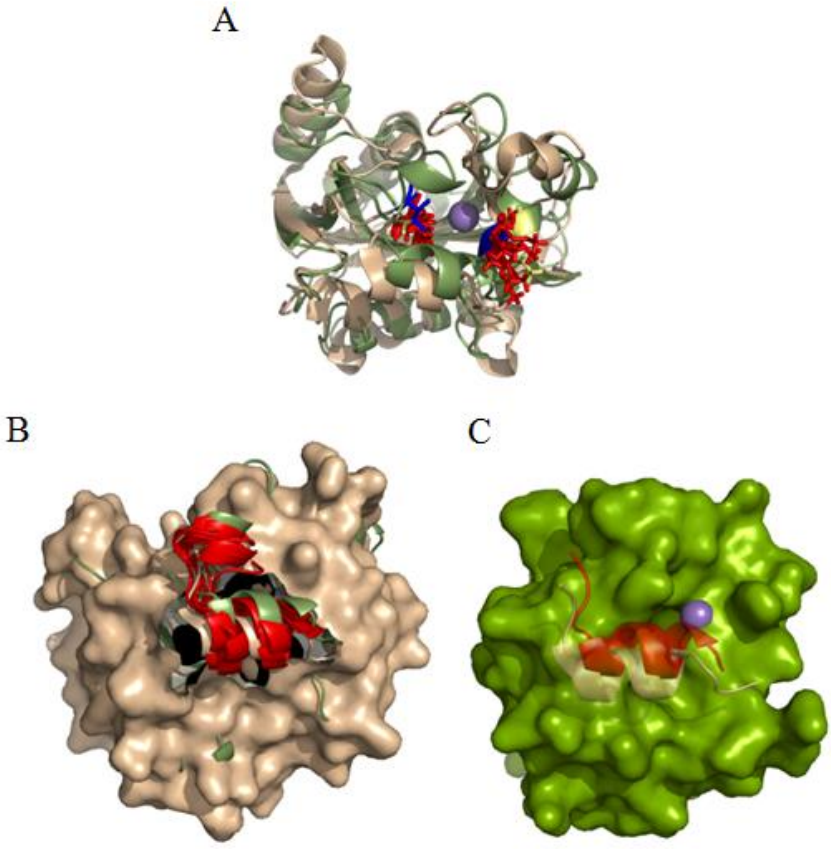


Figure 7.

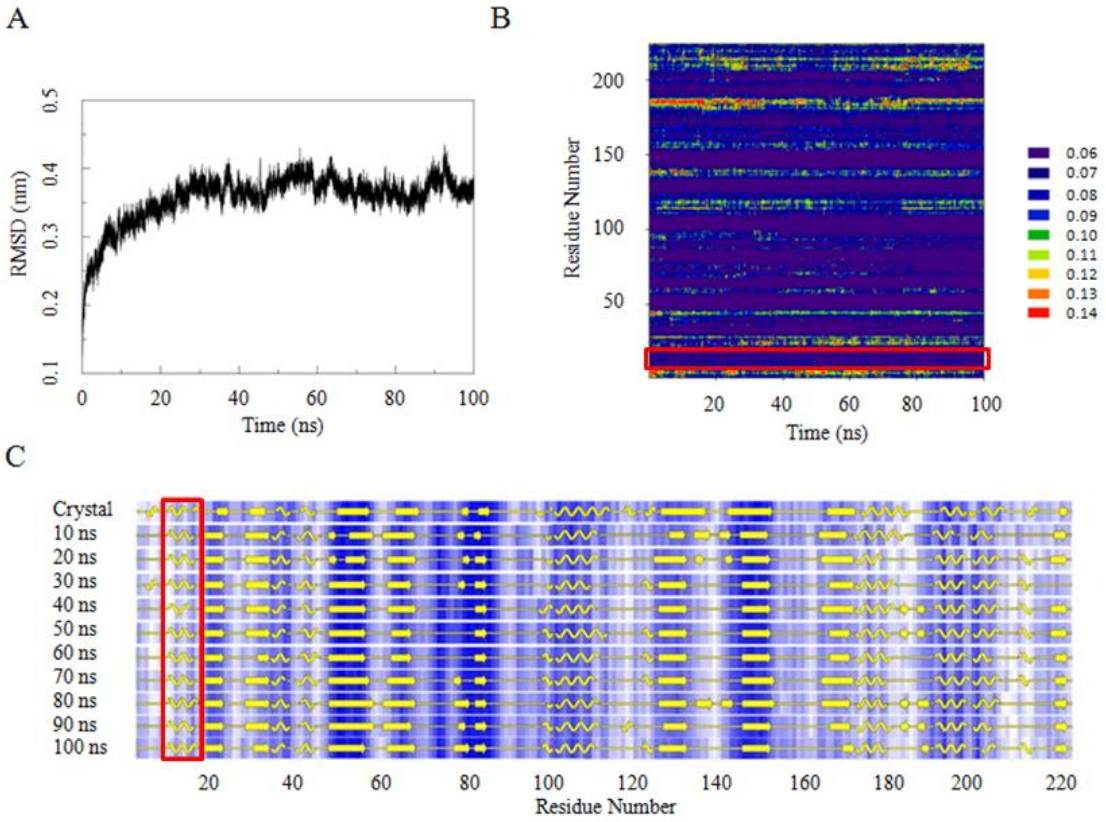

A
RMSD (nm)
0.1
0.2
0.3
0.4
0.5
0
20
40
60
80
100
Time (ns)
B
Residue Number
50
100
150
200
20
40
60
80
100
Time (ns)
0.06
0.07
0.08
0.09
0.10
0.11
0.12
0.13
0.14
C
Crystal
10 ns
20 ns
30 ns
40 ns
50 ns
60 ns
70 ns
80 ns
90 ns
100 ns
20
40
60
80
100
120
140
160
180
200
220
Residue Number

Figure 8.

CAPÍTULO IV

CONCLUSÕES E PERSPECTIVAS

Os óleos essenciais de *Eucalyptus* são terpenóides de elevado interesse para a indústria farmacêutica, de alimentos e de perfumaria (Vitti & Brito, 2003). Nos últimos anos, houve um grande aumento do nosso conhecimento acerca das rotas de biossíntese de compostos terpenóides (D´auria *et al.*, 2005). Especificamente para plantas de eucalipto, esse conhecimento abre novas perspectivas para uma maior produção de isoprenóides de importância médica e comercial em plantas modificadas por estratégias de bioengenharia (Verpoorte *et al.*, 2000; Pasquali *et al*., 2006).

Neste trabalho, nós objetivamos clonar, caracterizar e expressar heterologamente cinco genes de fundamental importância para a produção de compostos isoprenóides, os genes *dxs, dxr, mdc, ippi1* e *ippi2* de *Eucalyptus grandis*. Nós obtivemos as sequências completas dos genes *dxr, ippi1 e ippi2* denominados de *EgDXR, EgIPPI1* e *EgIPPI2*, respectivamente. A partir deste ponto, abrem-se novas portas para a produção de plantas de eucalipto capazes de produzir maior quantidade de óleos essenciais, visto que as enzimas clonadas e caracterizadas podem ser capazes de aumentar o fluxo pelas rotas metabólicas que produzem IPP e DMAPP.

Embora uma análise por RT-qPCR seja necessária para confirmação dos resultados, a análise de expressão diferencial pela técnica de hibridização de microarranjos de DNA revelou que os genes *EgDXR* e *EgIPPI1* são significativamente mais expressos em folhas do que em xilema, sugerindo que esses genes possam ser importantes para a produção óleos essências.

As análises *in silico* evidenciaram a presença dos principais domínios e resíduos catalíticos das proteínas e os estudos filogenéticos localizaram as proteínas dentro de grupos de plantas sugerindo que as sequências codificam proteínas funcionais. Conforme esperado, EgDXR possui a sequência de peptídeo sinal para plastídios onde a rota MEP está compartimentalizada (Kirby *et al.*, 2009). A proteína plastídica EgIPPI1 apresenta um peptídeo sinal para esse compartimento enquanto a sua isoforma EgIPPI2 não possui peptídeo sinal sendo, supostamente, direcionada para o citosol. No entanto, foi sugerido que as *IPPI*s possuem códons de início de tradução alternativos para localização em diferentes compartimentos (Cunningham *et al.*, 2000), sendo assim, é

possível que a enzima EgIPPI1 seja também direcionada para o citosol e não somente para um compartimento apenas.

Todas as IPPIs eucarióticas possuem uma extensão N-terminal que está ausente em IPPIs procarióticas (Zheng *et al.*, 2007). A função dessa extremidade ainda não está bem estabelecida, e diferentes hipóteses têm sido levantadas para elucidar essa característica. Zhang *et al.* (2007), com base em dados cristalográficos, sugeriram um mecanismo de flap no qual os aminoácidos presentes nessa região desenovelam-se quando a proteína está em seu estado livre e voltam a formar uma hélice quando ligada ao substrato o que fecharia a entrada para o sítio catalítico. Enquanto, Zheng *et al.*(2007), também baseados em dados cristalográficos, sugerem que essa hélice seja estável e que o local de entrada do substrato seja levemente deslocado quando comparado às IPPIs de *E. coli*. Para ajudar a elucidar o papel da extremidade N-terminal das IPPIs eucarióticas, nós realizamos uma análise de dinâmica molecular e evidenciamos que essa região permaneceu estável até o tempo de 100 ns. Dessa forma, nossos resultados corroboram a hipótese de uma entrada para substrato alternativa proposta por Zheng *et al*. (2007).

Como perspectivas de continuidade dessa linha de pesquisa, estão:

- Medir a atividade específica das proteínas EgDXR, EgIPPI1 e 2
- Realizar análises da expressão dos genes por RT-qPCR sob diferentes tratamentos para identificar um estado de expressão elevado. Medindo a quantidade de óleos essenciais nos mesmos tratamentos seria possível estabelecer uma relação entre os níveis de mRNA dos genes em estudo com a produção de óleos essenciais.
- Utilizar os genes identificados para o melhoramento de plantas por meio de suas superexpressões ou inibições por transgenia para a obtenção de metabólitos de interesse humano por intermédio de estratégias de bioengenharia.
- As proteínas EgIPPI1 e 2 heterologamente expressas em *E. coli* BL21 (DE3) poderão ser submetidas a experimentos cristalográficos para um melhor entendimento do papel conformacional e implicações de suas extremidades N-terminais.
- Realizar outras análises estruturais para elucidação de características conformacionais dessas proteínas incluindo a dinâmica molecular de EgDXR, de EgIPPI1 e de EgIPPI2.

REFERÊNCIAS (Capítulos I e IV)

ADAM, P.; HECHT, S.; EISENREICH, W.; KAISER, J.; GRAWERT, T.; ARIGONI, D.; BACHER, A.; ROHDICH, F. Biosynthesis of terpenes: Studies on 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate reductase. *Proceedings of the National Academy of Science USA*, 99: 12108-12113, 2000.

AGRANOFF, B. W.; EGGERER, H.; HENNING, U; LYNEN, F. Biosynthesis of Terpenes VII. Isopentenyl Pyrophosphate Isomerase. *The Journal of Biological Chemistry*, 236 (2): 326-332, 1959.

AHARONI, A.; GIRI, A. P.; DEUERLEIN, S.; GRIEPINK, F.; DE KOGEL, W.-J.; VERSTAPPEN, F. W. A.; VERHOEVEN, H. A.; JONGSMA, M. A.; SCHWAB, W.; BOUWMEESTER, H. J. Terpenoid metabolism in wild-type and transgenic *Arabidopsis* plants. *Plant Cell*, 15: 2866-2884, 2003.

ALEXANDROV, N. N.; TROUKHAN. M. E.; BROVER, V. V.; TATARINOVA, T.; FLAVELL, R. B.; FELDMANN, K. A. Features of Arabidopsis genes and genome discovered using full-length cDNAs. *Plant Molecular Biology*, 60: 71-87, 2006.

ARAKI, N.; KUSUMI, K.; MASAMOTO, K.; IBA, K. Temperaturesensitive Arabidopsis mutant defective in 1-deoxy-D-xylulose 5-phosphate synthase within the plastid non-mevalonate pathway of isoprenoid biosynthesis. *Physiologia Plantarum*, 108: 19-24, 2000.

BENNET, J. W.; BENTLEY, R. What's in a name? Microbial secondary metabolism. *Advances in Applied Microbiology*, 34: 1-28, 1989.

BASTOLLA, F. M. Seleção e avaliação de genes de referência para estudos de expressão gênica em *Eucalyptus*. Porto Alegre: UFRGS, 2007, 97 p. Dissertação (Mestrado) – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

BOUVIER, F.; RAHIER, A.; CAMARA, B. Biogenesis, molecular regulation and function of plant isoprenoids. *Progress in Lipid Research*, 44: 357-429, 2005.

BOUVIER, F.; D'HARLINGUE, A.; SUIRE, C.; BACKHAUS, R.A.; CAMARA, B. Dedicated roles of plastid transketolases during the early onset of isoprenoid biogenesis in pepper fruits. *Plant Physiology*, 117: 1423-1431, 1998.

CAMPOS, N.; LOIS, L.M.; BORONAT, A. Nucleotide sequence of a rice cDNA encoding a transketolase-like protein homologous to the Arabidopsis CLA1 gene product. *Plant Physiology*, 115: 1289-1293, 1997.

CARRETERO-PAULET, L.; AHUMADA, I.; CUNILLERA, N.; RODRIGUEZ-CONCEPCION, M.; FERRER, A.; BORONAT, A.; CAMPOS, N. Expression and molecular analysis of the *Arabidopsis* DXR gene encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. *Plant Physiology*, 129, 1581-1591, 2002.

CHAHED, K.; OUDIN, A.; GUIVARCH, N.; HAMDI, S.; CHENIEUX, J. C.; RIDEAU, M.; CLASTRE, M. 1-Deoxy-d-xylulose 5-phosphate synthase from periwinkle: cDNA identification and induced gene expression in terpenoidindole alkaloid-producing cells. *Plant Physiology and Biochemistry*, 38: 559-566, 2000.

CHEMLER, J. A.; KOFFAS, M. A. G. Metabolic engineering for plant natural product biosynthesis in microbes. *Current Opinion in Biotechnology*, 19: 597-605, 2008.

CHEMLER, J.A.; YAN, Y.; KOFFAS, M.K.G. Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in *Saccharomyces cerevisiae*. *Microbial Cell Factories*, *5*: 20-30, 2006.

CHENG, A. X.; LOU, Y. G.; MAO, Y. B.; LU, S.; WANG, L. J.; CHEN, X.Y. Plant Terpenoids: Biosynthesis and Ecological Functions. *Journal of Integrative Plant Biology*, *49*: 179-186, 2007.

COOPER, B.; CLARKE, J.; BUDWORTH, P.; KREPS, J.; HUTCHISON, D.; PARK, S.; GUIMIL, S.; DUNN, M.; LUGINBUHL, P.; ELLERO, C.; GOFF, S.A.; GLAZEBROOK, J. A network of rice genes associated with stress response and seed development. *Proceedings of the National Academy of Sciences*, 100: 4945-4950, 2003.

CRAGG, G. M.; NEWMAN, D. J.; AND SNADER, K. M. Natural products in drug discovery and development. *Journal of Natural Products, 60*: 52-60, 1997.

CROTEAU, R. Biosynthesis and catabolism of monoterpenoids. *Chemical Reviews,* 87: 929-954, 1987.

CUNNINGHAM, F. J. JR; LAFOND T. P.; GANTT E. Evidence of a role for LytB in the nonmevalonate pathway of isoprenoid biosynthesis. *The Journal of Bacteriology*, 182: 5841-5648, 2000.

D'AURIA, J. C.; GERSHENZON, J. The secondary metabolism of *Arabidopsis thaliana*: growing like a weed. *Current Opinion in Plant Biology*, 8: 308-316, 2005.

DE RUYCK, J. & WOUTERS, J. Structure-based Drug Design Targeting Biosynthesis of Isoprenoids: A Crystallographic State of the Art of the Involved Enzymes. *Current Protein and Peptide Science*, 9 (2): 117-137, 2008.

DE OLIVEIRA, L. A. Análise transcricional dos genes *ISA1*, *NFS1* e *ISU1* de *Eucalyptus grandis* sob estresse. Porto Alegre: UFRGS, 2008, 85 p. Dissertação (Mestrado) – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.

DHINGRA, V.; RAO, K. V.; NARASU, M. L. Current status of artemisinin and its derivatives as antimalaria drugs. *Life Sci.*, 66: 279-300, 2000.

DIXON R. A. Natural products and disease resistance. *Nature*, 411: 843-847, 2001.

DORAN, J. C. Eucalyptus leaf oils: use, chemistry, distillation and marketing. Melbourne: Inkata, 11-28, 1991.

DUBEY, V. S.: BHALLA, R.; LUTHRA, R. An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *Journal of Biosciences*, 28: 637-646, 2003.

DUDAREVA, N.; ANDERSSON, S.; ORLOVA, I.; GATTO, N.; REICHEL, M.; RHODES, D.; BOLAND, W.; GERSHENZON, J. The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. *Proceedings of the National Academy of Sciences*, 102: 933-938, 2005.

EISENREICH, W.; BACHER, A; ARIGONI, D; ROHDICH, F. Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cellular and Molecular Life Sciences*, 61: 1401–1426, 2004.

EISENREICH, W.; ROHDICH, F.; BACHER, A. Deoxyxylulose phosphate pathway to terpenoids. *Trends in Plant Science*, 6: 78–84, 2001.

EISENREICH, W.; SCHWARZ, M.; CARTAYRADE, A.; ARIGONI, D.; ZENK, M. H.; BACHER A. The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chemistry and Biology*, 5: R221–R233, 1998.

ENGPRASERT, S.; SHOYAMA, Y.; TAURA, F. Molecular cloning, expression and characterization of recombinant 1-deoxy-d-xylulose-5-phosphate reductoisomerase from *Coleus forskohlii* Brig. *Plant Science*, 169: 287–294, 2005.

ERSEK, T.; KIRALY, Z. Phytoalexins: warding-off compounds in plants. *Physiologia Plantarum*, 68: 343-346, 1986.

GANJEWALA, D.; KUMAR, S.; LUTHRA, R. An Account of Cloned Genes of Methyl-erythritol-4-phosphate Pathway of Isoprenoid Biosynthesis in Plants. *Current Issues in Molecular Biology*, 11 (Suppl. 1): i35–i45, 2008.

GAO, S.; LIN, J.; LIU, X.; DENG, Z.; LI, Y.; SUN, X.; TANG, K. Molecular Cloning, Characterization and Functional Analysis of a 2C-methyl-d-erythritol 2,4-cyclodiphosphate Synthase Gene from *Ginkgo biloba*. *Journal of Biochemistry and Molecular Biology*, 39: 502–510, 2006.

GONG, Y.; LIAO, Z.; CHEN, M.; ZUO, K.; GUO, L.; TAN, Q.; HUANG, Z.; KAI, G.; SUN,.X.; TAN, F.; TANG, K. Molecular cloning and characterization of a 1-deoxy-d-xylulose 5-phosphate reductoisomerase gene from *Ginkgo biloba*. *Mitochondrial DNA*, 16: 111-120, 2005.

GONG, Y.F.; LIAO, Z.H.; GUO, B.H.; SUN, X.F.; TANG, K.X. Molecular cloning and expression profile analysis of *Ginkgo biloba* DXS gene encoding 1-deoxy-d-xylulose 5-phosphate synthase, the first committed enzyme of the 2-C-methyl-d-erythritol 4-phosphate pathway. *Planta Medica*, 72: 329-335, 2006.

GUEVARA-GARCIA, A.; SAN ROMAN, C.; ARROYO, A.; CORTES, M. E.; DE LA LUZ GUTIERREZ-NAVA, M.; LEON, P. Characterization of the Arabidopsis clb6 mutant illustrates the importance of posttranscriptional regulation of the methyl-d-erythritol 4-phosphate pathway. *Plant Cell*, 17: 628-643, 2005.

HAHN, F. M.; HURLBURT, A. P.; POULTER C. D. *Escherichia coli* open reading frame 696 is idi, a non essential gene encoding isopentenyl diphosphate isomerase. *The Journal of Bacteriology*, 181: 4499-4504, 1999.

HAN, Y. S.; ROYTRAKUL, S.; VERBERNE, M. C.; VAN DER HEIJDEN, R.; LINTHORST, H. J. M.; VERPOORTE, R. Cloning of a cDNA encoding 1-deoxy-d-xylulose 5-phosphate synthase from *Morinda citrifolia* and analysis of its expression in relation to anthraquinone accumulation. *Plant Science*, 164: 911-917, 2003.

HANS, J.; HAUSE, B.; STRACK, D.; WALTER, M. H. Cloning, characterization, and immunolocalization of a mycorrhiza-inducible 1-deoxy-d-xylulose 5-phosphate reductoisomerase in arbuscule-containing cells of maize. Plant Physiology, 134: 614–624, 2004.

HARADA, H.; MISAWA, N. Novel approaches and achievements in biosynthesis of functional isoprenoids in Escherichia coli. *Applied Microbiology and Biotechnology*, 84: 1021-1031, 2009.

HARBORNE, J. B. Twenty-five years of chemical ecology. *Natural Product Reports*, *18*: 361-379, 2001.

HERZ, S.; WEEKUL, W. J.; SCHUHR, C.A.; HECHT, S.; LUTTGEN, H.; SAGNER, S.; FELLERMEIER, M.; EISENREICH, W.; ZENK, M. H.; BACHER, A.; ROHDICH, F. Biosynthesis of terpenoids: *Ygb*B protein converts 4-diphosphocytidyl-2C-methyl-D-erythritol2- phosphate to 2C-methyl-D-erythritol2, 4-cyclodiphosphate; *Proceedings of the National Academy of Science USA*, 97: 2486–2490, 2000.

HOSTETTMANN, K; TERREAUX, C. Search for new lead compounds from higher plants. *Chimia (Aarau)*, 54: 652-657, 2000.

KHEMVONG, S.; SUVACHITTANONT, W. Molecular cloning and expression of a cDNA encoding 1-deoxy-d-xylulose-5-phosphate synthase from oil palm in Elaeis guineensis Jacq. *Plant Science*, 169: 571-578, 2005.

JAHNS, M. T. Avaliação da expressão gênica diferencial entre folhas e tecidos vasculares de *Eucalyptus grandis*. Porto Alegre: UFRGS, 2007, 88 p. Dissertação (Mestrado) – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

JENNEWEIN, S.; WILDUNG, M. R.; CHAU, M.; WALKER, K.; CROTEAU, R. Random sequencing of an induced Taxus cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proceedings of the National Academy of Sciences*, 101: 9149-9154, 2004.

JIN, H.; GONG, Y.; GUO, B.; QIU, C.; LIU, D.; MIAO, Z.; SUN, X.; TANG, K. Isolation and characterization of a 2C-methyl-d-erythritol 2,4-cyclodiphosphate synthase gene from *Taxus media*. *Molecular Biology*, 40: 914-921, 2006.

KIM, S. M.; KUZUYAMA, T.; CHANG, Y. J.; KWON, H. J.; KIM, S. U. Cloning and functional characterization of 2-C-methyl-d-erythritol 4-phosphate cytidyltransferase (GbMECT) gene from *Ginkgo biloba*. *Phytochemistry*, 67: 1435-1441, 2006.

KIRBY, J.; KEASLING J. D. Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering. *Annual Review of Plant Biology*, 60: 335-355, 2009.

KISHIMOTO, S.; OHMIYA, A. Regulation of carotenoid biosynthesis in petals and leaves of chrysanthemum (*Chrysanthemum morifolium*). *Physiologia Plantarum*, 128: 436-447, 2006.

KUZUYAMA, T.; TAKAHASHI, S.; TAKAGI, M.; SETO H. Characterization of 1-deoxy-D-xylulose 5-phosphate reductoisomerase, an enzyme involved in isopentenyl diphosphate biosynthesis, and identification of its catalytic amino acid residue. *The Journal of Biological Chemistry*, 275: 19928-19932, 2000.

KUZUYAMA, T.; SHIMIZU, T.; TAKAHASHI, S.; SETO, H. Fosmidomycin, a speci¢c inhibitor of 1-deoxy-D-xylulose 5-phosphate reductoisomerase in the nonmevalonate pathway for terpenoid biosynthesis. *Tetrahedron Lett.*, 39: 7913-7916, 1998.

LANGE, B. M.; CROTEAU, R. Isoprenoid biosynthesis via a mevalonate-independent pathway in plants: cloning and heterologous expression of 1-deoxy-d-xylulose-5-phosphate reductoisomerase from peppermint. *Archives of Biochemistry and Biophysics*, 365: 170-174, 1999a.

LANGE, B. M.; CROTEAU, R. Isopentenyl diphosphate biosynthesis via a mevalonate-independent pathway: isopentenyl monophosphate kinase catalyzes the terminal enzymatic step. *Proceedings of the National Academy of Sciences*, 96: 13714-13719, 1999b.

LIAO, Z.; CHEN, R.; CHEN, M.; YANG, C.; WANG, Q.; GONG, Y. A new 1-deoxy-d-xylulose 5-phosphate reductoisomerase gene encoding the committed-step enzyme in the MEP pathway from *Rauvolfia verticillata*. *Journal of biosciences*, 62: 296-304, 2007.

LICHTENTHALER, H. K. The 1-deoxi-D-xilulose-5-phosphate pathway os isoprenoid biosynthesis in plant. *Annual Review of Plant Physiology and Plant Molecular Biology*. 50: 47-65, 1999.

LICHTENTHALER, H.K. Non-mevalonate isoprenoid biosynthesis: enzymes, genes and inhibitors. *Biochemical Society transactions*, 28: 785-789, 2001.

LIN, X.; KAUL, S. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature*, 402: 761-768, 1999.

LIU, J. R.; CHOI, D-W.; CHUNG, H-J.; WOO, S-S. Production of Useful Secondary Metabolites in Plants: Functional Genomics Approaches. *Journal of Plant Biology*, 45: 1-6, 2002.

LOIS, L.M.; RODRIGUEZ-CONCEPCION, M.; GALLEGO, F.; CAMPOS, N.; BORONAT, A. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-d-xylulose 5-phosphate synthase. *The Plant Journal*, 22: 503-513, 2000.

LOIS L. M.; CAMPOS N.; PUTRA S. R.; DANIELSEN K.; ROHMER M.; BORONAT A. Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase.like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. *Proceedings of the National Academy of Science*, 95: 2105-2110, 1998.

MAHMOUND, S. S.; CROTEAU, R. B. Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate

reductoisomerase and menthofuran synthase. *Proceedings of the National Academy of Science*, 98: 8915-8920, 2001.

MOEHS, C.P.; TIAN, L.; OSTERYOUNG, K.W.; DELLAPENNA, D. Analysis of carotenoid biosynthetic gene expression during marigold petal development. *Plant Molecular Biology*, 45: 281-293, 2001.

NUGROHO , L. H.; VERPOORTE, R. Secondary metabolism in tobacco. *Plant Cell, Tissue and Organ Culture*, 68**:** 105-125, 2002.

OHYANAGI, H.; TANAKA, T.; SAKAI, H.; SHIGEMOTO, Y.; YAMAGUCHI, K.; HABARA, T.; FUJII, Y.; ANTONIO, B. A.; NAGAMURA, Y.; IMANISHI, T.; IKEO, K.; ITOH, T.; GOJOBORI, T.; SASAKI, T. The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. japonica genome information. *Nucleic Acids Research*, 34: D741-D744, 2006.

PADUCH, R.; KANDEFER-SZERSZEŃ, M.; TRYTEK, M.; FIEDUREK, J. Terpenes: substances useful in human healthcare. *Archivum Immunologiae et Therapiae Experimentalis*, 55: 315-327, 2007.

PAGE, J. E.; HAUSE, G.; RASCHKE, M.; GAO, W.; SCHMIDT, J.; ZENK, M. H.; KUTCHAN, T. M. Functional Analysis of the Final Steps of the 1-Deoxy-D-xylulose 5-phosphate (DXP) Pathway to Isoprenoids in Plants Using Virus-Induced Gene Silencing. *Plant Physiology*, 134: 1-13, 2004.

PANG, Y.; SHEN, G.; BERGE, T.; WU, C. W.; SUN, X.; TANG, K. Molecular cloning, characterization and heterologous expression in *Saccharomyces cerevisiae* of a mevalonato diphosphate decarboxylase cDNA from *Ginkgo biloba*. *Physiologia Plantarum*, 127: 19-27, 2006.

PASQUALI G, PORTO DD, FETT-NETO AG. Metabolic engineering of cell cultures *versus* whole-plant complexity in the production of bioactive monoterpene indole

alkaloids: recent progress related to an old dilemma. *Journal of Bioscience and Bioengineering*, 101: 287-96, 2006.

PELLECCHIA, M.; MEININGER, D.; DONG, Q.; CHANG, E.; JACK, R.; SEM D. S. NMR-based structural characterization of large protein-ligand interactions. *Journal of Biomolecular NMR*, 22: 165-173, 2002.

PHILLIPS, M. A.; WALTER, M. H.; RALPH, S. G.; DABROWSKA, P.; LUCK, K.; UROS, E. M.; BOLAND, W.; STRACK, D.; RODRIGUEZ-CONCEPCION, M.; BOHLMANN, J.; GERSHENZON, J. Functional identification and differential expression of 1-deoxy-d-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (*Picea abies*). *Plant Mol. Biol.*, 65: 243-257, 2007.

PICHERSKY, E.; GANG, D. R. Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends in Plant Science*, 5: 439-445, 2000.

RODRÍGUEZ-CONCEPCIÓN, M. Early steps in isoprenoid biosynthesis: Multilevel regulation of the supply of common precursors in plant cells. *Phytochemistry Reviews*, 5: 1–15, 2006.

RODRÍGUEZ-CONCEPCIÓN, M.; BORONAT, A. Elucidation of the Methylerythritol Phosphate Pathway for Isoprenoid Biosynthesis in Bacteria and Plastids. A Metabolic Milestone Achieved through Genomics. *Plant Physiology.*, *130*: 1079-1089, 2002.

RODRÍGUEZ-CONCEPCIÓN, M.; AHUMADA, I.; DIEZ-JUEZ, E.; SAURET-GUETO, S.; LOIS, L. M.; GALLEGO, F.; CARRETERO-PAULET, L.; CAMPOS, N.; BORONAT, A. 1-Deoxy-d-xylulose 5-phosphate reductoisomerase and plastid isoprenoid biosynthesis during tomato fruit ripening. *The Plant Journal*, 27: 213–222, 2001.

RODWELL, V. W.; BEACH, M. J.; BISCHOFF, K. M. 3- Hidroxi-3-metilglutaril-CoA reductase. *Methods in Enzymology*. 324: 259-280, 2000.

ROHDICH, F.; WUNGSINTAWEEKUL, J.; LUTTGEN, H.; FISCHER, M.; EISENREICH, W.; SCHUHR, C. A.; FELLERMEIER, M.; SCHRAMEK, N.; ZENK, M. H.; BACHER, A. Biosynthesis of terpenoids: 4-diphosphocytidyl-2-C-methyl-d-erythritol kinase from tomato. *Proceedings of the National Academy of Sciences*, 97: 8251–8256, 2000.

ROHDICH, F.; WUNGSINTAWEEKUL, J.; FELLERMEIER, M.; SAGNER, S.; HERZ, S.; KIS, K.; EISENREICH, W.; BACHER, A.; ZENK, M. H. Cytidine 5´-triphosphate-dependent biosynthesis of isoprenoids: YgbP protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-*C*-methyl-Derythritol. *Proceedings of the National Academy of Sciences*, 96: 11758–11763, 1999.

ROHMER, M. The discovery of a mevalonate independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plant. *Natural Product Reports*, 16: 565-574, 1999.

ROHMER, M. Mevalonate-independent methylerythritol phosphate pathway for isoprenoid biosynthesis. Elucidation and distribution. *Pure and Applied Chemistry*, 75: 375–387, 2003.

ROHMER, M.; SEEMANN, M.; HORBACH, S.; BRINGER-MEYER, S.; SAHM H. Glyceraldehyde 3-phosphate and pyruvate as precursors of isoprenic units in an alternative pathway for terpenoid biosynthesis. *Journal of the American Chemical Society*, 118: 2564–2566, 1996.

SACCHETTINI, J. C.; POULTER, C. D. Creating isoprenoid diversity. *Science*, 277: 1788–1789, 1997.

SASAKI, T.; MATSUMOTO, T. The genome sequence and structure of rice chromosome 1. *Nature*, 420: 312–316, 2002.

SATO, S.; NAKAMURA, Y.; KANEKO, T.; KATOH, T.; ASAMIZU, E.; TABATA, S. Structural analysis of *Arabidopsis thaliana* chromosome 3. I. Sequence features of

the regions of *4, 504*, 864 bp covered by sixty P1 and TAC clones. *DNA Research*, 7: 131-135, 2000.

SCHWENDER, J.; MULLER, C.; ZEIDLER, J.; LICHTENTHALER, H. K. Cloning and heterologous expression of a cDNA encoding 1-deoxy-d-xylulose-5-phosphate reductoisomerase of *Arabidopsis thaliana*. *FEBS Letters*, 455: 140–144, 1999.

SEETANG-NUN, Y.; SHARKEY, T. D.; SUVACHITTANONT, W. Isolation and characterization of two distinct classes of DXS genes in *Hevea brasiliensis. DNA Sequence*, 19: 291-300, 2008a.

SEETANG-NUN, Y.; SHARKEY, T. D.; SUVACHITTANONT, W. Molecular cloning and characterization of two cDNAs encoding 1-deoxy-d-xylulose 5-phosphate reductoisomerase from *Hevea brasiliensis*. *Journal of Plant Physiology*, 165: 991-1002, 2008b.

SEKI, M.; IIDA, K.; SATOU, M.; SAKURAI, T.; AKIYAMA, K.; ISHIDA, J.; NAKAJIMA, M.; ENJU, A.; KAMIYA, A.; NARUSAKA, M.; CARNINCI, P.; KAWAI, J.; HAYASHIZAKI,Y.; SHINOZAKI, K. *Arabidopsis thaliana* full-length cDNA. Published Only in Database. http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi, 2002.

SHARKEY, T. D.; YEH, S.; WIBERLEY, A. E.; FALBEL, T. G.; GONG, D.; FERNANDEZ, D. E. Evolution of the Isoprene Biosynthetic Pathway in Kudzu. *Plant Physiology.*, 137: 700-712, 2005.

SPRENGER, G. A.; SCHÖRKEN, U.; WIEGERT, T.; GROLLE, S.; DE GRAAF, A. A.; TAYLOR, S. V.; BEGLEY, T. P.; BRINGER-MEYER, S.; SAHM, H. Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin and pyridoxol. *Proceedings of the National Academy of Science*, 94**:** 12857-12862, 1997.

STEINBACHER, S.; KAISER, J.; EISENREICH, W.; HUBER, R.; BACHER, A.; ROHDICH, F. Structural basis of fosmidomycin action revealed by the complex with IspC: implications for the catalytic mechanism and anti-malaria drug development. *The Journal of Biological Chemistry*, 278: 18401-18407, 2003.

SUMNER, L. W.; HUHMAN, D. V.; URBANCZYK-WOCHNIAK, E.; LEI, Z. Methods, applications and concepts of metabolite profi ling: Secondary metabolism. *Plant Systems Biology*, 2007.

TAIZ, L.; ZEIGER, E. Fisiologia Vegetal. 3ª. edição, Porto Alegre, Artmed, 2004.

TAKAHASHI, S.; KUZUYAMA, T.; WATANABE, H.; SETO, H. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proceedings of the National Academy of Sciences*, 95: 9879-9884, 1998.

TOTTE, N. M. L. C.; VAN DEN ENDE, W.; VAN DAMME, E. J. M.; COMPERNOLLE, F.; BABOEUF, I.; GEUNS, J.M.C. Cloning and heterologous expression of early genes in gibberellin and steviol biosynthesis via the methylerythritol phosphate pathway in *Stevia rebaudiana*. *Plant Physiology*, 81: 517-522, 2003.

VEAU, B.; COURTOIS, M.; OUDIN, A.; CHENIEUX, J. C.; RIDEAU, M.; CLASTRE, M. Cloning and expression of cDNAs encoding two enzymes of the MEP pathway in *Catharanthus roseus*. *Biochimica et Biophysica Acta*, 1517: 159-163, 2000.

VERPOORTE, R.; MEMELINK, J. Engineering secondary metabolite production in plants. *Current Opinion in Biotechnology*, 13: 181-187, 2002.

VERPOORTE, R.; VAN DER HEIJDEN, R.; MEMELINK, J. Engineering the plant cell factory for secondary metabolite production. *Transgenic Research*, 9: 323-343, 2000.

VITTI, A. M. S.; BRITO, J. O. Óleo essencial de eucalipto. Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz. Documentos Florestais, n. 17, 2003.

WU, S. J.; SHI, M.; WU, J. Y. Cloning and characterization of 1-deoxy-d-xylulose 5-phosphate reductoisomerase gene for diterpenoid tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots. *Biotechnology and Applied Biochemistry*, 52: 89-95, 2009.

WUNGSINTAWEEKUL, J.; SIRISUNTIPONG, T.; KONGDUANG, D.; LOSUPHANPORN, T.; OUNAROON, A.; TANSAKUL, P.; AND DE-EKNAMKUL, W. Transcription profiles analysis of genes encoding 1-deoxy-d-xylulose 5-phosphate synthase and 2C-methyl-d-erythritol 4-phosphate synthase in plaunotol biosynthesis from *Croton stellatopilosus*. *Biological and Pharmaceutucal Bulletin*, 31: 852-856, 2008.

XIAO, Y.; ZAHARIOU, G.; SANAKIS, Y.; LIU, P. IspG Enzyme Activity in the Deoxyxylulose Phosphate Pathway: Roles of the Iron-Sulfur Cluster. *Biochemistry*, 48: 10483-10485, 2009.

YAJIMA, S.; NONAKA, T.; KUZUYAMA, T.; SETO, H.; OHSAWA K. Crystal structure of 1-deoxy-D-xylulose 5-phosphate reductoisomerase complexed with cofactors: implifications of a flexible loop movement upon substrate binding. *The Journal of Biochemistry*, 131: 313-317, 2002.

YAO, H.; GONG, Y.; ZUO, K.; LING, H.; QIU, C.; ZHANG, F.; WANG, Y.; PI, Y.; LIU, X.; SUN, X.; TANG, K. Molecular cloning, expression profiling and functional analysis of a DXR gene encoding 1-deoxy-d-xylulose 5 phosphate reductoisomerase from *Camptotheca acuminata*. *Journal of Plant Physiology*, 165: 203-213, 2008.

YU, J.; WANG, J.; LIN, W.; LI, S.; LI, H. The Genomes of *Oryza sativa*: A History of Duplications. *PLoS Biology*, 3: E38-E42, 2005.

ZHANG, C.; LIU, L.; XU, H.; WEI, Z.; WANG, Y.; LIN, Y.; GONG, W. Crystal Structures of Human IPP Isomerase: New Insights into the Catalytic Mechanism. *Journal of Molecular Biology*, 366: 1437-1446, 2007.

ZHENG, W.; SUN, F.; BARTLAM, M.; LI, X.; LI, R.; RAO, Z. The Crystal Structure of Human Isopentenyl Diphosphate Isomerase at 1.7 Å Resolution Reveals its Catalytic Mechanism in Isoprenoid Biosynthesis. *Journal of Molecular Biology*, 366: 1447-1458, 2007.

ZHONG, J-J.; YUE, C-J. Plant Cells: Secondary Metabolite Heterogeneity and Its Manipulation. *Advances in Biochemical and Engineering/Biotechnology*, 100: 53-88, 2005.

ZINI, C. A. Estudo dos compostos voláteis de alguma espécies de eucalipto através do uso de micro extração de em fase sólida no modo headspace (HS-SPME). Porto Alegre: UFRGS, Instituto de Química, Tese (Doutorado), 2002.

CURRICULUM VITAE RESUMIDO

ATHAYDES, G. A.

1. DADOS PESSOAIS

Nome: Genaro Azambuja Athaydes

Local e data de nascimento: São Borja, Rio Grande do Sul, Brasil, 29 de julho de 1981.

Endereço profissional: Av. Bento Gonçalves, 9500, Centro de Biotecnologia, Laboratório de Biologia Molecular Vegetal, prédio 3431.

Telefone Profissional: (51) 3308.6094 **E-mail:** genaroathaydes@gmail.com

2. FORMAÇÃO: Farmácia (Universidade Federal do Rio Grande do Sul, 2002-2007).

3. ESTÁGIOS:

3.1.Grupo de Erros Inatos do Metabolismo (Bioquímica – UFRGS)
Orientação: Prof. Dr. Clóvis Milton Duval Wannmacher
Período: 2003-2005

3.2.Fundação de Ciência e Tecnologia (CIENTEC)/Departamento de Alimentos
Supervisão da Farmacêutica: M.Sc. Eliane Rossoni
Período: 2005-2007

4. ARTIGOS COMPLETOS PUBLICADOS

RECH, V. C.; **ATHAYDES, G. A.**; FEKSA, L. R.; DORNELLES, P. K. B.; RODRIGUES-JUNIOR, V.; DUTRA-FILHO, C. S.; WYSE, A. T. S.; WAJNER M.; WANMACHER C.M.D. Inhibition of Creatine Kinase Activity by Cystine in the Kidney of Young Rats. *Pediatric Research*, 60: 190-195, 2006.

RECH, V. C.; FEKSA, L. R.; FLECK, R. M. M.; **ATHAYDES, G. A.**; DORNELLES, P. K. B.; RODRIGUES-JUNIOR, V.; WANMACHER C.M.D.

Cysteamine prevents inhibition of thiol-containing enzymes caused by cystine or cystine dimethylester loading in rat brain cortex. *Metabolic Brain Disease*, 23: 133-145, 2008.