

杜仲 *HDR* 基因全长 cDNA 克隆与序列分析

刘攀峰¹, 杜红岩^{1*}, 乌云塔娜², 杜兰英¹, 孙志强¹

(1. 中国林业科学研究院经济林研究开发中心,河南 郑州 450003; 2. 国家林业局经济林育种与栽培重点实验室,湖南 长沙 410004)

摘要:以杜仲叶片 cDNA 为模板,采用反转录 RCR 及 RACE 技术分离出 HDR 基因的 cDNA 克隆,命名为 *EuHDR*。*EuHDR* 基因 cDNA 全长 1 653 bp,5' 端非编码区长 82 bp,3' 端非编码区长 188 bp,编码 460 个氨基酸,与喜树 *HDR* 基因序列相似性最高,达 82%;推导 *EuHDR* 氨基酸序列中包含转运肽序列(A1-A33)及植物 *HDR* 蛋白多个保守的功能位点(A117, A208, A262, A345);*EuHDR* 蛋白二级结构 α -螺旋占 35.65%, β -折叠占 19.78%,螺环结构占 44.57%;*EuHDR* 蛋白三级结构为单体形式,呈不规则的三叶草形状;系统进化分析表明 *EuHDR* 蛋白与葡萄 *HDR* 蛋白的亲缘关系最为接近。

关键词:杜仲;*HDR*;基因;序列分析

中图分类号:S718.4 文献标识码:A

Cloning and Sequence Analysis of 1-Hydroxy-2-Methyl-2-E-Butenyl-4-Diphosphate Reductase Gene cDNA from *Eucommia ulmoides*

LIU Pan-feng¹, DU Hong-yan¹, WUYUN Ta-na², DU Lan-ying¹, SUN Zhi-qiang¹

(1. Non-timber Forestry Research and Development Center, Chinese Academy of Forestry, Zhengzhou 450003, He'nan, China;

2. Key Laboratory of Non-wood Forest Products of State Forestry Administration, Changsha 410004, Hu'nan, China)

Abstract: 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase (*HDR*) synthesizes IPP and DMAPP in the last step of the plant 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway. Homologous *HDR* gene cDNA was isolated from the leaves of *Eucommia ulmoides* by the method of reverse transcription polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) technique, and named as *EuHDR*. With the highest gene sequence similarity to *Camptotheca acuminata* (82%), the full-length cDNA of *EuHDR* was 1 653 bp including 5' non-coding region of 82 bp and 3' non-coding region of 188 bp and encoded 460 amino acids. The transit peptide sequence (A₁-A₃₃) and multiple conserved functional sites (A₁₁₇, A₂₀₈, A₂₆₂, A₃₄₅) of plant *HDR* protein were found in the deduced coding sequence of *EuHDR*. The secondary structure of *EuHDR* protein was predicted with proportion of α -helix to 35.65%, β -sheet to 19.78% and loop/coil to 44.57%. The calculated protein tertiary structure of *EuHDR* was formed as monomer, which in space displayed asymmetrical shamrock-like shape. Phylogenetic analysis revealed that the evolutionary relationship of *EuHDR* protein was the closest to *Vitis vinifera* *HDR* protein.

Key words: *Eucommia ulmoides*; *HDR* gene; sequence analysis

萜(terpenoid)是结构最为古老和种类最为繁多种,因而此类化合物也统称为萜组^[1-2]。植物萜类的生物大分子,自然界中目前已发现有 65 000 多种作为重要的次生代谢产物,具有驱虫、杀虫、吸引授

收稿日期:2012-08-28

基金项目:国家公益性行业科研专项(201004029)

作者简介:刘攀峰(1984—),男,河南平顶山人。研究方向:杜仲栽培与育种;Email:pfengliu@126.com

* 通讯作者,研究员,博士生导师。主要从事杜仲育种、栽培与综合利用的研究

粉者、激素、香料以及染料等功能,在日常生活中广泛开发和应用在食品、保健品、化妆品、化工、医药以及橡胶等行业^[3]。异戊烯基焦磷酸(IPP)及其异构体二甲基烯丙基焦磷酸(DMAPP)是萜类化学组成与生物合成的基本结构单位和功能单位,植物中二者的形成主要经定位于胞质的甲羟戊酸途径(MVA pathway)和质体的2-甲基-D-赤藓醇-4-磷酸途径(MEP pathway)^[4]。1-羟基-2-甲基-2-E-丁烯基-4-焦磷酸还原酶(HDR)是MEP途径的末端酶,催化1-羟基-2-甲基-2-E-丁烯基-4-焦磷酸(HMBPP)发生还原反应并以5:1的比例生成IPP与DMAPP的混合物^[5]。有近5千余条与HDR相关的基因序列在GenBank中登录,包括细菌、植物及其它真核生物等,其中银杏(*Ginkgo biloba* L.)、马尾松(*Pinus massoniana* Lamb.)、苏铁(*Cycas revoluta* Thunb.)等裸子植物均有2~3个基因家族成员^[6]。HDR基因在萜类合成过程中起重要的调控作用^[7-11],被认为是萜类代谢工程的理想靶点,其基因结构、生物学功能与作用机理日益受到萜类研究的关注。

杜仲(*Eucommia ulmoides* Oliv.)适生于华中、华西、西南及西北各地,现广泛栽培^[12-13]。许多杜仲提取物在工业领域以及医药生产上有重要的经济价值,其中最重要的是萜类次生产物。目前对杜仲萜类生物合成、功能基因开发和利用开展的研究不多,对杜仲HDR基因的克隆及序列分析尚无报道。本研究以杜仲叶片为材料,分离HDR同源基因全长cDNA,通过生物信息学方法对基因序列及推导的氨基酸序列进行分析,以期研究杜仲HDR基因功能、阐释杜仲萜类生物合成机制和分子育种提供参考。

1 材料与方 法

1.1 材料

2011年4月下旬于中国林业科学研究院经济林研究开发中心院内采集‘华仲6号’杜仲良种叶片,清洗干净后投入液氮带回室内-80℃保存备用。

1.2 试剂

DEPC(Sigma,德国),PVP(Amresco,美国),SDS(上海生工,上海),氯化锂(Amresco,美国),CTAB(Amresco,美国),MightyAmp DNA Polymerase Ver. 2(Takara,大连),琼脂糖凝胶DNA回收试剂盒(天根,北京),3'-Full RACE Core Set(Takara,大连),

5'-Full RACE Kit(Takara,大连),M-MLV RTase cDNA Synthesis Kit(Takara,大连),pEASY-T1 Cloning Kit(全式金,北京),DH5 α 感受态细胞(天根,北京)。

1.3 试验方法

1.3.1 总RNA提取与单链cDNA的合成 采用改良的CTAB-LiCl法提取杜仲叶片RNA^[14-15],对满足实验要求的RNA样品保存于-80℃冰箱备用。按M-MLV RTase cDNA Synthesis Kit说明书进行单链cDNA的合成。

1.3.2 引物设计 根据一段已知杜仲HDR unigene序列,结合试剂盒锚定引物序列,设计适于3'RACE的巢式扩增的引物3P1:5'-TACGGTGAATAAGAG-GTTAGAGG-3',3P2:5'-ATCTTGCCTGCTTTCCGGT-GCTG-3'以及5'RACE巢式扩增引物5P1:5'-GTT-GAGCTAGATCCATCAAGTTGAC-3',5P2:5'-CTAT-AGTCTCCTCGTGCGAATAT-3'。

1.3.3 基因全长cDNA末端扩增 3'RACE及5'RACE的PCR反应体系与反应条件参照Takara 3'-Full RACE Core Set Ver. 2.0与Takara 5'-Full RACE Kit说明书操作。

1.3.4 目的片段回收与测序 按TIANGEN通用型DNA回收试剂盒说明进行目的PCR产物的回收;按pEASY-T克隆试剂盒说明将回收片段连接至克隆载体,鉴定后将阳性克隆送至南京金斯瑞公司测序。

1.3.5 生物信息学分析 利用NCBI Blast程序进行序列相似性检索,并用ORF Finder程序查找基因cDNA开放阅读框架;利用ExPASy ProtParam程序与ScanProsite程序分析氨基酸残基数目与组成、蛋白质相对分子量、理论等电点以及功能位点等;利用Predict Protein以及PSIPRED方法预测蛋白质的细胞定位及二级结构^[16];利用ChloroP 1.1 Server进行转运肽的预测^[17];利用SWISS-MODEL程序进行蛋白质的同源建模^[18];利用Lasergene软件进行蛋白质序列的多重比对并通过MEGA 5软件构建基因的系统进化树^[19]。

2 结果与分析

2.1 杜仲HDR基因全长cDNA的分离及序列特征

利用3'RACE与5'RACE端巢式扩增引物在逆转录的cDNA模板上分别扩增出1条1000bp左右和1条800bp左右的目的条带(图1),拼接后得到

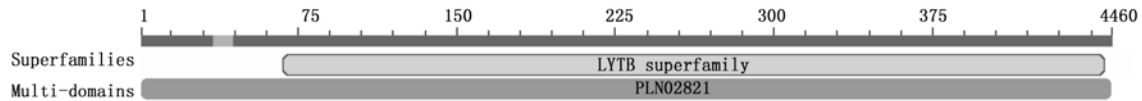


图5 EuHDR 蛋白的保守结构域预测

2.2.3 EuHDR 蛋白功能位点与翻译后磷酸化修饰预测 ExPaSy ScanProsite 程序分析 EuHDR 蛋白基序类型分为 5 种,含 29 个潜在的功能位点,包括 9 个酪蛋白激酶 II 磷酸化位点 (TstD, 12 ~ 15; TlpE, 17 ~ 20; SsgE, 35 ~ 38; TefD, 49 ~ 52; TlsE, 194 ~ 197; SheE, 234 ~ 237; StkE, 274 ~ 277; TicD, 342 ~ 345; StpD, 435 ~ 438); 6 个蛋白激酶 C 磷酸化位点 (SvR, 31 ~ 33; TvK, 106 ~ 108; SeK, 135 ~ 137; SeK, 196 ~ 198; StK, 274 ~ 276; SyK, 406 ~ 408); 5 个 N 端糖基化位点 (NLTR, 59 ~ 62; NVTV, 104 ~ 107; NQTT, 304 ~ 307; NSSN, 372 ~ 375; NTSH, 375 ~ 378); 8 个 N 端豆蔻酰化位点 (GNeyTW, 97 ~ 102; GVerAV, 118 ~ 123; GAavNE, 186 ~ 191; GQldGS, 267 ~ 272; GSssTK, 271 ~ 276; GGwnSS, 369 ~ 374; GNkiSY, 402 ~ 407; GVtsGA, 429 ~ 434); 1 个酪氨酸激酶磷酸化位点 (KnmaEam. Y, 252 ~ 259)。

NetPhos 2.0 server 预测出 27 个 EuHDR 蛋白翻译后磷酸化位点,包括 11 个丝氨酸磷酸化位点 (A_{31} 、 A_{35} 、 A_{36} 、 A_{39} 、 A_{42} 、 A_{135} 、 A_{227} 、 A_{234} 、 A_{274} 、 A_{432} 、 A_{435}), 10 个苏氨酸磷酸化位点 (A_{12} 、 A_{17} 、 A_{49} 、 A_{86} 、 A_{106} 、 A_{194} 、 A_{205} 、 A_{342} 、 A_{427} 、 A_{436}) 以及 6 个酪氨酸磷酸化位点 (A_{100} 、 A_{225} 、 A_{248} 、 A_{259} 、 A_{285} 、 A_{407}) (图 6)。

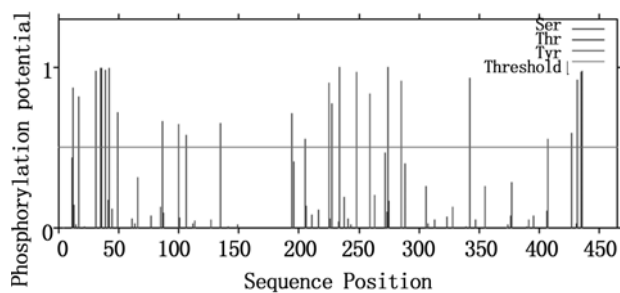


图6 EuHDR 蛋白翻译后磷酸化位点预测

2.2.4 EuHDR 蛋白三级结构分析 以嗜嗜热菌 (*Aquifex aeolicus*) HDR 蛋白 (3dnf) 为模板对 EuHDR 蛋白进行同源建模,如图 7 所示, EuHDR 蛋白三维结构为单体形式,在空间上呈不规则的三叶草形状。ExPaSy structure assessment 程序评测推导蛋白模型 QMEAN 4 得分为 0.503,与模板蛋白序列的相似性为 24.5%。

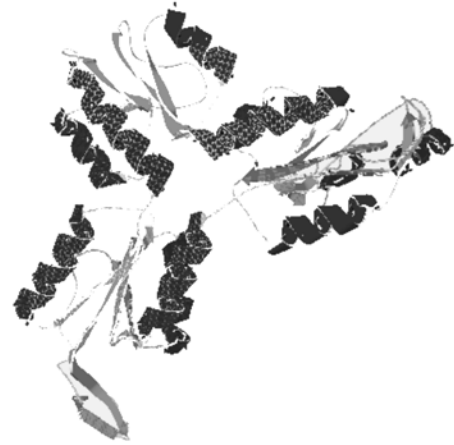


图7 EuHDR 蛋白三级结构预测

2.3 EuHDR 同源蛋白系统进化树的构建

利用 MEGA5 中 Clustal W 法对 22 个不同物种 HDR 蛋白序列进行比对,并用 Neighbor-joining 法构建了系统进化树 (图 8)。结果表明不同来源的 HDR 蛋白属不同的进化分支,裸子植物、单子叶植物及双子叶植物 HDR 之间有较为清晰的界限。EuHDR 蛋白与葡萄亲缘关系最为接近,进化距离为 0.084,其次为橡胶 (0.087)、喜树 (0.091)、番茄 (*Lycopersicon esculentum* Miller; 0.091)、萝芙木 (0.098) 和长春花 (0.098)。

3 讨论

萜类存在于近乎所有的生命形式中,生理功能涉及蛋白合成调节、电子转移、光合作用、质膜形成、信号传递以及植物防御等多个方面^[20]。分析代谢网络中的限速步骤,发掘鉴定关键功能基因,是应用基因工程以及分子育种的前提条件。相对于 MVA 途径,MEP 合成途径更广泛地存在于自然界中。DXS 和 DXR 酶基因最早被认为是 MEP 途径中重要的调控靶点,并在大肠杆菌、拟南芥 (*Arabidopsis thaliana* L.) 及番茄等生物系统中开展了一系列研究对其基因功能进行验证^[21-22]。而近期的一些研究表明 HDR 基因在 MEP 途径 IPP 与 DMAPP 合成的代谢流动中同样发挥关键作用,如在萜类物质含量很高的薄荷腺体中发现大量分布 HDR 基因表达的序列标签^[7];在模式物种拟南芥、大肠杆菌中也证

实 *HDR* 基因对萜类化合物的合成产生重要影响^[11];在银杏 *HDR* 基因转化系中银杏总内酯含量

是非转基因系的 2.5 倍,进一步说明了 *HDR* 对萜类合成有重要作用^[10]。基于生物信息学分析的结果

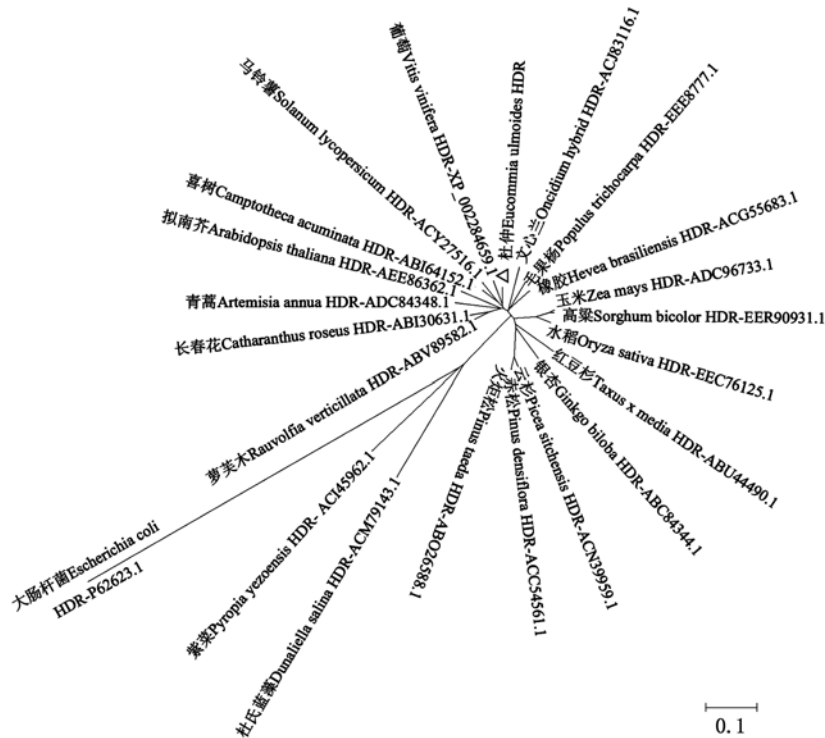


图8 HDR同源蛋白的系统进化树分析

表明 *EuHDR* 蛋白具有植物 *HDR* 典型的功能位点、基序及结构域,意味着所克隆的 *EuHDR* 是高等植物 *HDR* 基因家族的新成员,可为杜仲萜类基因工程候选基因的筛选提供基础信息,但对基因功能的判定有待深入研究。

目前 PDB 数据库中已注册 50 种 *HDR* 蛋白的晶体模型,包括细菌、古生菌、真核生物以及病毒等类型。利用 X-射线对大肠杆菌和嗜热菌分析的数据显示 *HDR* 蛋白由一条肽链折叠形成三个相似的结构域,在空间上排列成不完全对称的三叶草形状^[23]。而通过 Swiss-model 同源建模法预测的 *EuHDR* 蛋白三维结构与已解析的 *HDR* 蛋白空间特征基本吻合,这为 *EuHDR* 蛋白的功能预测以及认识植物 *HDR* 的催化机制提供了初步参考。

参考文献:

- [1] Köksal M, Hu R M, Coates R M, *et al.* Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase [J]. *Nat Chem Biol*, 2011, 7: 431 - 433
- [2] Eric O, Lin F Y. Terpene biosynthesis: modularity rules [J]. *Angew Chem*, 2012, 51(5): 1124 - 1137
- [3] Karine B, Yannick E, Alain D, *et al.* Isopentenyl diphosphate

isomerase: a checkpoint to isoprenoid biosynthesis [J]. *Biochimica*, 2012, doi:10.1016/j.biochi.2012.03.021; 1 - 14.

- [4] Daum M, Herrmann S, Wilkinson B, *et al.* Genes and enzymes involved in bacterial isoprenoid biosynthesis [J]. *Curr Opin Chem Biol*, 2009, 13: 180 - 188
- [5] Annegret A B, Karnjapan J, Juliusz A W, *et al.* Biosynthesis of isoprene units: mssobauer spectroscopy of substrate and inhibitor binding to the [4Fe-4S] cluster of the *LytB/IspH* enzyme [J]. *Angew Chem*, 2011, 50: 11976 - 11979
- [6] Sang-Min K, Tomohisa K, Akio K, *et al.* 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase (IDS) is encoded by multicopy genes in Gymnosperms *Ginkgo biloba* and *Pinus taeda* [J]. *Planta*, 2008, 227: 287 - 298
- [7] Cunningham Jr F X, Lafond T P, Gantt E. Evidence of a role for *LytB* in the nonmevalonate pathway of isoprenoid biosynthesis [J]. *J Bacteriol*, 2000, 182: 5841 - 5848
- [8] Guevara-Garcia A, San Roman C, Arroyo A, *et al.* Characterization of the *Arabidopsis clb6* mutant illustrates the importance of posttranscriptional regulation of the methyl-D-erythritol 4-phosphate pathway [J]. *Plant Cell*, 2005, 17: 628 - 643
- [9] Hsieh M H, Goodman H M. The *Arabidopsis IspH* homolog is involved in the plastid nonmevalonate pathway of isoprenoid biosynthesis [J]. *Plant Physiol*, 2005, 138: 641 - 653
- [10] 张雯, 王玉亮, 林娟, 等. 银杏 1-羟基-2-甲基-2-(E)-丁烯基-4-焦磷酸还原酶基因 (*hdr*) 转化银杏的研究 [J]. *复旦学报*,

- 2008,47(5):598-602
- [11] Botella-Pavía P, Óscar B, Phillips M A, *et al.* Regulation of carotenoid biosynthesis in plants: evidence for a key role of hydroxymethylbutenyl diphosphate reductase in controlling the supply of plastidial isoprenoid precursors[J]. *Plant J*,2004, 40(2):188-199
- [12] 崔克明. 杜仲研究的历史、现状和展望[J]. *西北林学院学报*, 1994,9(4):51-57
- [13] 杜红岩. 杜仲优质高产栽培[M]. 北京:中国林业出版社, 1996
- [14] 陈建. 几种提取杜仲 RNA 方法的比较[J]. *林业科技开发*, 2007,21(5):19-21
- [15] 周明兵,王红珍,赵德刚. 杜仲叶和树皮总 RNA 的快速提取法[J]. *山地农业生物学报*,2003,22(5):430-431
- [16] Jones D T. Protein secondary structure prediction based on position-specific scoring matrices [J]. *J Mol Biol*, 1999, 292(2):195-202
- [17] Emanuelsson O, Nielsen H, von Heijne G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites[J]. *Protein Sci*,1999,8: 978-984
- [18] Arnold K, Bordoli L, Kopp J, *et al.* The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling[J]. *Bioinformatics*, 2006,22:195-201
- [19] Tamura K, Peterson D, Peterson N, *et al.* MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J]. *Mol Biol Evol*, 2011,28: 2731-2739
- [20] 高文运. 同位素标记及未标记的 1-脱氧-D-木酮糖-5-磷酸和 2-甲基-D-赤藓糖醇-4-磷酸的合成方法[J]. *有机化学*,2010,30(1):23-37
- [21] 金蓉,朱长青,徐昌杰. 1-脱氧木酮糖-5-磷酸合成酶(DXS)及其编码基因[J]. *细胞生物学杂志*,2007,29:706-712
- [22] 郑洲翔,范燕萍,周纪刚,等. 植物萜类合成酶 1-脱氧-D-木酮糖-5-磷酸还原酶研究进展[J]. *安徽农业科学*, 2011,39(10):5695-5696
- [23] Tobias G, Ingrid S, Wolfgang E, *et al.* Probing the reaction mechanism of *IspH* protein by X-ray structure analysis[J]. *PNAS*,2010, 107(3):1077-1081