

鲍思元. 水稻 *PLA-I/gamma/I* 类似基因的电子克隆及生物信息学分析[J]. 江苏农业科学, 2014, 42(5): 23-29.

# 水稻 *PLA-I/gamma/I* 类似基因的电子克隆及生物信息学分析

鲍思元

(湖北科技学院基础医学院, 湖北咸宁 437100)

**摘要:** 利用 NCBI 数据库及电子克隆技术克隆了水稻 *PLA-I/gamma/I* 类似基因, 命名为 *OsPLA-I/gamma/I*。序列分析表明, 其开放阅读框为 1 374 bp, 只含有 1 个外显子, 没有内含子。利用生物信息学软件对 *OsPLA-I/gamma/I* 蛋白的理化性质、疏水性/亲水性、跨膜区、亚细胞定位、蛋白质结构及功能等进行预测分析。结果表明: 该蛋白分子式为  $C_{2138}H_{3409}N_{625}O_{637}S_{19}$ , 氨基酸总数为 457 个, 分子量为 48 670.5, 理论等电点为 pI 8.86, 为亲水性稳定蛋白;  $\alpha$  螺旋和规则卷曲是其蛋白质二级结构的主要构件, 延伸链散布于整个蛋白质中; 属于脂肪酶, 活性受可逆磷酸化或糖基化作用调控, 具有能量代谢功能, 是水稻的生长因子; 定位于叶绿体基质腔中; 含有 19 个蛋白激酶磷酸化位点, “Ser” 磷酸化位点 11 个, “Thr” 磷酸化位点 5 个, “Tyr” 磷酸化位点 3 个; 含有 17 个 “O- $\beta$ -GlcNAc” 位点, 丝氨酸 5 个, 苏氨酸 12 个; 与粗山羊草、蒺藜苜蓿、甘蓝型油菜、拟南芥的脂肪酶有较高的同源性。本研究结果为进一步研究该基因的功能和作用机制奠定基础。

**关键词:** 水稻; 脂肪酶; 生物信息学; 氨基酸序列

**中图分类号:** Q943.2; S511.03

**文献标志码:** A

**文章编号:** 1002-1302(2014)05-0023-07

脂肪酶(lipase, E. C. 3. 1. 1. 3) 别称甘油三酯酯水解酶, 是一类能催化长链脂肪酸甘油酯水解为甘油、长链脂肪酸的酶类(或是该反应的逆反应), 广泛存在于植物、动物、微生物中<sup>[1]</sup>。脂肪酶是脂肪代谢最基本的酶, 在脂质代谢中发挥重要作用。脂肪酶属于  $\alpha/\beta$  水解酶超家族<sup>[2]</sup>, 在活性中心绝大多数拥有 “Gly-X-Ser-X-Gly” 保守序列<sup>[3-4]</sup>, “Ser” 与其他区域的 “Asp/Glu” 和 “His” 组成了一个有催化作用的三联体<sup>[5]</sup>, 该三联体位于酶的催化部位。与单体底物作用时, 脂肪酶活性不是很高; 与临界胶束浓度以上的不溶底物接触时, 覆盖脂肪酶活性位点的盖子打开, 暴露出疏水口袋, 酶的催化作用在疏水口袋里完成, 这种现象被称为界面活化<sup>[6]</sup>。自然界中绝大多数脂肪酶都具有这种现象, 但也有例外, 如 *Candida antarctica* lipase B 由于不含  $\alpha$  螺旋结构盖子, 所以无界面活化特性<sup>[1, 7]</sup>。

目前已有许多关于脂肪酶基因的克隆、序列结构分析、基因表达以及生产和应用的研究报道<sup>[3]</sup>。Alberghina 等最早报道了脂肪酶基因的测序和克隆<sup>[8]</sup>。Winkler 等发现了猪胰脂肪酶一级结构和人胰脂肪酶(HPL)一级结构的相似性及其在催化特性上的一致性<sup>[9]</sup>。人胰脂肪酶、*Rhizomucor miehei* 脂肪酶(RML)和 *Geotrichum candium* 脂肪酶的蛋白晶体结构已经被研究清楚<sup>[9-11]</sup>。近年来一些在植物中编码甘油三酯酯水解酶活性蛋白的基因被克隆<sup>[12-15]</sup>。Eastmond 在鼠耳芥中获得 1 个脂肪酶基因 *SDPI*, *SDPI* 编码的蛋白拥有 patatin 样酰基水解酶结构域, 主要是在发育的种子中表达<sup>[16]</sup>。研究者

从甘蓝型油菜中克隆了 2 个 GDSL 家族脂肪酶基因 *BnLIP1*、*BnLIP2*, 发现它们很可能参与油菜种子的萌发、花的形成、角果的发育<sup>[17-18]</sup>。Padham 等发现质体相关的脂肪酶基因 *At2g31690*, 并指出该基因可能通过动员质体小球内的三脂酰甘油来维持叶绿体的完整性<sup>[19]</sup>。Hong 等指出, 胡椒中的 GDSL 脂肪酶基因 *CaGLIP1* 对疾病易感性和非生物胁迫耐受性有调控作用<sup>[20]</sup>。最近的研究表明, 拟南芥脂肪酶基因在鼠耳芥的种子、根、茎、叶、花中都有表达, 尤其是在衰老叶片中有高量表达<sup>[21]</sup>。这些说明脂肪酶在植物萌发、生长发育、衰老过程中发挥重要作用。

长期以来, 水稻等作物种子脂类与陈化的研究受到人们关注, 但是脂肪酶的作用一直未得到重视。水稻脂肪酶的研究也仅局限于米糠稳定化和大米风味营养的问题, 而关于水稻脂肪酶分子的研究则更少。本研究在 Blast 检索和生物信息学分析的基础上, 利用电子克隆方法获得水稻 *PLA-I/gamma/I* 类似基因, 并对其编码蛋白质的理化性质、结构、功能进行分析, 以期为进一步研究 *PLA-I/gamma/I* 类似基因在水稻中的作用机制奠定基础。

## 1 材料与方法

### 1.1 材料

数据资料来源于 NCBI 核酸及蛋白质数据库中已注册的核酸序列及对应的氨基酸序列: 节节麦 (*Aegilops tauschii*, EMT30786. 1)、蒺藜苜蓿 (*Medicago truncatula*, XP\_003621036. 1)、甘蓝型油菜 (*Brassica napus*, ACJ76846. 1)、鼠耳芥 (*Arabidopsis thaliana*, NM\_100557)。

### 1.2 方法

利用生物信息学数据库和互联网上的软件进行分析。用 FGENESH 软件进行基因预测; 用 EMBOSS Transeq 软件推导

收稿日期: 2013-08-26

基金项目: 湖北科技学院科研课题(编号: SK0612)。

作者简介: 鲍思元(1976—), 男, 湖北麻城人, 硕士, 讲师, 研究方向为植物分子遗传学和生物技术。E-mail: baosyuan@163.com。

氨基酸序列;用 EMBOSS Needle 软件分析序列同源性;用 DNAMAN V6 软件进行氨基酸的多序列比较和进化树分析;用 ProtParam 软件对蛋白质进行氨基酸组成、分子量、等电点及疏水性等理化性质分析;用 PREDATOR 软件预测蛋白质二级结构;用 Pepinfo 软件分析蛋白质疏水性/亲水性;用 TargetP 软件预测亚细胞定位;用 ChloroP 软件预测蛋白质转运肽;用 TMHMM 软件预测蛋白质跨膜结构;用 InterProScan 软件预测蛋白质功能;用 ProtFun 软件预测蛋白质功能类别和酶种类;用 NetPhos 软件预测磷酸化位点;用 YinOYang 软件预测糖基化位点;用 SWISS - MODEL 软件同源建模预测蛋白质的三维结构。

## 2 结果与分析

### 2.1 水稻脂肪酶 PLA-I/*gamma*/I 类似基因的电子克隆及基本信息

在 GenBank 的核苷酸数据库中获得拟南芥脂肪酶基因 *PLA-1 {gamma} 1* 的全长“CDS”(GenBank 登录号:

CGTACCCCGGTGTCAGACCCGCAACCGACCACCTTTTCCATTAAAGAAAACGCTGCTGCTCATCTCATCTCCACCCGTTTTCCTCGTC  
TCGCTAGTCAAATAATCCATACATCTCTCTCTTTCTCTCTCTCTCTCTCCTCAAGTTGCTTACGTGGCCAACCGCGTGCGTCGCCGTGCG  
CCTCCGATATAAAATATGTACGCAACGCGCGCTACCTTACCGCCAACCGCCACTCACTGTACTCATCCATCGATCGCTAGCTGCGCA  
CGCCACTGCGGCCGCGTCCGCTCACTTAACCTGGCTGCTCGATGCCGCCGCGCATGGCGGCAGCGCGGTGGCGCGGCTCACT  
GCGGCGACGACGTGCGCGATGCGCGTCACTGGCCCGGTCACTGCGCGACGACGAGTGGCTGCCTGCCGGTGACCGTCACTCTG  
GGTGATCTGCCACAGCAGAAGCGATCGGCCGCCACGGCGACGACGACGACGAGTAGTGGTGGTTCTCCTGTGGTGGTGGAGAAG  
AGAGCGGCGGCGCGACGCGCGGCGAGGAGGAGGAGGTGAGGAGGCGGCGGCGGTGGCGGCGCTGGCGAGCGCGTGCGCGGG  
AGGTGCAAGGGCGCGCGCATGGAGGGGACATGGTGAGAGCGCTCCACCCGCTGCTCCGGGACGAGGTGGTCCGGTACGGGGAG  
CTCGTCGCGCGCTGCTACCGGGGCTCGACCTGACCCCGGCTGCAAGCGCTACCTCAACTGCAAGCAGCGCAAGAAGCAGATG  
TGC CGCGCGGTGCGCATGGCCGGCGCGGTACACGGTACCAAGTACATCTACGCCGCCCGACGTCGCGTGCCTTCCGCTCGGC  
GTCGGCGGCGCGTCTCTGCGCCGGCAAGAGCCGGTGGATCGGCTACGTGCGCGTTCGCGTCCAACCGCGAGGCCGCCCGCTC  
GGCCGCCGCGACATCTCTGCTCTCTTCCGCGGCACGGTGACCGGATCGGAGTGGCTGGCCAACCTTCATGAGCGCGCTGTCCGCGG  
CGCGGTTTCGACCCGCGCCGACCCGCGCCCGACGTCGCGCTGAGTCGGGTTCCTCTCCCTCTACACCTCCGACGACCTCTCCGG  
CAAGTTCACTTCGCGGAGCTGCCGGAACCAAGCTCTCTCCGAGGTGACCCGCTCATGACAAGTACAAGACGACGACGTGACG  
CATCACCTCGCCGACACAGCATGGGGAGCTCGCTGGCGATCCTCTCGGCTACGACCTCGCCGAGCTGGGCTCAACCGCGG  
CGGCCGCGGCGCGCGCGCGGCCATCCGATCACCGTCTTCTCTTCCGCGGCCCGCGCGCTCGGCAACCTCGAGTTCAAGAG  
ACGGTGCGATGAGTCCGCGTCAAGGTGCTCCGCGTTCGCGAACGCCGCGACCCGGTGACGAGGATGCCCGGCGTGGTGTCAA  
CGAGGCCGCGCTAGGGTTTTCGGGGTTCGAGCTGCCGTGGAGCAAGGCGTGTACACCCACGTCGGGCTCGAAGTCCGCCCTCA  
TTTCTTCAACGGCCAGCCACGCGGCTGCGTCCACGACCTCGACGCTACATAAAATCACTCTTCGACGCGCGGCGCGCGGCTCG  
GCGACGACGACGACCTTGAGATCGTGAGGTGGCAGATGGCGGCGATTCTGTGAGGCGAGTGGCTGCAGACGCTAGGGATTGGA  
AGCGTCCACTAATTAATATCTGTAATCTGTACGATGTTTACAGTTGTATAGTTGTACTATTTTCATGTCTGTGGATCGATCAAGAT  
TCAAGCGGGGAGAAATTAAGAGGACAAAACAAGTAAACATCTAACCAAAATTAAGTCTGAGCTCTTGATTAGGGTTTGCAATTTCCG  
TTGCCGCTGAAATTTCCGTGGTTGCCGAAAAATTTGAAATTTCCGTTATTTCCGACTTTTCTGACTGAAATTAATTTGAATTTTAAAC  
ATATCTTATGTAATTTGATAAAATTTTAAACAAATTTACAAATATTGGACAAACAA

方框：启动子起始位点；下划线：起始密码子；灰色底纹：终止密码子；灰色底纹方框：“PolA”开始位点

图1 *OsPLA-I{gamma}1* 基因的DNA序列

MAAAAVARLTAATTCMRVSGPVTAAATTSASPVTVTGLDLPQQRSAATATTTTTSSGGSPVVVEKRAAAATAAEVEVEEAAA  
VAALASAWREVQGAGDWRGMVEPLHPLLREVVRYGELVAACYRAFDLPASKRYLNCKHGGKQMLPAVGMAGAGYTVTRY  
IYAAPDVALPFGVGGRCSCAGKSRWIGYVAVASNREARLGRRDILVSFRGTVTGSEWLANFMSALSAPRFPADPRPDVRVESG  
FLSLYTSDDLSGKFTCGSCRNQLLSEVTRLIDYKHDDVSITLAGHSMGSSLAILLGYDLAELGLNRGGRRGGGAIPITVFSFGGP  
RVGNLEKKRCRCDLRKVLVRVANARDPTVMPPGVVLDNEAAARVFRVELPWSKACYTHVGVEVALDFFKASHAACVHDLDAYIN  
HLLDGGAAASATTTTLESVRWQMAAIRAGEWLQTLGI

图2 *OsPLA-I{gamma}1* 基因推导的氨基酸序列

## 2.2 蛋白质序列同源性分析

在 GenBank 的蛋白质数据库中获得粗山羊草、蒺藜苜蓿、甘蓝型油菜、拟南芥脂肪酶基因编码蛋白质的氨基酸序列,登录号分别为 EMT30786.1、XP\_003621036.1、ACJ76846.1、NM\_100557。用 Needle 软件将其分别与 *OsPLA-1|gamma|1* 基因推导的氨基酸序列进行同源性分析,结果表明它们与 *OsPLA-1|gamma|1* 蛋白的相似性分别为 74.7%、52.7%、49.8%、50.3%。说明这些物种的脂肪酶基因氨基酸序列之

NM\_100557; locus\_tag: AT1G06800) 序列, 以“NM\_100557”序列为探针在水稻“Gene in MUS RGAP Release 7 - Genomic Sequence”数据库中进行“blastn”。在“blastn”结果中选取“LOC\_Os11g19290”作为水稻脂肪酶 *PLA-1* {*gamma*} 1 类似基因序列, 其“Score”值为 496, “E”值为  $1.0 \times 10^{-13}$ , “Top Id”为 57.19%。在 RGAP 网站上对“LOC\_Os11g19290”进行“Genome Browser”, 显示其位于第 11 号染色体 11 050 768 ~ 11 052 141 bp 的区段上, 且位于负链上, 全长 1 374 bp。

为了获得水稻脂肪酶基因的完整核苷酸序列,选取第 11 号染色体负链上 11 050 455 ~ 11 052 454 bp 的 2 kb 基因组序列用 FGENESH 软件预测水稻脂肪酶基因,预测到 1 个“ORF”,位于 314 ~ 1 687 bp,全长 1 374 bp,只含有 1 个外显子,没有内含子(图 1)。这说明“LOC\_Os11g19290”就是水稻脂肪酶基因的“CDS”,将其命名为 *OsPLA - I / gamma / 1*。用 EMBOSS Transeq 软件推导 *OsPLA - I / gamma / 1* 基因编码蛋白质的氨基酸序列,结果发现蛋白质全长 457 个氨基酸残基(图 2)。

间具有较高的相似性,它们在初级结构上具有相当的保守性。蛋白质氨基酸序列有较高的同源性,说明它们具有相似功能,因此推测 *OsPLA-1/γ1* 是水稻的脂肪酶基因。

### 2.3 蛋白质理化性质分析和疏水性/亲水性预测

用 ProtParam 软件预测 OsPLA-1 蛋白质的理化性质,推测该蛋白质的分子式为  $C_{2138}H_{3409}N_{625}O_{637}S_{19}$ ,氨基酸总数为 457 个,分子量为 48 670.5,理论等电点 pI 为 8.86。负电荷氨基酸残基 (Asp + Glu) 总数为 45 个,正电荷氨基酸

残基( Arg + Lys)总数为 52 个,净电荷总数为 7 个。总的亲水性平均系数( GRAVY)为 -0.001,预测该蛋白属于亲水性蛋白。

白。预测该蛋白不稳定指数是 38.59,属于稳定蛋白。其氨基酸组成见表 1。

表 1 OsPLA-I{gamma}I 蛋白质氨基酸组成

氨基酸	数量(个)	占比(%)	氨基酸	数量(个)	占比(%)	氨基酸	数量(个)	占比(%)
Ala(A)	69	15.1	Arg(R)	38	8.3	Asn(N)	9	2.0
Asp(D)	22	4.8	Cys(C)	10	2.2	Gln(Q)	7	1.5
Glu(E)	23	5.0	Gly(G)	42	9.2	His(H)	8	1.8
Ile(I)	11	2.4	Leu(L)	41	9.0	Lys(K)	14	3.1
Met(M)	9	2.0	Phe(F)	13	2.8	Pro(P)	19	4.2
Ser(S)	31	6.8	Thr(T)	30	6.6	Trp(W)	8	1.8
Tyr(Y)	12	2.6	Val(V)	41	9.0			

疏水性/亲水性分析对于预测蛋白质二级结构和功能域具有重要生物学意义。采用 Pepinfo 软件对 OsPLA-I{gamma}I 蛋白进行疏水性/亲水性预测,结果表明多肽链 164 位的 Arg(R)氨基酸亲水值最高,为 0.624,77~79 区域的 Glu(E)、Glu(E)、Val(V)等 3 个氨基酸亲水值最低,为 -0.801。

总体来看,亲水氨基酸均匀分布在整条肽链中,且数量远多于疏水氨基酸。总的亲水性平均系数为 -0.085 99,依据氨基酸分值越低、亲水性越强的规律,可推测 OsPLA-I{gamma}I 蛋白属于亲水性蛋白(图 3)。这与用 ProtParam 软件预测的结果一致。

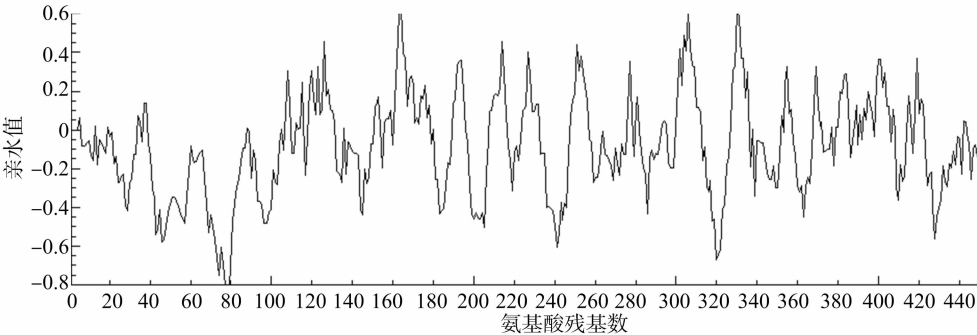


图3 OsPLA-I{gamma}I蛋白疏水性/亲水性分析

2.4 蛋白质二级结构预测

用 PREDATOR 软件预测蛋白质二级结构,结果表明该氨基酸序列中共有  $\alpha$  螺旋(Hh) 164 处,占总二级结构的 35.89%;延伸链(Ee) 59 处,占总二级结构的 12.91%;无规

则卷曲(Cc)234 处,占总二级结构的 51.20%。 $\alpha$ 螺旋和无规则卷曲是 OsPLA-I{gamma}I 蛋白主要的构成元件,延伸链散布于整个蛋白质中(图 4)。

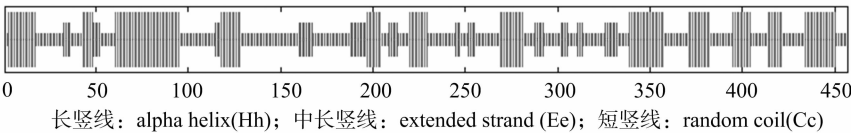


图4 OsPLA-I{gamma}I蛋白二级结构预测

2.5 亚细胞定位预测和蛋白质跨膜结构预测

用 TargetP 软件预测亚细胞定位,结果表明 OsPLA-I{gamma}I 蛋白位于叶绿体,转运肽位于 1~67 区段,全长 67

个氨基酸(图 5)。与多肽链氨基酸序列相比,该区段氨基酸富含碱性氨基酸,苏氨酸、丝氨酸含量较多(13 个苏氨酸和 7 个丝氨酸)。这些特征符合叶绿体转运肽的特征。

名称	长度	叶绿体 转运肽	线粒体 靶向肽	分泌途径 的信号肽	其他	位置	可靠性 等级	运输肽 长度
IOC_Osllg19290.1	457	0.718	0.534	0.002	0.014	C	5	67
下限		0.000	0.000	0.000	0.000			

图5 用TargetP软件对OsPLA-I{gamma}I蛋白质进行亚细胞定位预测

用 ChloroP 软件对 OsPLA-I{gamma}I 蛋白进行转运肽预测,该预测结果与 TargetP 软件预测结果一致(图 6)。说明 OsPLA-I{gamma}I 蛋白由核基因组编码,在细胞质中合成后,在转运肽引导下进入叶绿体,然后转运肽在肽酶作用下被切除。

用 TMHMM 软件进行跨膜结构预测,结果表明 OsPLA-I{gamma}I 蛋白不含跨膜结构,说明 OsPLA-I{gamma}I 蛋白不是膜蛋白(图 7)。结合该蛋白质的亲水特性,推断 OsPLA-I{gamma}I 蛋白位于叶绿体基质腔中,其催化功能在基质腔中完成。

名称	长度	分值	叶绿体转运肽	裂解位点得分	叶绿体转运肽长度
IOC_Osllg19290.1	457	0.578	是	8.276	67

图6 用ChloroP软件对OsPLA-I{gamma}1蛋白质进行转运肽预测

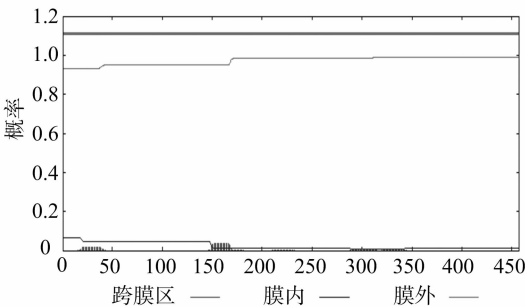


图7 OsPLA-I{gamma}1蛋白跨膜结构预测

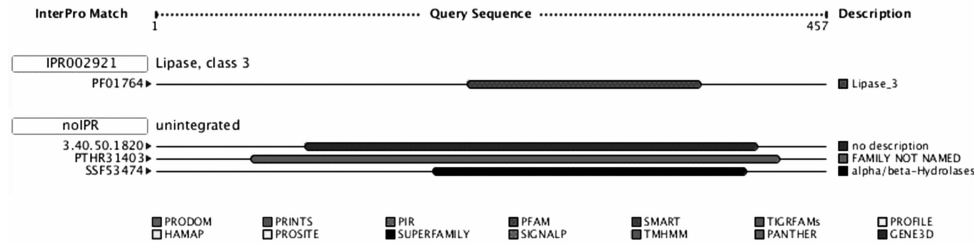


图8 OsPLA-I{gamma}1蛋白功能结构域预测

酶,具有能量代谢功能,是水稻生长因子。由于脂肪酶催化长链脂肪酸甘油酯水解为甘油和长链脂肪酸,并且能进行逆反应,所以脂肪酶为连接酶,这与预测结果一致。脂肪酶催化长链脂肪酸甘油酯水解生成的脂肪酸通过氧化分解产生能量,具有能量代谢功能,这也与预测结果一致。所以推测 OsPLA-I{gamma}1 蛋白属于脂肪酶。

2.7 翻译后修饰位点预测

用 NetPhos 软件进行“serine”、“threonine”、“tyrosine”磷酸化位点预测,结果显示 OsPLA-I{gamma}1 蛋白具有 19 个蛋白激酶磷酸化位点,“Ser”磷酸化位点(32、56、57、60、198、213、232、249、261、297、438)11 个,“Thr”磷酸化位点(10、14、55、217、434)5 个,“Tyr”磷酸化位点(193、392、418)3 个(图 9)。说明磷酸化是该蛋白翻译后修饰的主要方式之一。

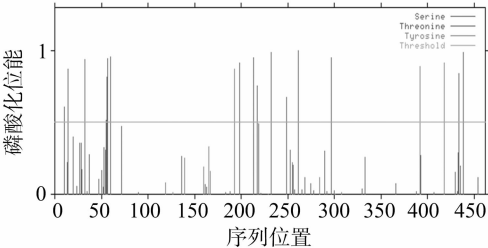


图9 OsPLA-I{gamma}1蛋白磷酸化位点预测

大多数脂肪酶是糖蛋白,但也有一些不含糖基的脂肪酶,主要取决于脂肪酶一级结构中是否存在糖基化识别序列。用 YinOYang 软件进行糖基化位点预测,结果表明 OsPLA-I{gamma}1 蛋白有 17 个“O-β-GlcNAc”位点,丝氨酸 5 个,苏氨酸 12 个,说明 OsPLA-I{gamma}1 蛋白也是糖蛋白。其

2.6 蛋白质功能结构域及蛋白质功能预测

用 InterProScan 软件进行蛋白质功能结构域预测,结果表明 OsPLA-I{gamma}1 蛋白多肽链 212~371 区域含有“Lipase\_3”(IPR002921)保守结构域,189~403 区域含有“alpha/beta-Hydrolases”(SSF53474)保守结构域,65~425 区域含有 1 个未被命名的家族(PTHR31403)(图 8)。故 OsPLA-I{gamma}1 蛋白属于脂肪酶家族,并归属于 α/β 水解酶超家族。

用 ProtFun 软件进行功能类别和酶的种类预测,结果表明 OsPLA-I{gamma}1 蛋白是酶蛋白,在酶分类上属于连接

中 55 位苏氨酸、56 位丝氨酸、434 位苏氨酸在细胞生长不同时期可动态可逆地被糖基化或磷酸化修饰(图 10)。这说明 OsPLA-I{gamma}1 蛋白活性受可逆磷酸化作用或糖基化作用调控,在信号分子作用下 OsPLA-I{gamma}1 蛋白处于活性状态或失活状态。

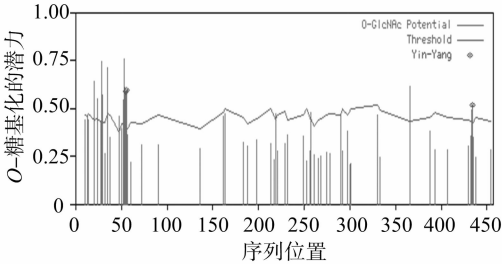


图10 OsPLA-I{gamma}1蛋白糖基化位点预测

2.8 氨基酸的多序列比较和进化树分析

将水稻(Oryza sativa, LOC\_Os11g19290)、节节麦、藜苳苳、甘蓝型油菜、鼠耳芥脂肪酶基因氨基酸序列用 DNAMAN V6 软件进行氨基酸的多序列比对分析。如图 11 所示,其氨基酸一致性为 56.69%。以往研究表明,脂肪酶在活性位点具有“GXSXG”特征序列。序列对比发现,5 种植物都具有“GXSXG”特征序列,这进一步说明 OsPLA-I{gamma}1 是水稻的脂肪酶基因。在多序列比对分析的基础上,用 DNAMAN V6 软件绘制系统发育树(图 12),结果显示水稻脂肪酶与节节麦脂肪酶的亲缘关系最近,序列同源性最大;水稻脂肪酶与甘蓝型油菜、鼠耳芥脂肪酶的亲缘关系最远,序列同源性最小。这是因为水稻与节节麦同属于禾本科植物,在进化上亲缘关系较近。这与 Needle 软件分析结果一致。

LOC_0=11g19290.1	ISAAAVARLT AATT CRMRVSGPUTAATTS AASPUTUTLGD	40
EMT30786.1	ISVAPSS...PTQRLRACMP...AAASATHVALRA	30
XP_003621036.1	ISSITHLPN..KTSM INITFPQN.FHPKTHPFGQVSLPN	37
ACJ76846.1	ISTIPSHNLLLPNPTINQSAHSL.S.FKP...QSTLLNFFA	36
AT1G06800.1_1	ISTIPSHN.LRPHTTNQRTQYSL.S.FRPHTFSRSTLITFFA	38
Consensus	ma	
LOC_0=11g19290.1	LPQKRSAAATATTT.....SSGGSPUUVKRA	68
EMT30786.1	SNEPHKAAATTATT.....VAUPUP	50
XP_003621036.1	NMHLSPSLKSSIVASMSALLEIILPTTTNNNNNIYNTIS	77
ACJ76846.1	RSSPAAUTRAVSRTDG.....ASISRLPEVEK	64
AT1G06800.1_1	RSSPA...RAMSRTDEE.....ASISTRLE.QES	63
Consensus		
LOC_0=11g19290.1	AAATAAEVEEVEEAASVAA.....LASAOREVQAGD	100
EMT30786.1	TAVVAPADRRSDAQAAPS.....LACTMEDVHCADD	82
XP_003621036.1	TSITPKHTQKITHKISTSTTITTTKHNANIOREIQGCNN	117
ACJ76846.1	YEITAAGDURRDREAKET.....KSRDRTORKIQGEDD	98
AT1G06800.1_1	YGLTTAEDIRRDGEAKES.....KSRDRTORKIQGEDD	97
Consensus		l w g
LOC_0=11g19290.1	WRGLVEPLHPLRLDDVURYGELVAAACYPALDLPASKRYL	140
EMT30786.1	WRGLVEPLHPLRLADVURYGELGGACYPALDLNPSKRYL	122
XP_003621036.1	WENLLNPLHPLRLREIIRYGEFUTSYKAFDLTNSKRYL	157
ACJ76846.1	WAGLMDPMDPDLRSELIRYGEHAQACVDFDFDFPSRYCG	138
AT1G06800.1_1	WAGLMDPMDPDLRSELIRYGEHAQACVDFDFDFPSRYCG	137
Consensus	w p p l r e r y g e y a d s	
LOC_0=11g19290.1	NCKKGKKQMLPAVCHAGAGYTUTRNIVASAPDVALPFGVGG	180
EMT30786.1	NCKHGKRQILQAVCHADAGYQUTKNTVSAPDVPLPMGICR	162
XP_003621036.1	NCKYGGKKNMLKEVCHENCYETKNTVATP...PMIMEN	193
ACJ76846.1	SCRITRKKLFDSLGIIDSGYEAAALYATSNINLPMFFSK	178
AT1G06800.1_1	SCRITRRHLFDSLGIIDSGYEAAALYATSNINLPMFFSK	177
Consensus	c g gy y y p	
LOC_0=11g19290.1	RCSC...AGKSRMIGYUAVASNREAA..RLGRRDILUSFR	215
EMT30786.1	PCS.....KSRMIGYUAVASDREAG..RLGRRDILUSFR	194
XP_003621036.1	NSS.....GRMIGYUAVSSDDSYK..KLGRRDIVUTFR	224
ACJ76846.1	SRWSKVWSKNANMIGYUAVSSDDSEATRHLGRRDIAIAWR	218
AT1G06800.1_1	SRWSKVWSKNANMIGYUAVSSDDNEATRCLGRRDIAIAWR	217
Consensus	w gyvav lgrrdi r	
LOC_0=11g19290.1	GTUTGSEDLANFMSALSPARFD.PADPRPDORUESGFLSL	254
EMT30786.1	GTUTGSEDLANFMSALAPARFD.PADPRPDORUESGFLSL	233
XP_003621036.1	GTUTNQEWISNLMSSLTASLD.PNNQLPNQKVESGFLSL	263
ACJ76846.1	GTUTQLEWIAIDLKDLKPVSGNGFRCPDFAOKAESGFLDL	258
AT1G06800.1_1	GTUTRLEWIAIDLKDLKPVSGNGFRCPDFAOKAESGFLDL	257
Consensus	gtvt ew l p p v esgfl l	
LOC_0=11g19290.1	YTSDDLSGKFTCGSCRNQLLSDUTRLIDKYKH...DVSII	291
EMT30786.1	YTSDNGTGKFTAGSCRNQLLSDISRLIAENKQE...DMSII	270
XP_003621036.1	YTSDESSSKFGLQSCREQLLSDUSRLMKKHKGK...NVSI	301
ACJ76846.1	YTDKDTSCNFSKFSAREQLTEUKLVERYGDEEGGDLII	298
AT1G06800.1_1	YTDKDTSCNFSKFSAREQLTEUKLVERYGDEEGEELSI	297
Consensus	yt f s r q l e r l si	
LOC_0=11g19290.1	TLAGHSMGSSALHLLGYDLDELGLMRGGGGRGGAIPITU	331
EMT30786.1	TLAGHSMGSSALHLLGYDLDELGLMRGVHGG...AIPITU	307
XP_003621036.1	SLAGHSMGSSALHLLAYDISLGLMKNKNDKN...DASVTV	338
ACJ76846.1	TUTGHSLEGALAL SAYDVDEGLMRTKNGK...VUPUTV	335
AT1G06800.1_1	TUTGHSLEGALAL SAYDVDEGLMRTKNGK...VIPUTA	334
Consensus	ghs g l a l y d e g n t	

方框表示脂肪酶活性位点特征序列“GXSGXG”

图11 水稻脂肪酶与其他植物脂肪酶的氨基酸序列比较

LOC_0s11g19290.1	FSPGGPRUGNLEFKRRCDLGRUKULRUSNARDPVTMPGV	371
EMT30786.1	FSPGGPRUGNLEFKMRCDLGRUKULRUSNARDPVTMPGV	347
XP_003621036.1	FSPGGPRUGNLEFKRRCEDLGRUKULRISNAMPITKLPGV	378
ACJ76846.1	FTYSAPRUGNIRFKERMEELGRUKULRVUNKNHUVPKSPGL	375
AT1G06800.1_1	FTYGGPRUGNIRFKERIEKLGRUKULRVUNKNHUVAKSPGL	374
Consensus	f p r v g n f k r l v k v l r n d p g	
LOC_0s11g19290.1	VLNFAARVTF..RVELPWSKACYTHUGUEVALDFFK....	405
EMT30786.1	LNIESARVLGG.RYEMPWSKACYAHUGUEVALDFFK....	382
XP_003621036.1	VFINENFRVLNGGRYEFPWSCSCYAHUGUELMLDFFN....	414
ACJ76846.1	FLNEHAPHAL..KQLAGGLPWICYCHUGGEKLALDHQNSPFL	413
AT1G06800.1_1	FLNERAPQAL..MKLAGGLPWICYSHUGEMLPIDHMQKSPFL	412
Consensus	ne cy hvg ld	
LOC_0s11g19290.1	..ASHAACVHDLDAYINHLDG.....GAAAS.....	430
EMT30786.1	..AGDIACVHDLLEVYIDQLKIP..KVDVATASSVM.....	414
XP_003621036.1	..MQNPSCVHDLDTYIGLRCPK..KKIEEVINSQSDGGGV	451
ACJ76846.1	KPSVDISTAHMLEALLNLDGYHGXGQRFVLSGGRDPALV	453
AT1G06800.1_1	KPTVDLSTAHMLEALLNLDGYHGXGQRFVLSGGRDPALV	452
Consensus	h l l	
LOC_0s11g19290.1	.....ATTTLESWRQMAAIRAG.....EMLQTLGI.	457
EMT30786.1	.....RUPSMFESWRQMAAIRAG.....EMLRALGI.	441
XP_003621036.1	NFLEKARMTLLKSSISQVFAAGGWN...YHDFLSSVSTD	488
ACJ76846.1	NKASDFLKDHIMVPPYWRQDANKGMVRNTDGRWIQPDRIR	493
AT1G06800.1_1	NKASDFLKDHIMVPPYWRQDANKGMVRNTDGRWIQPDRIR	492
Consensus		
LOC_0s11g19290.1	.....	457
EMT30786.1	.....	441
XP_003621036.1	ILCSWSDELLFGLVLLLL....	506
ACJ76846.1	AEDNHAPDINHLLTQLNHPS..	513
AT1G06800.1_1	ADDQHAPDINHLLTQLNHPSQL	514
Consensus		

续图 11

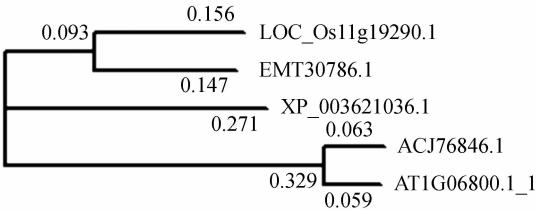


图12 水稻脂肪酶氨基酸序列与其他植物脂肪酶氨基酸序列的系统发育树

2.9 三级结构预测

采用 SWISS - MODEL 软件同源建模的方法预测 OsPLA - I|gamma|1 蛋白的三级结构。以在“PDB”晶体库中与 OsPLA - I|gamma|1 蛋白一致性达 31% 的“Fungal lipases”晶体结构(PDB id:1tibA)作为模板,预测 OsPLA - I|gamma|1 蛋白三级结构(图 13)。在预测的 3D 分子结构中,β 折叠位于脂肪酶分子的中央,α 螺旋包围在 β 折叠外面,形成近似球形的蛋白质分子,该结构符合 α/β 水解酶的结构特点。

3 结论与讨论

电子克隆技术具有快速、低成本、目标明确等优点,特别适用于条件不完善的小型实验室,但也受到生物信息资源的

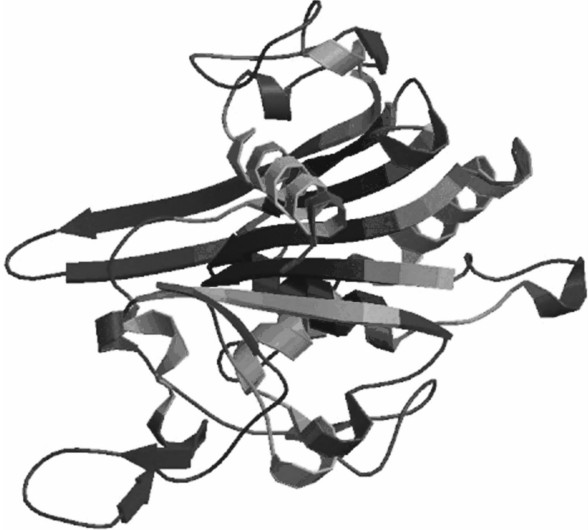


图13 OsPLA-I|gamma|1蛋白的三级结构预测模型

限制。随着生物基因组序列、EST 序列信息的增加,运用电子克隆技术识别克隆新基因已经成为基因克隆的常用方法。电子克隆技术已经在人类和动物克隆中得到大量应用<sup>[22-26]</sup>。随着植物基因组核苷酸数据库、EST 数据库的日益丰富,目前

电子克隆技术在一些植物中也得到成功应用<sup>[27-31]</sup>。目标基因一旦被成功识别,将为新基因的结构分析、功能鉴定等分子生物学研究奠定良好基础。

本研究结合分子生物学与生物信息学,运用电子克隆技术获得了水稻 *PLA-I/γ1* 类似基因的核苷酸序列。应用生物信息学分析技术对水稻 *OsPLA-I/γ1* 蛋白的理化性质、二级结构、疏水性/亲水性、亚细胞定位、蛋白质跨膜结构、翻译后修饰、功能结构域及蛋白质功能、同源性、高级结构进行了预测分析,结果表明,*OsPLA-I/γ1* 蛋白属于带正电荷的亲水性稳定糖蛋白; $\alpha$ 螺旋和不规则卷曲是其蛋白质二级结构的主要构件;存在多个磷酸化位点和糖基化位点;定位于叶绿体基质腔中;属于脂肪酶家族,并归属于 $\alpha/\beta$ 水解酶超家族,具有能量代谢功能,是水稻的生长因子;在参入代谢时,受可逆磷酸化作用或糖基化作用调控。

为了提高预测的准确性,本研究在预测 *OsPLA-I/γ1* 蛋白的疏水性/亲水性、亚细胞定位、氨基酸序列同源性、蛋白质功能时均采用 2 个不同软件。由于不同预测程序分析时所依据的原理和算法有不同的侧重性,其预测结果中的一致性部分具有较高的可信度。本研究结果为深入研究 *OsPLA-I/γ1* 基因的表达信息、生物学功能奠定基础。

#### 参考文献:

- [1] Jaeger K, Reetz M T. Microbial lipases from versatile tools for biotechnology[J]. Trends in Biotechnology, 1998, 16(9): 396-403.
- [2] Ollis D L, Cheah E, Cygler M, et al. The alpha/beta hydrolase fold[J]. Protein Engineering, 1992, 5: 197-211.
- [3] 邹显章, 邹敏辰. 脂肪酶分子生物学的研究进展[J]. 无锡轻工大学学报, 2002, 21(1): 94-98.
- [4] 邹敏辰. 脂肪酶基因结构和氨基酸序列的比较[J]. 江南学院学报, 2001, 16(2): 1-4.
- [5] McKillican M E. Lipid changes in maturing oil-bearing plants. IV. Changes in lipid classes in rape and crambe oils[J]. J American Oil Chemists' Society, 1965, 43(7): 461-465.
- [6] Svendsen A. Lipase protein engineering[J]. Biochimica et Biophysica Acta, 2000, 1543: 223-238.
- [7] Ferrato F, Carriere F, Sarda L, et al. A critical reevaluation of the phenomenon of interfacial activation[J]. Methods in Enzymology, 1997, 286: 327-347.
- [8] Alberghina I, Schmid R D, Verger R. Lipases: structure, mechanism and genetic engineering[J]. GBF Monographs, 1991, 16: 361-364.
- [9] Winkler F K, D'arcy A, Hunziler W. Structure of human pancreatic lipase[J]. Nature, 1990, 343: 771-774.
- [10] Brady L, Brzozowski A M, Derewenda Z S. A serine protease triad forms the catalytic centre of a triacylglycerol lipase[J]. Nature, 1990, 343: 767-770.
- [11] Schrag J D, Li Y G, Wu S, et al. Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*[J]. Nature, 1991, 351(6329): 761-764.
- [12] Eastmond P J. Cloning and characterization of the acid lipase from castor beans[J]. Journal of Biological Chemistry, 2004, 279(44): 45540-45545.
- [13] El-Kouhen K, Blangy S, Ortiz E, et al. Identification and characterization of a triacylglycerol lipase in *Arabidopsis* homologous to mammalian acid lipases[J]. FEBS Letters, 2005, 579(27): 6067-6073.
- [14] Matsui K, Fukutomi S, Ishii M, et al. A tomato lipase homologous to DAD1 (LeLID1) is induced in post-germinative growing stage and encodes a triacylglycerol lipase[J]. FEBS Letters, 2004, 569(1/2/3): 195-200.
- [15] 张慧莉, 高剑峰, 刘伟. 陆地棉脂肪酶基因克隆及氨基酸序列分析[J]. 西北植物学报, 2008, 28(9): 1734-1737.
- [16] Eastmond P J. SUGAR-DEPENDENT1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating *Arabidopsis* seeds[J]. The Plant Cell, 2006, 18(3): 665-675.
- [17] Jacks T J, Yatsu L Y, Altschul A M. Isolation and characterization of peanut spherosomes[J]. Plant Physiology, 1967, 42(4): 585-597.
- [18] Ling H, Zhao J, Zuo K, et al. Isolation and expression analysis of a GDSL-like lipase gene from *Brassica napus* L[J]. Journal of Biochemistry and Molecular Biology, 2006, 39(3): 297-303.
- [19] Padham A K, Hopkins M T, Wang T W, et al. Characterization of a plastid triacylglycerol lipase from *Arabidopsis*[J]. Plant Physiology, 2007, 143(3): 1372-1384.
- [20] Hong J K, Choi H W, Hwang I S, et al. Function of a novel GDSL-type pepper lipase gene, CaGLIPI, in disease susceptibility and abiotic stress tolerance[J]. Planta, 2008, 227(3): 539-558.
- [21] Kaup M T, Froese C D, Thompson J E. A role for diacylglycerol acyltransferase during leaf senescence[J]. Plant Physiology, 2002, 129(4): 1616-1626.
- [22] Hosoi T, Koguchi Y, Sugikawa E. Identification of a novel human eicosanoid receptor coupled G(i/o)[J]. The Journal of Biological Chemistry, 2002, 277(35): 31459-31465.
- [23] 冯爱芹, 何晓红, 叶绍辉, 等. 猪原癌基因 *JunB* 的电子克隆及比较基因组学分析[J]. 江苏农业科学, 2012, 40(10): 33-36.
- [24] West A B, Lockhart P J, O'Farrell C, et al. Identification of a novel gene linked to parkin via a bi-directional promoter[J]. Journal of Molecular Biology, 2003, 326(1): 11-19.
- [25] 杨胡, 杨军厚. 斑节对虾 1 种 *lectin* 基因的电子克隆及分析[J]. 江苏农业科学, 2012, 40(8): 34-36.
- [26] 张力鹏, 张东, 蔡欣. 牛 *MRF4* 基因的电子克隆与分析[J]. 中国牛业科学, 2007, 33(6): 1-4.
- [27] 李蕊, 孟宪萍, 胡英考, 等. 大豆吡哆醇生物合成蛋白基因 (*PDX*) 的电子克隆和进化分析[J]. 华北农学报, 2007, 22(1): 64-68.
- [28] 巩元勇, 郭书巧, 束红梅, 等. 陆地棉 *GhNIP5. J* 基因的电子克隆及生物信息学分析[J]. 江苏农业学报, 2013, 29(3): 682-684.
- [29] 王璐, 冯彩平. 玉米中一种新的蛋白激酶电子克隆与序列分析[J]. 生物技术, 2007, 17(5): 14-17.
- [30] 王俊生, 范小芳, 李成伟, 等. 小麦 *Ta-UBX1* 基因的电子克隆和生物信息学分析[J]. 江苏农业科学, 2012, 40(6): 29-32.
- [31] 张会敏, 姜民国, 冯友军. 稻瘟菌 I 型烯醇化酶基因全长 cDNA 的电子克隆[J]. 生物信息学, 2006, 4(2): 58-61.