

植物法呢基焦磷酸合成酶的生物信息学分析

李 嵘, 王喆之

(陕西师范大学生命科学院, 西安 710062)

摘要:采用生物信息学的方法和工具对已在 GenBank 上注册的番茄 (*Lycopersicon esculentum*)、橡胶 (*Hevea brasiliensis*)、香蕉 (*Musa acuminata*)、苹果 (*Malus × domestica*)、向日葵 (*Helianthus annuus*)、葡萄 (*Vitis vinifera*) 等的法呢基焦磷酸合成酶的核苷酸及氨基酸序列进行分析, 并对其组成成分、转运肽、跨膜拓扑结构域、疏水性/亲水性、蛋白质二级及三级结构、分子系统进化关系等进行预测和推断。结果表明: 该酶基因的全长包括 5'、3' 非翻译区和一个开放阅读框; 其编码的蛋白是一个无跨膜结构域、不具转运肽的亲水性蛋白, 包括 5 个保守的结构 motif, 其中两个是富含天冬氨酸 (Asp) 的 motif, α -螺旋和不规则盘绕是蛋白质二级结构最大量的结构元件, β -转角和延伸链散布于整个蛋白质中, 蛋白质的功能域在空间布局上是由 α -螺旋围绕形成的中间具一个“大空穴”的立体结构, “大空穴”的表面藏有五个功能保守 motif, 其中两个 Asp-motif 位于“空穴”的内壁。

关键词: 植物萜类合成; 法呢基焦磷酸合成酶; 生物信息学

中图分类号: Q71

文献标识码: A

文章编号: 1005-3395(2007)02-0126-09

A Bioinformatics Analyses on Farnesyl Diphosphate Synthase

LI Rong, WANG Zhe-zhi

(College of Life Sciences, Shaanxi Normal University, Xi'an 710062, China)

Abstract: The nucleic acid sequences and amino acid sequences of farnesyl diphosphate synthase (FPPS) from *Lycopersicon esculentum*, *Hevea brasiliensis*, *Musa acuminata*, *Malus × domestica*, *Helianthus annuus* and *Vitis vinifera*, which were registered in GenBank, were analyzed and predicted by the tools of bioinformatics in the following aspects: the composition of nucleic acid sequences and amino acid sequences, transit peptides, transmembrane topological structures, hydrophobicity or hydrophilicity, secondary and tertiary structure of protein, molecular phylogenetic evolution. The full-length gene of FPPS contains an opening reading frame, a 5'-untranslated region, and a 3'-untranslated region. FPPS is a hydrophilic and non-transmembrane protein without transit peptides at its N-terminal end. The amino acid sequences of FPPS include 5 conserved structural regions, out of which two are Asp-rich motifs. α -helix and random coil are the major motifs of predicted secondary structure of FPPS. β -turn and extended strands spread in the secondary structure of the protein. A large central cavity formed from the α -helices is present within the predicted tertiary structure of functional regions of FPPS. Five conserved amino acid sequences are on the surface of this cavity, two of which are conserved Asp-rich motifs on the inner walls of the cavity.

Key words: Plant isoprenoid biosynthesis; Farnesyl diphosphate synthase (FPPS); Bioinformatics

萜类化合物 (terpenoids) 是一类由异戊二烯 (isoprene) 为结构单元的天然烃类化合物, 分布广、

种类多、结构多样。迄今, 大约有 22 000 种萜及萜类衍生物的化学结构已被鉴定^[1-3]。据估计, 植物体内

收稿日期: 2006-08-22 接受日期: 2006-10-30

基金项目: 国家“十五”科技攻关项目 (2004BA721A35) 资助

的萜类化合物有 30 000 种之多^[9],它们不仅在植物的生命活动中发挥重要作用,而且还被广泛应用于工业、医药卫生等领域。植物体内萜类化合物的合成有甲羟戊酸 (mevalonate pathway, MVA) 和丙酮酸 / 磷酸甘油醛 (pyruvate/glyceraldehydes-3-phosphate pathway, DXP) 两条途径,两者都以异戊烯基焦磷酸 (isopentenyl diphosphate, IPP) 为主要中间产物^[5-10]。法呢基焦磷酸合成酶 (farnesyl diphosphate synthase, FPPS) 属于短链烯丙基转移酶 (prenyltransferase) 家族,在植物萜类合成中催化牻牛儿基焦磷酸 (geranyl diphosphate, GPP) 与异戊烯基焦磷酸 (IPP) 或二甲基烯丙基焦磷酸 (dimethylallyl diphosphate, DMAPP) 缩合成为法呢基焦磷酸 (farnesyl diphosphate, FPP),该酶是植物萜类合成途径中的一个关键酶^[11],其催化终产物 FPP 是许多重要的次生代谢产物如甾体、多萜醇、参与线粒体电子传递的泛醌等的生物合成的前体,同时在倍半萜环化酶的作用下,FPP 还可生成包括植保素在内的倍半萜衍生物^[9]。目前 FPPS 酶基因已从树棉 (*Gossypium arboreum*)、银胶菊 (*Parthenium argentatum*)、辣椒 (*Capsicum annuum*)、水稻 (*Oryza sativa*)、拟南芥 (*Arabidopsis thaliana*)、山艾树 (*Artemisia tridentata* subsp. *spiciformis*)、薄荷 (*Mentha × piperita*)、黄花蒿 (*Artemisia annua*)、白羽扇豆 (*Lupinus albus*)、玉米 (*Zea mays*) 等植物中分离克隆^[12-21]。

生物信息学是一门由分子生物学和信息科学技术相结合的交叉学科。由于人类基因组计划的完成,相关生物科学的发展也已从基因组测序转向对基因组表达的分析,对蛋白质组结构与功能的预测。利用国际互联网上的各类分子生物学数据库及分析软件对已知的核酸和蛋白质序列进行比对、分析、建立计算模型,从而对其结构、功能进行推断及预测,是当今生物信息学发展的必然潮流。

本文利用生物信息学的方法,以番茄 (*Lycopersicon esculentum*) 为重点,对橡胶 (*Hevea brasiliensis*)、香蕉 (*Musa acuminata*)、苹果 (*Malus × domestica*)、向日葵 (*Helianthus annuus*)、葡萄 (*Vitis vinifera*) 等植物的法呢基焦磷酸合成酶的基因及相应氨基酸序列的理化性质、结构特征、功能及系统演化关系等进行预测和分析,以期今后深入开展该酶的相关研究提供理论依据。以往对

FPPS 的研究为实验性研究,主要集中在基因的调控与表达、表达水平及表达部位等方面,类似该文的研究还未见文献报道。

1 材料和方法

数据资料来源于 National Center for Biotechnology Information (NCBI) 核酸及蛋白质数据库中已注册的法呢基焦磷酸合成酶的核酸序列及其对应的氨基酸序列:番茄 (AF048747、AAC73051)、橡胶 (AY349419、AAQ56011)、苹果 (AY083165、AAM08927)、向日葵 (AF019892、AAC78557)、葡萄 (AY966012、AAX76910)、香蕉 (AF470318、AAL82595)。

依据 Vector NTI Suite 8.0 软件及 <http://www.ncbi.nlm.nih.gov/>、<http://www.ebi.ac.uk/>、<http://www.cbs.dtu.dk/>、<http://cn.expasy.org/> 等网站提供的各类生物信息学软件进行在线分析^[22-23]。核酸及氨基酸序列的组成成分分析、理化性质分析、开放阅读框 (Open reading frame, ORF) 的查找和翻译利用 Vector NTI Suite 8.0 软件及 ProtParam、pI/MV、ORF Finder 在线工具上进行;核酸及氨基酸序列的同源性比对、多序列比对及分子系统进化关系利用 Blast 和 Clustal W 在线工具及 MEGA 3.1 软件完成;蛋白质转运肽的预测、跨膜结构域及亲水性 / 疏水性的分析利用在线工具 TargetP、TMHMM、ProtScale 完成;蛋白质二级及三级结构的预测利用 SOPMA 和 SWISS-MODEL 在线工具完成。

2 结果和分析

2.1 不同植物 FPPS 基因的核酸序列及对应氨基酸序列的组成成分及理化性质分析

依据 Gasteiger 等^[24]的方法,用 ORF Finder、Vector NTI Suite 8.0、ProtParam 和 pI/MV 对番茄、橡胶、苹果、向日葵 FPPS 基因的 cDNA 序列及氨基酸序列进行分析,结果见表 1。其中,番茄 FPPS 基因的全长 cDNA 序列及其推导的氨基酸序列如图 1 所示。从表 1 可以看出,不同植物 FPPS 基因的全长均包括 5'、3' 非翻译区 (Untranslated region, UTR) 和一个开放阅读框,起始密码子均为 atg,终止密码子为 tag 或 taa,开放阅读框的碱基数及所编码的氨基酸残基数基本相同。

表 1 不同植物 FPPS 基因的 cDNA 序列及对应氨基酸序列的组成成分及理化性质分析比较

Table 1 Composition and physical and chemical properties of the cDNA sequences and deduced amino acid sequences of FPPSs from 4 species

	番茄 <i>Lycopersicon esculentum</i>	橡胶 <i>Hevea brasiliensis</i>	苹果 <i>Malus × domestica</i>	向日葵 <i>Helianthus annuus</i>
基因全长 Full-length (bp)	1274 bp	1130 bp	1401 bp	1347 bp
开放阅读框 Open reading frame (bp)	1029 bp	1029 bp	1029 bp	1026 bp
起始位点及密码子 Initiation site and codon (bp)	33, ATG	17, ATG	143, ATG	17, ATG
终止位点及密码子 Termination site and codon (bp)	1061, TAG	1045, TAA	1171, TAG	1042, TAG
5'/3'非翻译区长度 The lengths of the 5' and 3' untranslated regions (bp)	32 bp / 213 bp	16 bp / 85 bp	142 bp / 230 bp	16 bp / 305 bp
推导氨基酸残基数 Number of amino acids	342 aa	342 aa	342 aa	341 aa
分子量 Molecular weight (kDa)	39.33	39.42	39.57	39.06
pI	5.44	5.75	5.36	5.96
主要氨基酸 Major amino acids	Leu (13.45%) Lys (9.94%) Asp (8.77%) Ile (6.73%) Val (6.43%)	Leu (12.87%) Lys (8.77%) Val (7.60%) Glu (7.60%) Asp (6.43%)	Leu (11.70%) Lys (8.48%) Glu (7.89%) Asp (7.31%) Val (7.31%)	Leu (12.32%) Lys (8.50%) Val (7.92%) Asp (7.33%) Ser (6.45%)
酸性/碱性氨基酸比例 Acidic/basic amino acids ratio (%)	14.91%/12.87%	14.04%/12.28%	15.20%/12.57%	13.49%/11.73%
带电氨基酸比例 Charged amino acids (%)	36.84%	35.67%	36.55%	34.60%
极性氨基酸比例 Polar amino acids (%)	23.39%	23.68%	23.68%	24.05%
疏水性氨基酸比例 Hydrophobic amino acids	38.60%	38.01%	35.96%	36.66%
蛋白质不稳定性指数 Instability index	26.24%	35.58%	40.53%	31.94%

pI: 理论等电点 Theoretical isoelectric point

不同植物 FPPS 的分子量、理论等电点、酸性和碱性氨基酸的比例、总氨基酸带电荷的比例、极性氨基酸的比例、疏水性 (Hydrophobicity) 氨基酸的比例基本一致; 含量最丰富的氨基酸基本均为 Leu、Lys、Asp、Val; 除苹果 FPPS 属于不稳定类蛋白质外, 其余均属稳定类蛋白质。

2.2 不同植物 FPPS 的核酸及氨基酸序列的比对分析

用 Blast 程序对番茄 FPPS 的 cDNA 及氨基酸序列进行同源性比对^[25], 结果表明番茄与其它植物 FPPS 的 cDNA 序列, 如树棉 (Y12072)、银胶菊 (X82543)、辣椒 (X84695)、拟南芥 (X75789)、山艾树 (AY308476)、水稻 (D85317)、薄荷 (AF384040)、

黄花蒿 (AF112881)、白羽扇豆 (U20771)、玉米 (L39789) 等具有较高的同源性, 分别为 65%、71%、81%、64%、55%、60%、65%、64%、66%、57%; 但相对于蛋白质-蛋白质的 Blastp 比对来说, 番茄与其它植物 FPPS 的氨基酸序列有着更广泛和更高的相似性, 如与树棉 (CAA72793)、银胶菊 (CAA57893)、辣椒 (CAA59170)、拟南芥 (CAA53433)、山艾树 (AAP74719)、水稻 (BAA19856)、薄荷 (AAK63847)、黄花蒿 (AAD17204)、白羽扇豆 (AAA87729)、玉米 (AAB39276) 的一致性分别为 79%、79%、88%、77%、75%、69%、80%、79%、78%、67%, 相似性分别为 88%、88%、92%、86%、85%、82%、89%、87%、86%、81%。

```

1                                     tatacataaaccttttttttctctgtcaaaa
33 atggctgatctgaagaagaatttttggatgtgactctgttcttaaatctgatctgttagaagacacagctttttaa
M A D L K K K F L D V Y S V L K S D L L E D T A F E
111 tttcacagatgattctcgtaaatgggtcgacaagatgctggactacaatgtacctggagggaagctgaaccgaggacta
F T D D S R K W V D K M L D Y N V P G G K L N R G L
189 tctgttattgacagccttagtttggtaagatggaaaagaatfaaccgctgatgaaatctttaaagcactctgctctt
S V I D S L S L L K D G K E L T A D E I F K A S A L
267 ggctgggtgcatgaaatggcttcaagcacttccctgttcttgatgatataatggatggatctcacaacgctcaggtt
G W C I E W L Q A Y F L V L D D I M D G S H T R R G
345 caaccatgctgggtacaacttagaaaaggttgggatgatgcttataatgatggcattcttcttcgcaaccacatcacc
Q P C W Y N L E K V G M I A I N D G I L L R N H I T
423 agaattctgaagaagtacttcaggccagagctcttattatggtgatctcttagatttgtttaacgaggtagagttccag
R I L K K Y F R P E S Y Y V D L L D L F N E V E F Q
501 actgctctggacaaatgatagatttgatcacacacttgttaggagagaagaatttacaagtactcattgtctatt
T A S G Q M I D L I T T L V G E K D L S K Y S L S I
579 catgcccggattgtccaglataaaactgcctattatcattttacctcccagtgccatgtgcacttcttatggtagga
H R R I V Q Y K T A Y Y S F Y L P V A C A L L M V G
657 gagaatcttgacaagcatgttgatgtcaagaaaatcttattgacatgggaatctactttcaagttcaggatgattat
E N L D K H V D V K K I L I D M G I Y F Q V Q D D Y
735 ctggactgcttctgaccagaggtgctgggcaagatggcactgatattcaggactttaaagtgctcttggttggta
L D C F A D P E V L G K I G T D I Q D F K C S W L V
813 gttaaagctctagaactatgcaacgaggagcaaaagaagatattttgagaactatggaaaagataacgctgctgtt
V K A L E L C N E E Q K K I L F E N Y G K D N A A C
891 attgctaaaataaggcactctataatgatctcaacttgaggaagtatttctggaatacagaagaagcctcattgaa
I A K I K A L Y N D L K L E E V F L E Y E K T S Y E
969 aaattgaccactctatgacgctcacaagcaaacgctacagcgagtgctgctatcattcttgggaagatttat
K L T T S I A A H P S K A V Q A V L L S F L G K I Y
1047 aagaggcagaag taggataaaaataatgaatgaacttcattctccattgggaatgagatgctgtatttcatgcat
K R Q · K *
1125 ttttctgttttttttctctatttttagtttgggactctgatgatggattgttatttttttcaagacatggaact
1202 cttatcctttttcactttaataaacagtagaaaagtccaatgaaatcttcagtaaaaaaaaaaaaaaaaa
    
```

图 1 番茄 FPPS 基因的全长 cDNA 序列及推导的氨基酸序列

Fig. 1 Full-length cDNA sequence and deduced amino acid sequence of FPPS in *Lycopersicon esculentum*
 起始密码子用方框显示, 终止密码子用斜体下划线显示, 5' 和 3' 非翻译区加粗显示, 保守结构功能域用下划线显示。
 The initiation codon (ATG) is boxed and the termination codon (TAG) in italic is underlined. The untranslated regions of
 5' and 3' are shown in bold. Conserved function motifs are underlined.

对橡胶、香蕉、苹果、向日葵、葡萄 FPPS 的 cDNA 及氨基酸序列进行类似同源性比对, 也得到上述规律, 即不同植物氨基酸序列之间的比对相对于核酸序列之间的比对来说, 有着更高和更广泛的相似性。

依据 Higgins 等^[26]的方法, 用 Clustal W 程序对树棉、银胶菊、辣椒、拟南芥、水稻、薄荷、黄花蒿、白羽扇豆、玉米、番茄、橡胶、香蕉、苹果、向日葵、葡萄的氨基酸序列进行多序列比对, 结果表明 (图 2): 番茄、橡胶、香蕉、苹果、向日葵、葡萄 FPPS 的功能域与其它植物有着几乎一致的氨基酸组成^[12-21], 均具有 FPPS 活性所必需的典型的多肽位点, 即 5 个保守的结构 motif^[27], 其中有两个是富含天冬氨酸 Asp 的 motif (DDXXD), 这两个 Asp-motif 被认为是烯丙基转移酶与底物异戊烯基焦磷酸相结合的位点^[28]。这一结果暗示功能结构域在分子进化中具有较高的稳定性。

2.3 FPPS 转运肽的预测和分析

转运肽 (Transit peptides 或 Transit sequences) 位于蛋白质的 N 端, 蛋白质首先在细胞质基质游离核糖体上起始合成前体蛋白, 然后在转运肽的引导下, 转运至叶绿体基质中发挥作用^[29]。因此, 转运肽的预测和分析, 对正确认识和理解蛋白质的细胞定位及结构域的分选有着重要的意义。用 TargetP 1.1 Server 对番茄、橡胶、香蕉、苹果、向日葵、葡萄 FPPS 氨基酸序列进行预测^[30], 结果表明, 所有序列含有叶绿体转运肽、线粒体目标肽 (Targeting peptide) 及分泌途径信号肽 (Signal peptide) 的分值均较低, 无氨基酸残基分裂位点, 由此可以看出, 番茄、橡胶、香蕉、苹果、向日葵、葡萄 FPPS 可能不存在转运肽酶切位点, 不具有转运肽。可以推断, 植物 FPPS 在细胞质中合成后, 不进入其它亚细胞器, 而是留在细胞质中与相关代谢底物相互作用。由于 FPPS 是植物体内倍半萜和三萜合成中的一个限速酶, 而倍

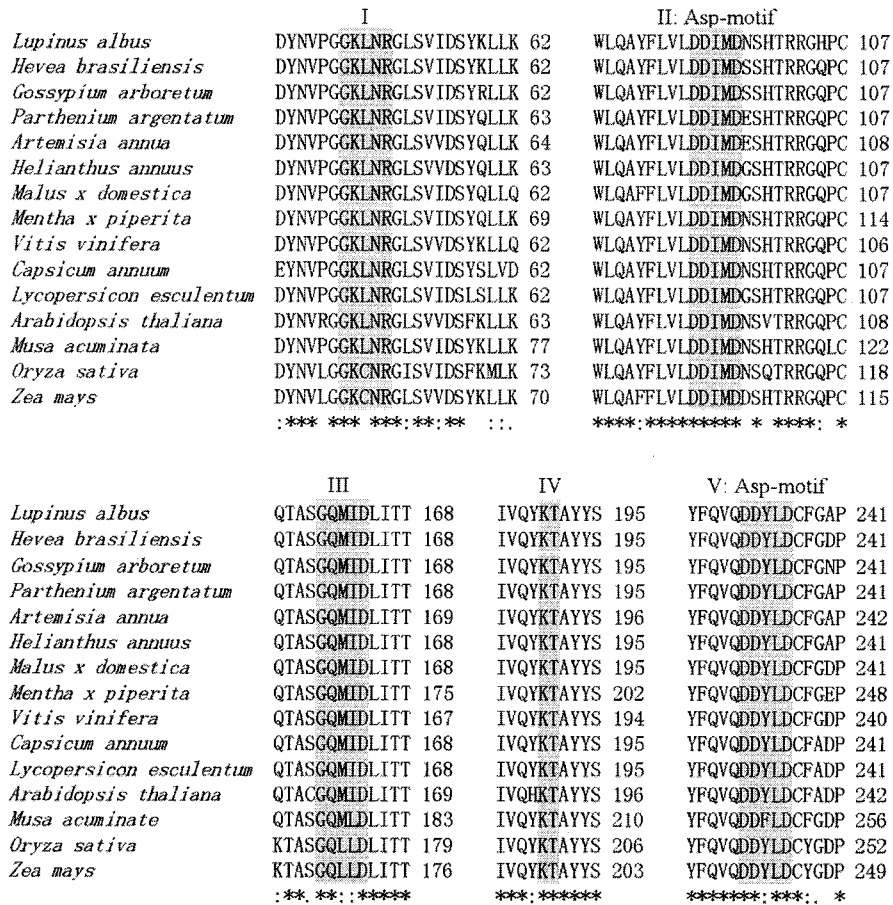


图 2 不同植物 FPPS 氨基酸序列多重比对

Fig. 2 Multiple sequence alignment of FPPS from 15 species

符号“*”、“:”、“.”分别表示氨基酸完全相同、保守和半保守三种情况。保守结构域 motif 用阴影表示。

* signifies identical, : signifies conservation, and . signifies 50% conservation. The conservative domains motifs are shown with gray shading.

半萜和三萜是由 MVA 途径产生，该途径在细胞质中进行^[10,31]，由此看出，这一推断与相关萜类代谢途径在细胞中的作用部位是相吻合的。

2.4 不同植物 FPPS 跨膜结构域的预测和分析

跨膜结构域是膜内在蛋白与膜脂相结合的主要部位，一般由 20 个左右的疏水氨基酸残基组成，形成 α 螺旋，它固着于细胞膜上起“锚定”作用^[29]。跨膜结构域的预测和分析，对正确认识和理解蛋白质的功能、结构、分类、方位及细胞中的作用部位均有着重要的指示意义。用 TMHMM 2.0 Server 对番茄、橡胶、香蕉、苹果、向日葵、葡萄 FPPS 氨基酸序列的跨膜结构域进行预测，结果表明，预测结果均与番茄一致（图 3），FPPS 整条肽链都位于膜外，可能不存在跨膜结构域。结合上述转运肽的预测，可

以推断，植物 FPPS 在细胞质中合成前体蛋白后，不进入其它亚细胞器，而是留在细胞质中，经过翻译后蛋白质的加工成为成熟蛋白质发挥催化作用，而不与膜脂结合。

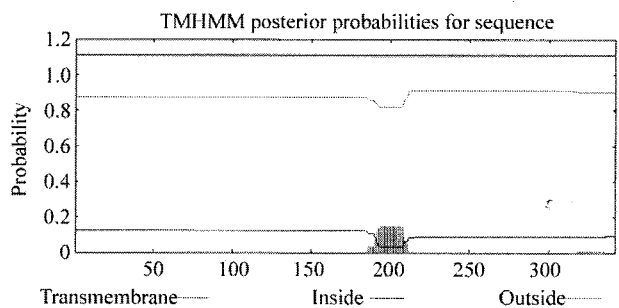


图 3 番茄 FPPS 跨膜结构域的预测

Fig. 3 Predicted transmembrane domains of FPPS in *Lycopersicon esculentum*

2.5 不同植物 FPPS 疏水性/亲水性的预测和分析

20 种氨基酸都具有疏水残基,蛋白质的折叠总是倾向于把疏水残基埋藏在分子的内部,这一趋势加上空间立体条件和其它一些因素,决定了一个蛋白质最终折叠成的三维空间构象。此外,通过了解肽链中不同肽段的疏水性,可以对跨膜蛋白的跨膜结构域进行预测。因此,疏水性/亲水性的预测和分析,对蛋白质次级结构的预测及功能结构域的分选提供了重要的理论参考。用 ProtScale 对番茄、橡胶、香蕉、苹果、向日葵、葡萄 FPPS 氨基酸序列的疏水性/亲水性进行预测^[32],结果表明,预测结果均与番茄一致(图 4),整个多肽链表现为亲水性,没有明显的疏水区域。综合前述跨膜结构域的预测结果,可以推断,植物 FPPS 不存在明显的疏水区域,与植物 FPPS 不存在跨膜结构域相一致。

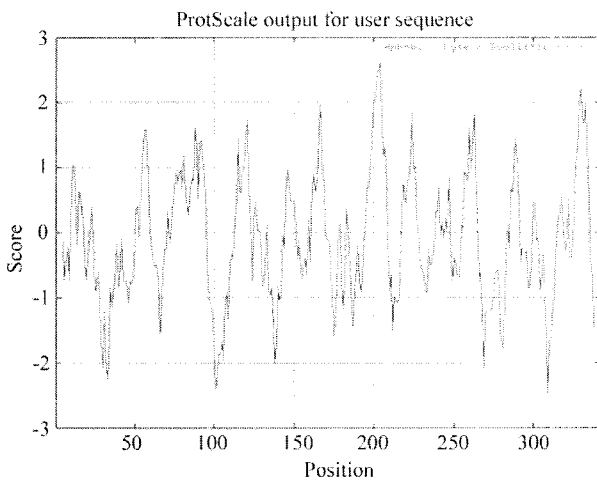


图 4 番茄 FPPS 疏水性/亲水性的预测

Fig. 4 Predicted hydrophobicity/hydrophilicity of FPPS in *Lycopersicon esculentum*

2.6 不同植物 FPPS 二级及三级结构的预测和分析

蛋白质高级结构的预测和分析,对理解蛋白质结构与功能之间的相关性有着极其重要的意义。依据 Geourjon 和 Deléage^[33]及 Combet 等^[34]的方法,用 SOPMA 对番茄、橡胶、香蕉、苹果、向日葵、葡萄 FPPS 氨基酸序列的二维结构进行预测,结果表明,预测结果均与番茄一致(图 5), α -螺旋和不规则盘绕是 FPPS 整体结构中的主要组成结构元件, β -转角和延伸链散布于整个蛋白质中。由此可以推断,植物 FPPS 整体蛋白结构中, α -螺旋和不规则盘

绕是最大量的结构元件, β -转角和延伸链散布于整个蛋白质中。

依据 Peitsch^[35]、Guex 和 Peitsch^[36]及 Schwede 等^[37]的方法,用 SWISS-MODEL 对番茄、橡胶、香蕉、苹果、向日葵、葡萄 FPPS 氨基酸序列进行蛋白质三维结构同源性建模(molecular homologous modeling),结果表明,预测结果均与番茄一致(图 6),FPPS 的功能域在空间布局上是由 α -螺旋围绕形成的中间具一个“大空穴”的立体结构,“大空穴”的表面藏有五个功能保守 motif,其中与底物相结合的两个 Asp-motif(DDXXD)位于“空穴”的内壁,FPPS 的酶促催化反应就在空穴中进行,这一结果与哺乳动物鸟类 FPPS 的三维结构非常相似^[38]。由此可以推断,植物 FPPS 是一个在萜类生物合成途径中起作用的功能蛋白。

2.7 FPPS 氨基酸序列的分子系统进化分析

在 MEGA 3.1 软件平台上采用 Neighbor-Joining 方法对树棉、银胶菊、辣椒、水稻、拟南芥、山艾树、薄荷、黄花蒿、白羽扇豆、玉米、番茄、橡胶、香蕉、苹果、向日葵、葡萄、曼地亚红豆杉(*Taxus x media*)、人参(*Panax ginseng*)、三七(*Panax notoginseng*)、积雪草(*Centella asiatica*)、银杏(*Ginkgo biloba*)、忽布(*Humulus lupulus*)的氨基酸序列进行分子系统进化分析^[39],结果如图 7 所示,22 种不同植物的 FPPS 氨基酸序列聚成若干不同的分支,这一结果虽与其他学者的研究不完全一致^[40-45],但还是反映了这些植物之间的亲缘关系。如曼地亚红豆杉与银杏的亲缘关系较近,聚成裸子植物(Gymnosperm)支,得到 100%的支持率;水稻、玉米、香蕉聚成单子叶植物(Monocotyledones)支,其中,水稻与玉米同属禾本科(Grammineae)又聚成一支;余下的 17 种植物聚成双子叶植物(Dicotyledones)支,得到 99%的支持率。在双子叶植物支内,人参与三七同属五加科(Araliaceae),先聚在一起,再与亲缘关系较近的伞形科(Umbelliferae)积雪草聚为一支;辣椒与番茄同属茄科(Solanaceae),银胶菊、山艾树、黄花蒿、向日葵同属菊科(Compositae),它们也都分别聚为一支。由此看出,依据氨基酸序列所得出的系统演化关系虽不能真实地反映植物的自然进化关系,但其结果对判断不同植物之间的亲缘关系仍具有一定的借鉴意义。

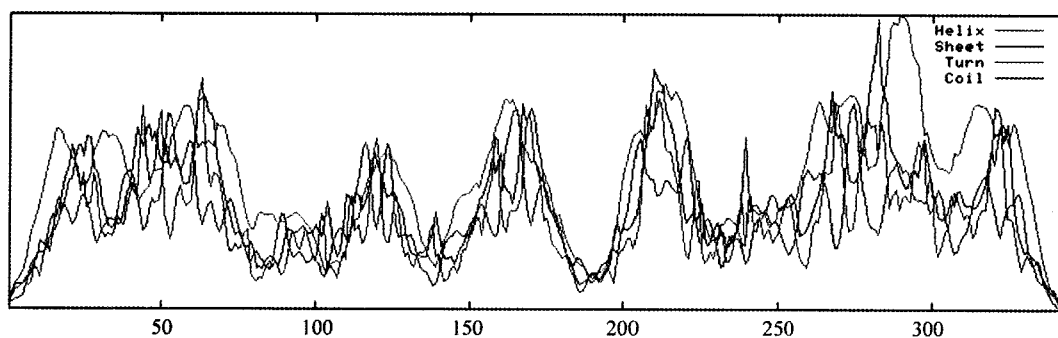


图 5 番茄 FPPS 二级结构的预测

Fig. 5 Predicted secondary structure of FPPS in *Lycopersicon esculentum*

Helix: α -螺旋(蓝线); Sheet: 延伸链(红线); Turn: β -转角(绿线); Coil: 不规则盘绕(紫线)。

Helix: Alpha helix (blue); Sheet: Extended strand (red); Turn: Beta turn (green); Coil: Random coil (purple)

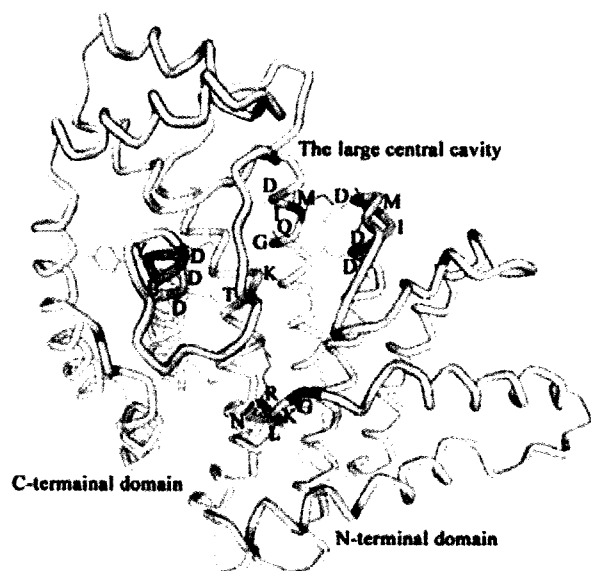


图 6 基于同源性建模获得的番茄 FPPS 功能结构域的立体结构

Fig. 6 Three-dimensional structure of the functional domain of the FPPS

from *Lycopersicon esculentum* using homology-based models

蓝色英文大写字母示功能结合 motif. The substrate-binding

motifs are in blue and marked with capital letters

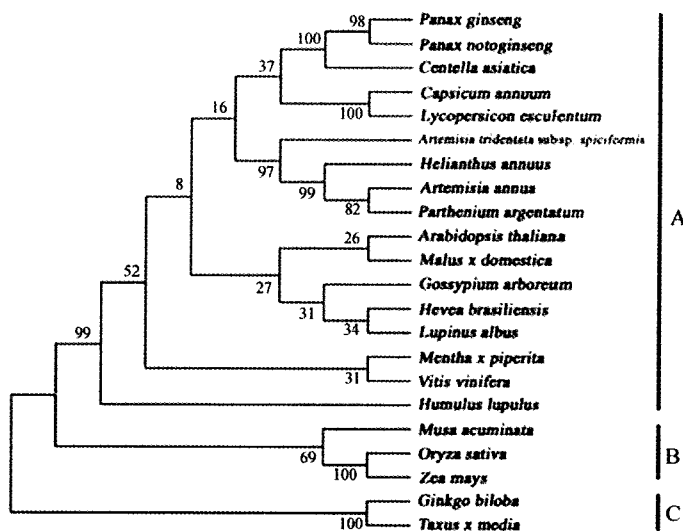


图 7 不同植物 FPPS 的氨基酸序列分子系统发生树

Fig. 7 Phylogenetic tree of the amino acid sequences of the FPPSs from 22 species

分支上的数字是靴带值(重复 1000 次)。The bootstrap values are shown in arabic (1000 replicates). A: 双子叶植物 Dicotyledones; B: 单子叶植物 Monocotyledones; C: 裸子植物 Gymnosperm.

3 讨论

目前对法呢基焦磷酸合成酶的研究主要用特异性寡聚核苷酸作探针, 从不同植物组织或细胞(根、茎皮、叶、乳汁、胚乳、油腺细胞)的 cDNA 文库中分离筛选该酶的编码基因, 通过大肠杆菌 (*Escherichia coli*) 原核表达系统和酵母 (*Saccharomyces cerevisiae*) 缺陷功能互补法对表达蛋白的功能进行验证和检测^[12-21]。已有的实验研究表明: FPPS 属于烯丙基转移酶家族, 其催化终产物以法

呢基焦磷酸为主, 是一个组成型表达蛋白, 其氨基酸序列与植物来源的氨基酸序列具有较高的同源性, 而非植物来源的氨基酸序列同源性较低, Southern blot 显示它在拟南芥和水稻基因组中是一个单拷贝基因, 在玉米基因组中有至少两个拷贝, Northern blot 显示它在水稻叶中具有高的表达水平, 其表达受蓝光刺激而增强, 它在橡胶树和银胶菊的橡胶生物合成中起着重要的作用^[12-18, 46]。对比前述生物信息学的预测和分析可以看出, 就 FPPS 的氨基酸序列与其它生物来源的氨基酸序列之间的

同源性及其功能确认来说, 实验研究结果与生物信息学的预测分析结论相一致, 但对 FPPS 的实验研究主要集中在基因的调控与表达、基因的拷贝数、基因的表达水平、丰度及基因在植物组织中的表达部位等方面, 而对该酶基因核酸及氨基酸序列的结构、理化性质、转运肽、跨膜拓扑结构域、疏水性 / 亲水性及蛋白质的二级及三级结构、分子系统进化关系的实验研究还未见报道。

生物大分子的功能是由其基本组成成分及空间结构所决定, 也即“结构迎合功能”, 因此, 生物信息学的预测和分析, 为后期 FPPS 氨基酸序列的结构、性质、功能域及蛋白质空间结构的实验研究提供了一定的理论参考依据, 同时也为该酶在植物体内的作用机制、过程及基因的调控与表达提供了一定的线索。通过生物信息学的预测和分析, 结合实验的验证, 深入了解结构与功能之间的相关性, 对正确认识“基因 - 编码蛋白 - 功能”之间的表达规律与调控机理有着深远的意义。

参考文献

- [1] Sacchettini J C, Poulter C D. Creating isoprenoid diversity [J]. *Science*, 1997, 277:1788-1799
- [2] Verpoorte R, Alfermann A W. *Metabolic Engineering of Plant Secondary Metabolism* [M]. Dordrecht: Kluwer Academic Publishers, 2000:1-30.
- [3] Lange B M, Rujan T, Martin W, et al. Isoprenoid biosynthesis: the evolution of two ancient and distant pathways across genomes [J]. *Proc Natl Acad Sci USA*, 2000, 97:13172-13177.
- [4] Page J E, Hause G, Raschke M, et al. 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 [J]. *Plant Physiol*, 2004, 134:1401-1413.
- [5] Lange B M, Wilding M R, Mccaskill D, et al. A family of transketolases that directs isoprenoid biosynthesis via a mevalonate independent pathway [J]. *Proc Natl Acad Sci USA*, 1998, 95: 2100-2104.
- [6] Hartmul K L. The 1-deoxyxylulose-5-phosphate pathway of isoprenoid biosynthesis in plant [J]. *Annu Rev Plant Physiol Plant Mol Biol*, 1999, 50:47-65.
- [7] Rohmer M. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and high plants [J]. *Nat Prod Rep*, 1999, 16:565-574.
- [8] Chen D H (陈大华), Ye H C (叶和春), Li G F (李国凤), et al. Advances in molecular biology of plant isoprenoid metabolic pathway [J]. *Acta Bot Sin (植物学报)*, 2000, 42:551-558. (in Chinese)
- [9] Dudarevan N, Andersson S, Orlova I, et al. The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers [J]. *Proc Natl Acad Sci USA*, 2005, 102: 933-938.
- [10] Liu Y, Wang H, Ye H C, et al. Advances in the plant isoprenoid biosynthesis pathway and its metabolic engineering [J]. *J Integ Plant Biol*, 2005, 47:769-782.
- [11] Cunillera N, Arró M, Delourme D, et al. *Arabidopsis thaliana* contains two differentially expressed farnesyl diphosphate synthase genes [J]. *J Biol Chem*, 1996, 271:7774-7780.
- [12] Delourme D, Lacroute F, Karst F. Cloning of an *Arabidopsis thaliana* cDNA coding for farnesyl diphosphate synthase by functional complementation in yeast [J]. *Plant Mol Biol*, 1994, 26: 1867-1873.
- [13] Attucci S, Aitken S M, Gulick P J, et al. Farnesyl pyrophosphate synthase from white lupin: molecular cloning, expression, and purification of the expressed protein [J]. *Arch Biochem Biophys*, 1995, 321:493-500.
- [14] Huguency P, Bouvier F, Badillo A, et al. Developmental and stress regulation of gene expression for plastid and cytosolic isoprenoid pathways in pepper fruits [J]. *Plant Physiol*, 1996, 111: 619-626.
- [15] Li C P, Larkins B A. Identification of a maize endosperm-specific cDNA encoding farnesyl pyrophosphate synthetase [J]. *Gene*, 1996, 171:193-196.
- [16] Matsushita Y, Kang W, Charlwood B V. Cloning and analysis of a cDNA encoding farnesyl diphosphate synthase from *Artemisia annua* [J]. *Gene*, 1996, 172:207-209.
- [17] Pan Z, Herickhoff L, Backhaus R A. Cloning, characterization, and heterologous expression of cDNAs for farnesyl diphosphate synthase from the guayule rubber plant reveals that this prenyltransferase occurs in rubber particles [J]. *Arch Biochem Biophys*, 1996, 332:196-204.
- [18] Sanmiya K, Iwasaki T, Matsuoka M, et al. Cloning of a cDNA that encodes farnesyl diphosphate synthase and the blue-light-induced expression of the corresponding gene in the leaves of rice plants [J]. *Biochem Biophys Acta*, 1997, 1350:240-246.
- [19] Liu C J, Heinstejn P, Chen X Y. Expression pattern of genes encoding farnesyl diphosphate synthase and sesquiterpene cyclase in cotton suspension-cultured cells treated with fungal elicitors [J]. *Mol Plant Microb Interact*, 1999, 12:1095-1104.
- [20] Lange B M, Wildung M R, Stauber E J, et al. Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes [J]. *Proc Natl Acad Sci USA*, 2000, 97:2934-2939.
- [21] Hemmerlin A, Rivera S B, Erickson H K, et al. Enzymes encoded by the farnesyl diphosphate synthase gene family in the big sagebrush *Artemisia tridentata* subsp. *spiciformis* [J]. *J Biol Chem*, 2003, 278:32132-32140.
- [22] Baxevanis A D, Ouellette B F F. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* [M]. Translated by Li Y D (李衍达), Sun Z R (孙之荣). Beijing: Tsinghua University Press,

- 2000:231–250. (in Chinese)
- [23] Mount D W. Bioinformatics: Sequence and Genome Analysis [M]. Translated by Zhong Y (钟扬), Wang L (王莉), Zhang L (张亮), et al. Beijing: Higher Education Press, 2003:301–345. (in Chinese)
- [24] Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the ExPASy server [C]//Walker J M. The Proteomics Protocols Handbook. Herts, UK: Humana Press, 2005: 571–607.
- [25] Altschul S F, Madden T L, Schaffer A A, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs [J]. *Nucleic Acids Res*, 1997, 25:3389–3402.
- [26] Higgins D, Thompson J, Gibson T, et al. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice [J]. *Nucleic Acids Res*, 1994, 22: 4673–4680.
- [27] Sandmann G. Carotenoid biosynthesis in microorganisms and plants [J]. *Eur J Biochem*, 1994, 223:7–24.
- [28] Ohnuma S I, Hirooka K, Hemmi H, et al. Conversion of product specificity of archaeobacterial geranylgeranyl-diphosphate synthase [J]. *J Biol Chem*, 1996, 271:18831–18837.
- [29] Zhai Z H (翟中和), Wang X Z (王喜忠), Ding M X (丁明孝). Cell Biology [M]. Beijing: Higher Education Press, 2000:79–240. (in Chinese)
- [30] Emanuelsson O, Nielsen H, Brunak S, et al. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence [J]. *J Mol Biol*, 2000, 300:1005–1016.
- [31] Yang T (杨涛), Zeng Y (曾英). Plant terpenoid synthases: molecular biology and genetic engineering [J]. *Acta Bot Yunnan (云南植物研究)*, 2005, 27:1–10. (in Chinese)
- [32] Kyte J, Doolittle R F. A simple method for displaying the hydrophobic character of a protein [J]. *J Mol Biol*, 1982, 157:105–132.
- [33] Geourjon C, Deléage G. SOPMA: Significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments [J]. *Comput Appl Biosci*, 1995, 11: 681–684.
- [34] Combet C, Blanchet C, Geourjon C, et al. NPS@: Network Protein Sequence Analysis [J]. *Trends Biochem Sci*, 2000, 25: 147–150.
- [35] Peitsch M C. Protein modeling by E-mail [J]. *Bio/Technology*, 1995, 13:658–660.
- [36] Guex N, Peitsch M C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling [J]. *Electrophoresis*, 1997, 18:2714–2723.
- [37] Schwede T, Kopp J, Guex N, et al. SWISS-MODEL: an automated protein homology-modeling server [J]. *Nucleic Acids Res*, 2003, 31:3381–3385.
- [38] Tarshis L C, Proteau P J, Kellogg B A, et al. Regulation of product chain length by isoprenyl diphosphate synthases [J]. *Proc Natl Acad Sci USA*, 1996, 93:15018–15023.
- [39] Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment [J]. *Briefings in Bioinformatics*, 2004, 5:150–163.
- [40] Takhtajan A. Diversity and Classification of Flowering Plants [M]. New York: Columbia University Press, 1997:43–265.
- [41] Thorne R F. The classification and geography of the flowering plants: dicotyledons of the class Angiospermae [J]. *Bot Rev*, 2000, 66:442–664.
- [42] Soltis D E, Soltis P S, Chase M W, et al. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences [J]. *Bot J Linn Soc*, 2000, 133:381–461.
- [43] ANGIOSPERM PHYLOGENY GROUP (APG II). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II [J]. *Bot J Linn Soc*, 2003, 141:399–436.
- [44] Wu Z Y (吴征镒), Lu A M (路安民), Tang Y C (汤彦承), et al. The Families and Genera of Angiosperms in China, A Comprehensive Analysis [M]. Beijing: Science Press, 2003:57–1071. (in Chinese)
- [45] Wu Z Y (吴征镒), Chen X Q (陈心启). *Flora Reipublicae Popularis Sinicae*, Tomus I [M]. Beijing: Science Press, 2004: 95–573. (in Chinese)
- [46] Adiwilaga K, Kush A. Cloning and characterization of cDNA encoding farnesyl diphosphate synthase from rubber tree (*Hevea brasiliensis*) [J]. *Plant Mol Biol*, 1996, 30:935–946.