

向日葵异戊烯焦磷酸异构酶基因(*ipi*)的电子克隆和生物信息学分析

化文平, 王喆之*

陕西师范大学生命科学学院, 教育部药用植物资源与天然药物化学重点实验室, 西安 710062

摘要: 用电子克隆方法获得向日葵异戊烯焦磷酸异构酶(IPI)基因, 并用生物信息学方法对此基因编码产物从氨基酸组成、理化性质、二级和三级结构以及功能进行预测和分析的结果表明, 向日葵IPI为亲水性 α/β 蛋白, 含有IPP异构酶结构域, 其与其他植物的IPI在序列组成、结构及活性位点均有高度的相似性。

关键词: 异戊烯焦磷酸异构酶; 电子克隆; 向日葵; 生物信息学

Electronic Cloning and Characterizations of *ipi* Gene from *Helianthus annuus* L. Using Bioinformatics Tools

HUA Wen-Ping, WANG Zhe-Zhi*

Key Laboratory of Medicinal Resources and Natural Pharmaceutical Chemistry, Ministry of Education, College of Life Sciences, Shaanxi Normal University, Xi'an 710062, China

Abstract: An isopentenyl diphosphate isomerase (IPP) gene from *Helianthus annuus* was cloned by electronic cloning based on the EST sequences from UniGene of NCBI. Some characters of the *ipi* deduced amino acid were analyzed and predicted by the tools of bioinformatics in the following aspects: the composition of amino acid sequences, hydrophobicity or hydrophilicity, secondary and tertiary structure of protein and function, and so on. Analysis results showed that *ipi* gene contained an IPP isomerase domain, belonging to α/β protein, despite sequence variations (mainly at the N-terminus), the core structure, sequence and activity sites were highly similar with other plant IPI.

Key words: isopentenyl diphosphate isomerase (IPI); electronic clone; *Helianthus annuus*; bioinformatics

大多数萜类化合物(isoprenoid 或 terpenoid)是由C5结构的异戊二烯结构聚合而成, 其前体是C5结构的异戊烯焦磷酸(isopentenyl diphosphate, IPP)和其同分异构体二甲基丙烯焦磷酸(dimethylallyl diphosphate, DMAPP)。在植物体内DMAPP是IPP的活化状态, 它失去一个无活性的磷酸基团后与IPP缩合成聚异戊烯链或其他次生代谢产物(Campbell等1997)。异戊烯焦磷酸异构酶(isopentenyl diphosphate isomerase, IPI)催化IPP和DMAPP之间的可逆转化(Ramos-Valdivia等1997), 分为两类: I型IPI发现较早, 存在于大多数物种中(Zhang等2007), 其蛋白单体是一紧密的球形蛋白, 在二价阳离子作用下发挥功能; II型IPI是近几年发现的, 主要存在于一些革兰氏阳性菌、蓝藻和古生菌中, 为TIM桶状黄素蛋白, 不仅需要阳离子还需要FMN和NAD(P)H辅助(Siddiqui等2005)。Kajiwar等(1997)认为, 植物体内的IPI可能是一关键性的限速酶, 他们将酵

母 *Pharffi arhosozyma* 和雨生红球藻中分离的IPI cDNA转入大肠杆菌菌株JMI01后番茄红素产量增加3.6~4.5倍。植物*ipi*基因在大肠杆菌中超量表达后, 类胡萝卜素合成量增加(Gallagher等2003)。许多植物的*ipi*基因已经克隆, 如烟草(*Nicotiana tabacum*)、拟南芥(*Arabidopsis thaliana*)、伯惠绣衣(*Clarkia breweri*)、*Taxus brasiliensis*、龙胆草(*Gentiana lutea*)、喜树(*Camptotheca acuminata*)、茶树(*Melaleuca alternifolia*)和玉米(*Zea mays*)等。本文以NCBI数据库为基础, 从向日葵中电子克隆得到编码IPI基因的cDNA序列, 采用生物信息学方法对其推导的氨基酸序列从氨基酸组成、理化性质、疏水性、亚细胞定位及结构功能等进行了预测和分析。

收稿 2007-11-14 修定 2007-12-27

资助 “十一五”科技支撑计划(2006BA106A12)。

* 通讯作者(E-mail: zzwang@snnu.edu.cn; Tel: 029-85310260)。

材料与amp;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 IPI 的核酸序列及其对应的氨基酸序列。向日葵 EST 序列来自于 NCBI 的 UniGene 数据库, 其序列号为 BU030901.1、AJ437873、AJ541418。用 CAP3 程序(<http://bioweb.pasteur.fr/seqanal/interfaces/cap3.html>)对 EST 序列进行拼接。依据该序列采用 Primer Premier 5.0 设计了 RT-PCR 扩增引物 IPI-S (5' ACCATCACTTCCACTTGCAACA 3') 和 IPI-A (5' AAAAAAGCAATACCAAAATCACACT 3')。向日葵 (*Helianthus annuus* L. cv. RHA280) 材料采自西安植物园。总 RNA 提取试剂盒 (BIOZOL) 购自于 BioFlux 公司; RevertAid First Strand cDNA Synthesis 试剂盒购自 Fermentas 公司, RNA 提取和第一链 cDNA 合成严格按照试剂盒说明书完成。PCR 反应程序: 95 °C 5 min; 95 °C 30 s, 55 °C 30 s, 72 °C 90 s, 40 个循环; 72 °C 8 min。扩增片段经凝胶回收 (试剂盒购自安徽优品生物工程公司) 插入 T 载体 (购自 Promega 公司) 后转入大肠杆菌 DH5 α 后送上海捷瑞生物公司测序。

依据 DNASTar 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。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 DNASTar 软件及 ProtParam、pI/MW、ORF Finder 在线工具进行; 核酸及氨基酸序列的同源性比对和多序列比对用 Blast 和 Clustal W 在线工具完成; 蛋白质亲水性 / 疏水性的分析用在线工具 ProtScale 完成; 蛋白质二级及三级结构的预测用 SOPMA 和 SWISS-MODEL 在线工具完成, 三维视图用 SweetMollyGrace 1.2 软件完成。

结果与amp;讨论

1 向日葵 *ipi* 基因 cDNA 序列的电子克隆及 RT-PCR 验证

依据 Letondal (2001) 的方法, 电子克隆出一条长 1 219 bp 的向日葵 *ipi* 基因序列, 该基因序列包含一个 888 bp 的 ORF 框, 编码 295 个氨基酸 (图

1)。以起始密码子 ATG 的 A 碱基为第一位 (标记为 +1), 在 +4 位核苷酸为 G, -1 位核苷酸为 A, 符合典型的 Kozak 序列规则, 此外, 在 3' -UTR 区有典型的 polyA 尾, 说明克隆序列符合一个真核基因全长 cDNA 的基本特征 (李桂源和钱骏 2004)。

提取向日葵总 mRNA 经 RT-PCR 反应扩增出一特异性条带 (图 2)。经测序验证该片段 917 bp 包含该基因完整 ORF 框, 与拼接序列完全吻合。

2 向日葵 *ipi* 基因序列相似性比对分析

用 Blastn 程序对向日葵 IPI 的核酸序列进行同源性比对 (Altschul 等 1997) 的结果表明向日葵 IPI 的核酸序列与其他植物, 如茶树 (AF031080)、杜仲 (*Eucommia ulmoides*, AB041629)、莴苣 (*Lactuca sativa*, AF188063)、万寿菊 (*Tagetes erecta*, AF188064)、杠柳 (*Periploca sepium*, AB091677)、甘薯 (*Ipomoea batatas*, DQ150100.1)、烟草 (AB049815.1) 的一致性 (identities) 分别为 81%、81%、80%、80%、80%、80%、80%。其推导氨基酸序列用 Blastp 程序进行同源比对 (Altschul 等 1997) 的结果表明, 向日葵 *ipi* 所推导氨基酸序列与其他植物的 IPI 具有很高的同源性, 如龙胆草 (BAE92732)、茶树 (AAB94132)、烟草 (BAB40974、BAB40973)、伯惠绣衣 (CAA57947.1) 等的同源性分别为 91%、95%、96%、90%、90%。由此说明克隆到的基因为 *ipi* 基因, 并揭示出不同物种中 *ipi* 基因及其编码氨基酸序列间具有高度的相似性, 在初级结构上是保守的。

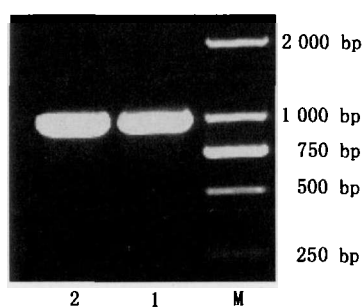
3 向日葵 *ipi* 基因及其推导的氨基酸序列组成成分和理化性质分析

采用 Gasteiger 等 (2005) 的方法, 用 ORF Finder、ProtParam 和 pI/MW 对向日葵、拟南芥、伯惠绣衣、烟草和大肠杆菌氨基酸序列进行分析的结果 (表 1) 表明, 各植物 *ipi* 基因编码氨基酸序列长度变化较大, 但等电点、酸碱性氨基酸比例、极性氨基酸比例、疏水性氨基酸比例等基本一致; 除向日葵 IPI 和拟南芥 IPI 属于不稳定蛋白外, 其他蛋白都属于稳定蛋白分子。说明向日葵 IPI 与其他植物及大肠杆菌的 IPI 的基本性质是相似的。

4 向日葵 IPI 氨基酸序列的亲水性 / 疏水性分析

疏水性是氨基酸固有的特性, 疏水性氨基酸在蛋白质内部, 由于其疏水的相互作用, 而在形

TTAGCGACGCTAAAAACCTTCAAATATCTTGCAGCCATAGCCCCATGGGCATCAACACTCCCAACTAACC	72
GCGTTCAGTACTCCATGAGTATTTCTCTCGGTTGGCGGACCCACCATCACTTCCACTTGCACCAATGGCT	144
	M A
CAATCTTTGATCTCAAAATGTTCAAATTCGCCATTCTCAAAGCCCTAAGCTTTTCTTATCCCTTTTCATTGCT	216
Q S L I S K C S I R H S Q S P K L S Y P L S L S	
TCATCTCAACCCCCACCTCAGTTCCTCTCTTTCAAACCCCGCTTACCTGTTTTCTCTTTCACCGCCGGC	288
S S S T P P P Q F L S F K P R L P V F S F T A G	
GTCCTCTTCTCTCTTCCGGTCAGTCTCCTCCGCTGGTGTGATGCAGATCCACCATGGATGCTGTC	360
V S S S S S P V T S P P L G A D A D S T M D A V	
CAACGCGGCTTATGTTTGAAGACGAGTGCATTCTGGTTGATGCTAATGATGGTGTGTGGTTCATGATACA	432
Q R R L M F E D E C I L V D A N D G V V G H D T	
AAATACAATTTGTCACTTGATGGAGAAGATTGAATCCGAGAACTTGTTCACCGTCTTTCAGTGTATTCTTG	504
K Y N C H L M E K I E S E N L L H R A F S V F L	
TTCAACTCAAAATATGAGTTACTCTTCAACAAGATCTCAACAAGGTGACATTCCCTTTGGTATGGACA	576
F N S K Y E L L L Q Q R S S T K V T F P L V W T	
AACACATGTTGCAGCCATCCACTATACCGAGAATCTGAGCTTATTGAAGAGAATTCCTTGGTGCAGAAAT	648
N T C C S H P L Y R E S E L I E E N C L G V R N	
GCTGCACAGAGAAAGCTTCTAGACGAACTCGGTATTCTTTCAGAAAGAACTTCCAGTTGATGAGTTTATCCC	720
A A Q R K L L D E L G I P S E E L P V D E F I P	
CTAAGTCGCATTCTCTACAAAGCACCATCTGATGAAAGTGGGGCAACATGAACTTGATTACCTGCTCTTC	792
L S R I L Y K A P S D G K W G E H E L D Y L L F	
ATTGTACGGGATGTGAGCATGAACCCCAACCCAGATGAGGTTGCAGACGTTAAGTATGTGAACCGGAGCAA	864
I V R D V S M N P N P D E V A D V K Y V N R E Q	
TTAAAAGAGCTGGTGAAGAAAGCAGACGCGGGGAGGAGGTTCTGAAGCTGTCTCCATGGTTTCAGGATAGTT	936
L K E L V R K A D A G E E G L K L S P W F R I V	
GTAGATAAATTCTTGTTCAGTGGTGGATCATGTGAAAAGGGAAAGCCTACATGAAGCTGCTGATATGAAG	1008
V D N F L F K W W D H V E K G S L H E A A D M K	
ACCATCCATAACCTGTAATTTTCATCATCTTTGAAAAAGATGTTTGAITTTATAAAAATTATTAGTGTGATTTT	1080
T I H N L	
GGTATTGCTTTTTTCCCACTTATATGTTTGAATAATTGGATGACTGAGGTTATGCAITTAAGGTTATGTAAT	1152
ATTTCAAAAAATATATAAATATAGCCCTTTGGATTTTAAAAAAGGATGTTTGAITTTATAAAAATTATTAGTGTGATTTT	1219

图1 向日葵 *ipi* 基因 cDNA 序列及其推导的氨基酸序列Fig.1 The full-length cDNA sequence and deduced amino acid of *ipi* from *H. annuus*图2 向日葵 *ipi* 基因 RT-PCR 扩增Fig.2 The RT-PCR amplification of *ipi* from *H. annuus*M: 分子量标准; 1 和 2: 扩增出的向日葵 *ipi* cDNA 条带。

成和保持蛋白质的三级结构上起作用。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分析均可提供参考。采用 Gasteiger 等(2003)的方法和 ProtScale 程序对向日葵 IPI 亲水性/疏水性分析的结果(图3)表明,第202位色氨酸(Trp)亲水性

最强,第19位丝氨酸(Ser)疏水性最强,N端存在一疏水区域。从总体上看,大部分氨基酸(分值为负值)属于亲水性氨基酸,氨基酸序列属亲水蛋白。进一步比较拟南芥、伯惠绣衣和烟草中 IPI 的结果均显示亲水特性。

5 向日葵 IPI 的亚细胞定位

蛋白质在细胞中的定位与蛋白执行的功能密切相关。植物中 *ipi* 基因都是以基因家族形式存在(Cunningham 和 Gantt 2000), *ipi* 基因或酶不仅存在于细胞质,同时在线粒体、过氧化物体(Ramos-Valdivia 等 1997)和质体(Cunningham 和 Gantt 2000)都有发现。在模式植物拟南芥多种亚细胞结构(线粒体、质体、细胞质和内质网)中均发现有 IPI (Campbell 等 1997; Laule 等 2003)。采用 PSORT 程序对向日葵 *ipi* 基因编码产物进行定位的结果(表2)显示,向日葵 *ipi* 基因编码产物定位

表1 不同植物和大肠杆菌的 *ipi* 基因推导的氨基酸序列的组成成分和理化性质分析Table 1 An analysis of composition and physical and chemical characters of deduced amino acid sequence of *ipi* among the different plants and *E. coli*

注册号	推导的氨基酸残基数	分子量 / kDa	理论等电点	碱性与酸性氨基酸比例 / %	极性氨基酸比例 / %	疏水性氨基酸比例 / %	蛋白质不稳定性指数	
向日葵	—	295	33.46	5.42	10.51/13.90	25.76	34.58	43.96 (属于不稳定类蛋白)
拟南芥	(1) AAB67741.1	291	33.21	5.69	12.37/15.12	22.34	36.43	43.88 (属于不稳定类蛋白)
	(2) AAC49932.1	284	32.61	6.10	13.03/14.44	21.48	37.32	33.70 (属于稳定蛋白)
烟草	(1) BAB40973.1	290	33.20	5.94	12.41/14.14	23.45	36.55	34.63 (属于稳定蛋白)
	(2) BAB40974.1	235	27.18	4.87	11.06/17.45	20.85	37.02	26.80 (属于稳定蛋白)
伯惠绣衣	CAA57947	287	32.95	5.26	11.50/15.33	22.65	36.93	37.40 (属于稳定类蛋白)
橡胶	AF111843.1	234	26.75	4.90	11.11/16.67	20.94	37.18	23.51 (属于稳定蛋白)
玉米	AF330034.1	237	27.23	4.86	11.39/18.14	18.98	36.71	23.63 (属于稳定蛋白)
大肠杆菌	AAD26812.1	182	20.51	5.94	9.89/11.54	26.37	36.81	26.17 (属于稳定蛋白)

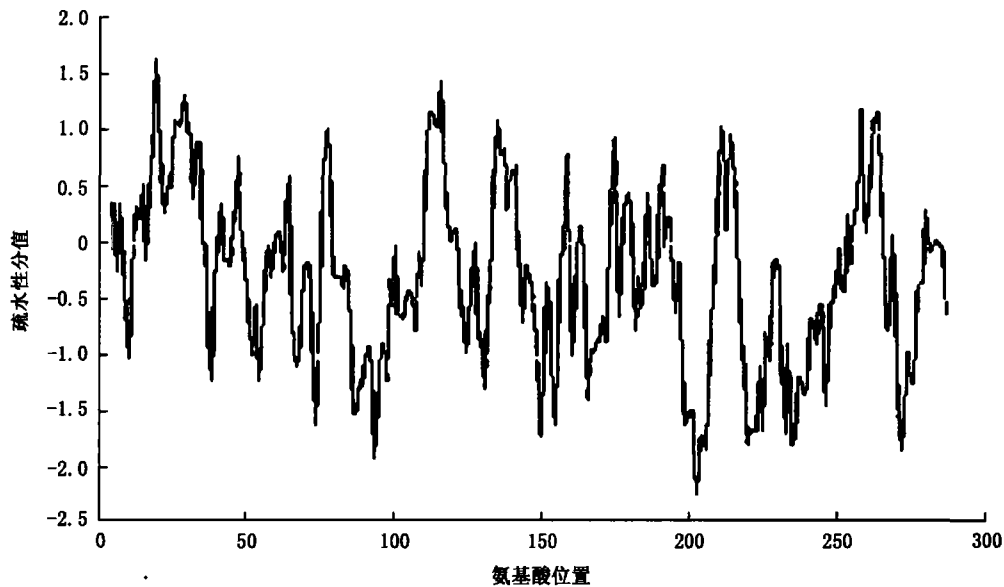
图3 向日葵 *ipi* 推导的氨基酸序列的疏水性 / 亲水性预测Fig.3 Predicted hydrophobicity/hydrophilicity of the sequence of *ipi* deduced amino acid from *H. annuus*

表2 向日葵 IPI 亚细胞定位预测

Table 2 The result of predicting subcellular location of IPI from *H. annuus*

	叶绿体 基质	叶绿体 类囊体膜	叶绿体 类囊体腔	线粒体 内膜间隙
可信度	0.950	0.800	0.800	0.467
结论	叶绿体基质蛋白			

于叶绿体基质。

6 向日葵 IPI 二级及三级结构的预测与分析

自然界中的功能蛋白都具有一定的三维结构, 具有空间结构的蛋白质才是生命的有效载

体。蛋白质的高级结构由其一级结构即氨基酸序列所决定。利用 SOPMA 程序(Geourjon 和 Deléage 1995; Combet 等 2000)对向日葵 IPI 氨基酸序列的二维结构进行预测的结果显示, 向日葵 IPI 由 41.02% 的 α -螺旋、18.98% 的 β -折叠、35.25% 的不规则卷曲和 4.75% 的 β -转角组成。

采用 SWISS-MODEL 在线同源建模方法对向日葵 IPI 进行分析, 并用 SweetMollyGrace 软件视图得到的结果(图4)显示: 向日葵 IPI 单体主要由 7 个螺旋、10 个不同大小的折叠和一些转角构成, 这与已研究清楚的人、大肠杆菌的 IPI 蛋白的高级结构一致, 都为紧密球形的 α/β 蛋白



图4 向日葵 IPI 三维结构

Fig.4 3-D structure of IPI from *H. annuus*

(Zhang 等 2007; Durbecq 等 2001)。与拟南芥 (Campbell 等 1997)、伯惠绣衣(预测结果, 未显示)等植物的 IPI 相比较, 向日葵 IPI 在高级结构上没有明显变化, 这预示其可能具有相同的功能。

7 向日葵 IPI 功能域的预测分析

用 Blast-Conserved Domains Search (Letondal 2001; Marchler-Bauer 和 Bryant 2004)分析向日葵 IPI 结构域的结果(图 5)显示, 在 81~268 位氨基酸间含有 IPP 异构酶结构域。该结构域蛋白属于 NUDIX 水解酶超家族成员, 含 NUDIX 基序, 需要 Mg^{2+} 或者 Mn^{2+} 的辅助构成具有活性的空间构像。采用 pfam 网站对向日葵 IPI 基序(motif)进行搜索, 发现其含 NUDIX 基序。但 NUDIX 基序的位点分布不如大肠杆菌集中, 在初级结构上没有集中的活性位点区域。

将向日葵 IPI 与其他几种植物及大肠杆菌 IPI

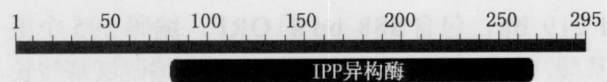


图5 向日葵 IPI 氨基酸序列功能结构域分析

Fig.5 Structure domain analysis of IPI from *H. annuus*

拟南芥IPI (1)	1	VMTDSSNDAGMDAVQRRLLMFEDECILVDENDRVVGHDKYNCHELMKIESENLLHRAFVSFVFLNFSKYELL
拟南芥IPI (2)	1	TMTDTKDAGMDAVQRRLLMFEDECILVDENDRVVGHDKYNCHELMENIEAKNLLHRAFVSFVFLNFSKYELL
伯惠绣衣	1	--MGEVVDAGMDAVQRRLLMFEDECILVDENDRVVGHESKYNCHELMKIESENLLHRAFVSFVFLNFSKYELL
向日葵	1	PPFGADADSTMDAVQRRLLMFEDECILVDENDRVVGHDKYNCHELMKIESENLLHRAFVSFVFLNFSKYELL
大肠杆菌	1	-----MQTEHVIILINAGVPTITLTKYAAHTADTR-----LHRAFSSMLFNKAGQLL
保守序列	1	m d dagmdavqrrllmfedecilvdendrvvgdtkync helmkie enllhrafsvflfnfskyl
		* * *
拟南芥IPI (1)	71	LQQRSTKVTFFPLVWNTCCSHPLYRESELIENMCLGVRMAAQRKLLDELGIWAEDVPVDEFTPLGRMLY
拟南芥IPI (2)	71	LQQRSTKVTFFPLVWNTCCSHPLYRESELIQDMALGVRMAAQRKLLDELGIWAEDVPVDEFTPLGRMLY
伯惠绣衣	69	LQQRSTKVTFFPLVWNTCCSHPLYRESELIENMCLGVRMAAQRKLLDELGIWAEDVPVDEFTPLGRMLY
向日葵	71	LQQRSTKVTFFPLVWNTCCSHPLYRESELIENMCLGVRMAAQRKLLDELGIWAEDVPVDEFTPLGRMLY
大肠杆菌	48	VTRFALSCKAMPGVWNTNSVCGHEQLGES-----NEDAVIRRCRYELGVEIT--PPESIYED--FRY
保守序列	71	lqqRs tkvtffplvwntccshplyrEselieen lgvrnAaqRkllldELGi aedvPvdef Pl rmlY
		Δ Δ Δ *
拟南芥IPI (1)	141	KAPSDGKWEHELDYLLFIVRDVWLPMPDEVAEIKYVREELKELVKADAGEEGKLSPWFRVLVDNPF
拟南芥IPI (2)	141	KAPSDGKWEHELDYLLFIVRDVWLPMPDEVAEIKYVREELKELVKADAGEEGKLSPWFRVLVDNPF
伯惠绣衣	139	KAPSDGKWEHELDYLLFIVRDVWLPMPDEVAEIKYVREELKELVKADAGEEGKLSPWFRVLVDNPF
向日葵	141	KAPSDGKWEHELDYLLFIVRDVWLPMPDEVAEIKYVREELKELVKADAGEEGKLSPWFRVLVDNPF
大肠杆菌	105	RAIDPSGIVEMEVCPVFAARTTSALQINDDEVMDYQCCD---LADVYLGIDATP--WAFSPWVQLQAT--
保守序列	141	kApsdgkwgEhEldyllfivrdv lqpNpDEVAeikyv reeLkelyvkaDAgeegvklSPWfrlrvldnfp
		* * *
拟南芥IPI (1)	211	LNKMWDHVEKGLTEAADMKTIHKL
拟南芥IPI (2)	211	LNKMWDHVEKGLTEAADMKTIHKL
伯惠绣衣	209	LNKMWDHVEKGLTEAADMKTIHKL
向日葵	211	LNKMWDHVEKGLTEAADMKTIHKL
大肠杆菌	168	-----NREARKKLSAFTQLK-----
保守序列	211	l kwdhvekgtl eadmKtihkl

图6 向日葵 IPI 与其他几种植物及大肠杆菌 IPI 的多序列比对

Fig.6 The Multi-alignment of IPI deduced amino acid sequences with other IPI proteins

保守序列标记氨基酸为活性位点氨基酸(三角符号标记), 星号标记位点为金属离子结合位点。

氨基酸序列利用 Clustal W 进行多序列比对分析的结果(图 6)显示: 植物 IPI 的 N 端区域没有同源性, 在这个区域后是高度同源的区域, 即 IPI 的功能区段, 可判定向日葵 IPI 同几种 IPI 类似, 为 I 型 IPI。参照 Durbecq 等(2001)对大肠杆菌 IPI 活性位点的研究发现向日葵 IPI 的活性位点为: H₁₀₃、H₁₁₅、S₁₁₉、R₁₃₄、K₁₃₈、C₁₅₀、H₁₅₂、K₁₇₅、E₁₇₉、Y₂₀₀、E₂₁₀、E₂₁₂、E₂₃₁ 和 W₂₆₂, 其中 H₁₀₃、H₁₁₅、H₁₅₂、E₂₁₀、E₂₁₂ 为几种 IPI 蛋白金属离子结合位点。其中只有 K₁₇₅ 与大肠杆菌 IPI 的 R₈₃ 有所变化, 但与所比对的其他植物 IPI 完全一致, 总体上看, 植物的 IPI 在功能位点上是非常保守的。

综上可以得到以下结论: 向日葵 *ipi* 基因全长 1 219 bp, 包含 888 bp 的 ORF, 编码 295 个氨基酸, 与拟南芥 IPI 和伯惠绣衣等在编码氨基酸长度上相似, 但与其他氨基酸序列在 N 端差别较大, 其他区域高度保守, N 端差异可能与各 IPI 定位相关。向日葵 IPI 与其他植物在核酸水平和氨基酸水平上都有很高的同源性, *ipi* 基因相对比较保守, 可能与其重要功能相关。与其他植物的 IPI 相比较, 向日葵 IPI 同其他植物一样, 除 N 端存在一疏水的区域外, 整体表现为亲水蛋白。IPI 在各个亚细胞结构中都有发现, 亚细胞定位预测显示, 向日葵 IPI 定位于叶绿体基质中, 属于质体 IPI。进一步功能位点分析显示, 向日葵 IPI 与其他植物及大肠杆菌功能活性位点上是一致的。向日葵 IPI 为紧密球型的 α/β 蛋白, 在高级结构上与其他植物 IPI 表现出一致性。

参考文献

- 李桂源, 钱骏主编(2004). 基于 WWW 的生物信息学应用指南. 长沙: 中南大学出版社, 143~146
- Altschul SF, Madden TL, Schäffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25: 3389~3402
- Campbell M, Hahn FM, Poulter CD, Leustek T (1997). Analysis of the isopentenyl diphosphate isomerase gene family from *Arabidopsis thaliana*. *Plant Mol Biol*, 36: 323~328
- Combet C, Blanchet C, Geourjon C, Deléage G (2000). NPS@: network protein sequence analysis. *Trends Biochem Sci*, 25 (3): 147~150
- Cunningham FX, Gantt E (2000). Identification of multi-gene families encoding isopentenyl diphosphate isomerase in plants by heterologous complementation in *Escherichia coli*. *Plant Cell Physiol*, 41: 119~123
- Durbecq V, Sainz G, Oudjama Y, Clantin B, Bompard-Gilles C, Triot C, Cailet J, Stalon V, Droogmans L, Villeret V (2001). Crystal structure of isopentenyl diphosphate: dimethylallyl diphosphate isomerase. *EMBO J*, 20: 1530~1537
- Gallagher CE, Cervantes-Cervantes M, Wurtzel ET (2003). Surrogate biochemistry: use of *Escherichia coli* to identify plant cDNAs that impact metabolic engineering of carotenoid accumulation. *Appl Microbiol Biotechnol*, 60: 713~719
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*, 31: 3784~3788
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005). Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed). *The Proteomics Protocols Handbook*. Totowa: Humana Press, 571~607
- Geourjon C, Deléage G (1995). SOPMA: significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci*, 11: 681~684
- Kajiwarra S, Fraser PD, Kondo K, Misawa N (1997). Expression of an exogenous isopentenyl diphosphate isomerase gene enhances isoprenoid biosynthesis in *Escherichia coli*. *Biochem J*, 324: 421~426
- Laule O, Fürholz A, Chang HS, Zhu T, Wang X, Heifetz PB, Gruißem W, Lange M (2003). Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA*, 100: 6866~6871
- Letondal C (2001). A web interface generator for molecular biology programs in Unix. *Bioinformatics*, 17: 73~82
- Marchler-Bauer A, Bryant SH (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res*, 32: 327~331
- Ramos-Valdivia AC, van der Heijden R, Verpoorte R (1997). Isopentenyl diphosphate isomerase: a core enzyme in isoprenoid biosynthesis. A review of its biochemistry and function. *Nat Prod Rep*, 14: 591~603
- Siddiqui MA, Yamanaka A, Hirooka K, Bamaba T, Kobayashi A, Imanaka T, Fukusaki E, Fujiwara S (2005). Enzymatic and structural characterization of type II isopentenyl diphosphate isomerase from hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Biochem Biophys Res Commun*, 331: 1127~1136
- Zhang C, Liu L, Xu H, Wei Z, Wang Y, Lin Y, Gong W (2007). Crystal structures of human IPP isomerase: new insights into the catalytic mechanism. *J Mol Biol*, 366: 1437~1446