

doi:10.3969/j.issn.1672-5565.2013.03.10

齿肋赤藓热激蛋白60基因的电子克隆和生物信息学分析

蔡明^{1,2}, 高贝^{1,2}, 张道远^{1*}

(1. 中国科学院干旱区生物地理与生物资源重点实验室, 中国科学院新疆生态与地理研究所, 新疆乌鲁木齐 830011; 2. 中国科学院大学, 北京 100049)

摘要:用电子克隆方法获得耐旱苔藓齿肋赤藓的热激蛋白60基因,采用生物信息学方法,对该基因编码蛋白从氨基酸组成、结构保守域、理化性质、信号肽、疏水性/亲水性、亚细胞定位、跨膜结构域、二级结构、功能域、活性位点、及同源性等方面进行了预测和分析。结果表明:齿肋赤藓热激蛋白60基因全长1 841bp,开放阅读框1 581bp,编码526个氨基酸残基;编码蛋白含有GroEL保守域,是chaperon-like superfamily家族;亚细胞定位分析显示,编码蛋白位于内质网中;活性化位点分析表明,编码蛋白存在6类活性位点;同源性分析表明,齿肋赤藓热激蛋白60与小立碗藓预测的HSP60同源性最高,达到92%,与卷柏的HSP60次之,同源性达88.83%。研究结果为该基因的实验克隆奠定基础。

关键词:齿肋赤藓;热激蛋白60;电子克隆;生物信息学分析

中图分类号:Q785 **文献标识码:**A **文章编号:**1672-5565(2013)-03-216-08

Electronic cloning and characterization of HSP60 gene from *Syntrichia caninervis* using bioinformatics tool

CAI Ming^{1,2}, GAO Bei^{1,2}, ZHANG Dao-yuan^{1*}

(1. Key Laboratory of Biogeography and Bioresources in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: A novel heat shock protein 60 from *Syntrichia caninervis* was cloned in silico based on the transcriptome database of *Syntrichia caninervis*. Some characters of amino acids encoded by ScHSP60 gene, include the composition of amino acid sequence, conserved domain, physical and chemical properties, Signal P-NN, hydrophobicity/hydrophilicity, subcellular localization, transmembrane domain, secondary structure of protein, the functional domain, signatures plus the homologies were analyzed by bioinformatics tool. Bioinformatic analysis showed that the full-length HSP60 gene from *S. caninervis* was 1 841bp, and it contained a complete ORF which encoded 526 amino acids. In addition, ScHSP60 contained a GroEL domain and belonged to chaperon-like superfamily. Furthermore, the encoded protein located in endoplasmic reticulum and there were six signatures in the encoded protein. Homology comparison and phylogenetic analysis showed that the amino acid encode by ScHSP60 was highly homologous with the predicted HSP60 in *Physcomitrella patens* and that in *Selaginella tamariscina*, with homologies of 92% and 88.83%, respectively. The above results will provide basis for molecular cloning of HSP60 gene in *S. caninervis*.

Keywords: *Syntrichia Caninervis*; Heat Shock Protein 60 (HSP60); In Silico Cloning; Bioinformatics

热激蛋白是高水平表达蛋白,占正常生长条件下所有类型细胞中蛋白总量的5%~10%。当细胞

处于应激条件时(如高温、干旱、高盐、重金属、缺氧等)热激蛋白表达量明显提高^[1-4]。根据分子量的

收稿日期:2013-01-31;修回日期:2013-04-17。

基金项目:国家自然科学基金资助项目(U1170304)。

作者简介:蔡明,女,新疆乌鲁木齐,硕士研究生;E-mail:kaoyanming@163.com。

* 通讯作者:张道远,女,新疆乌鲁木齐,博士,研究员;E-mail:zhangdy@ms.xjb.ac.cn。

不同,热激蛋白主要划分为 HSP100、HSP90、HSP70、HSP60、HSP40 和 small HSP 6 个家族^[5-6],其中 HSP70 和 HSP60 高度保守。关于 HSP70 的研究相对较多,而 HSP60 的研究却相对较少。HSP60 作为一种分子伴侣,其重要的生物学功能已日益引起人们的重视。许多研究已证实,在离体条件下 HSP60 蛋白可以保护酶不被高温失活^[7];能恢复多种被温度、尿素等变性因素解折叠的蛋白^[8];能协助变性、不可溶的凝聚蛋白重新恢复天然构象^[9]。但有关该基因生物信息学方面的研究未见报道。利用电子克隆技术从苔藓植物齿肋赤藓转录组信息中获得 HSP60 基因,并利用生物信息学软件分析该基因的核苷酸序列组成,预测该基因编码蛋白的结构和功能,为深入研究 HSP60 的基因结构和生物学功能提供一定的理论基础。

1 材料和方法

1.1 齿肋赤藓 HSP60 基因的电子克隆技术

表1 实验所用生物信息学在线分析工具

Table 1 The softwares of bioinformatics analysis in this study

检索内容 (Items)	分析工具 (Softwaers)
保守结构域 (Conserved domain)	http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml
理化性质 (Physical and chemical properties)	http://www.expasy.org/tools/protparam.html
信号肽 (Signal peptide)	http://www.cbs.dtu.dk/services/SignalP-3.0
疏/亲水性 (Hydrophobicity/hydrophilicity)	http://cn.expasy.org/tools/protscale.html
亚细胞定位 (Subcellular localization)	http://psort.hgc.jp
跨膜结构域 (Transmembrane region)	http://www.ch.embnet.org/software/TMPRED_form.html
二级结构 (Secondary structure)	http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html
功能域分析 (Function prediction)	http://www.cbs.dtu.dk/services/ProtFun/
活性位点分析 (Active site prediction)	http://npsa-pbil.ibcp.fr

2 结果与分析

2.1 新基因的识别

以拟南芥 (*Arabidopsis thaliana*) 的 TCP-1/cpn60 chaperonin family 蛋白作为探针,以经过拼接的转录组信息库为基础,进行本地 BLAST 检索,得到一个 1 841bp 的 cDNA 序列。经 ORF 软件分析,该 cDNA 序列包括一个 1 581bp 完整的开放阅读框,编码 526 个氨基酸(图 1)。以起始密码子 ATG 为第一位(标记 +1),在 +4 位置的核苷酸为 G,符合典型的 Kazak 序列规则。同时,在 3'-UTR 区有典型的 poly A 尾,也表明拼接并搜索得到的序列符合一个真

以齿肋赤藓的转录组信息为基础,用 SOAPdenovo 软件进行拼接;用拟南芥 (*Arabidopsis thaliana*) 的 TCP-1/cpn60 chaperonin family 蛋白作为探针,以经过拼接的转录组信息库为基础,进行本地 BLAST 检索;获得的序列利用 NCBI 中相关软件(包括 blastp 和 blastx)进行搜索鉴定;判断正确的开放阅读框(ORF),翻译成编码蛋白,用 PFAM 和 SMART 数据库鉴定,直至得到正确完整的 HSP60 基因序列。

1.2 齿肋赤藓 HSP60 基因及其编码蛋白的生物信息学分析方法

利用在线生物信息学分析软件对齿肋赤藓 HSP60 基因的核苷酸序列组成进行分析,对基因编码的蛋白进行结构保守域、理化性质、信号肽、疏水性/亲水性、亚细胞定位、跨膜结构域、二级结构、功能域、活性位点等生物信息学分析和预测(见表 1)。

核基因全长 cDNA 的基本特征。此外,利用 NCBI 中相关软件(包括 blastp 和 blastx)进行搜索,并用 PFAM 和 SMART 数据库鉴定,该 cDNA 序列是一个齿肋赤藓的热激蛋白 60 基因。

2.2 齿肋赤藓 HSP60 结构保守域分析

使用 NCBI 保守结构域数据库 (conserved Domain Database, CDD),分析齿肋赤藓 HSP60 基因编码蛋白中的保守结构域的结果显示,齿肋赤藓 HSP60 含有 GroEL 保守域,是 chaperon-like superfamily 家族(见图 2),这进一步表明了本研究通过电子克隆所获得的基因是一条完整的基因,属于热激蛋白 60 基因家族。

ATGCAGCGCGGAATTGACAAGCTGGCGGACTCCGTCCGCGTACGCTGGCCCGCGAGGTCGCAATGTGGTGTGGACGACTTCGG
 CGCGCCAGGGTGA TCAACGA TGGCGTACTATCGCCCGGGCGATCGAGTTGCCGAATGCAATGGAGAACGCGGGCGTTCTCTTA
 TCCGCGAGGTTGCCAGCAGGACGAATGACTCTGCGGAGATGGCACGACCACCGCATGTGTGCTGGCGAGGAGCTGATCAAGATG
 GGGCTCTTGAGCGTGACGTCCGGTGCAGAACCCCGTGGCTATCAAGAAAGGAATTGACAAGACCGTGGCTGCTCTCATCGAGGAGCT
 GAAGGAGAGGTCGTTCCCGTGAGGGCCGCGAGACCATCAAGCCGTGGCTCCATTTCTGCTGGAACGACGAGCTGATTGGCA
 CCATGATTGCTGACGCCATCGACAAGTCCGACCCGATGGTGTCTGTCCATCGAGTCCTCTTCTCTTCGAGACCACTGTGGAC
 GTGGAGGAGGATGGAGATTGACAGGGGCTACGTCCTCGCCCAAGTTCGTGACGAAACAACGAGAAGCTGATTGTGGAGTTCTCGAA
 CGCTCGGTGCTGGTGACCGACCAAGAATACGTCATCAAGGAGATCGTCCCGGTGCTGGAGAAACGACCCAGATGAACGTGC
 CACTGCTGATCGTGGCGGAGGACATCTCCGGCGAGGCTCTGGCGACGCTGGTGGTGAACAAGTTGAGGGGCGTGATCCAGGTGGCG
 GCCATCAAGGCGCCCGGTTCCGGGAGCGGCAAGGCTCTTCTGCAGGACATTGCCATCATGACTGGGTCGGAGTTCATCGCCGG
 CGACCTGGGCATGAAGTGGAGAGCACGAGGTGACCAGCTCGGGACTGCGAGGAAGATCACGGTGAAGACGGCACGACCACCA
 TCATCGCTGACGCCGCGAGCAAGGACGAGATCCAGGCCGCATTTCTCAGATCAAGAAAGGAGTTGCAAGAGACTGACTCCGTGTAC
 GACACGAGAAGTTGTCGGAGCGCATTGCGAAGCTGTACAGGAGGAGTGGCGGTGATCAAGTGGAGCCGCGACAGAGACTGAGCT
 CGAGGACCGTAAGCTGAGAATTGAGGACGCGAAGAACGCGACGTTCCGCGGATCGAGGAGGCGATCGTCCCGGTGGCGGTGCGG
 CGTTGGTGCACCTGTCGGCGTGGTCCAGCCATCAAGAACGATCAGGACCCGAGGAGAAGCTGGGAGCCGACATCGTGACG
 AGGGCCCTGAGCTCTCCGGCGAGTCTGATCGCCAACACGCGCGTGGAGGGCAGGTAGTGGTGGAGAAGATCTGGACAGCGA
 GTGGCAGATGGGTACAACCGGATGACGGACGTCAGGACCTGCTGGTGGCGGCGTGATCGACCCGCGAAGGTGACGCGGT
 GCGCGCTGCAGAACGACGCTCCGTGGCCGGCATGGTGTGACGACGAGGCCATCGTGTGCGAGAAGCCGAGAGAAGTCCGGCT
 GTACCGATGCAGCCCAGGCCATGACCATGTAAATTTTTGACGGTGGTAGGCGGTAGAGGTAAGCACACTCATTTCGACACTCATT
 TACATTCATAGTACTTGTAGTGTAGGCTAGGTGACAAAGACCCAGATTTTCATTTGTAGCGAGAGTCCGTGCTTATGAGGAAGT
 TTATTTCAAGCTGTATTAGGTTGAATTCTTGAACCATAATTTGTTTCTCTTGAAGTGTCCAGATAAACATGCAATACCATGAAG
 GTCTGTTTTCGAAAAAAAAAAAAAATGGAA

MQAGIDKLDVSVVTLGPRGRNVVLDLDFGAPRVIINDGVTIARAIELPNAMENAGASLI REVASRTNDSAGDGTTACVLAREL IKM
 GLLSVTSGANPVAIKKGIDKTVAALIEELKERSVPVEGRET IKAVASISAGNDELIGTMIADAIDKVGPDGVLSESSSFETTVD
 VEEGMEIDRGYVSPQFVTNNEKLI VEFVSNARVLVTDQKITSIKEIVPVLEKTTQMNVPVLLI VAEDI SGEALATLVVNKLRGVIQVA
 AIKAPGFGERRKALLQDIAIMTGSEFIAGDLGMKVESTEVDQLGTARKITVRSQTTI IADAASKDEIQARISQIKKELQETDSVY
 DTEKLSERIAKLSGGVAIVKVGAAETELEDRLRIEDAKNATFAAIEEGI VPGGGAALVHLSALVPAIKNTIQDPEEKLGDIVQ
 RALSSPASP IANNAGVEGEVVVEKILDSEWQMGYNAMTDTYEDLLVAGVIDPAKVTRCALQNAASVAGMVLTTQAI VCEKPEKKS
 VPMQPQGMTM

图 1 齿肋赤藓 HSP60 基因的开放阅读框及其编码蛋白序列

Fig. 1 The sequences of ScHSP60 open reading frame and its protein

注:横线以上为基因 ORF 序列,横线以下为编码蛋白序列

Notes: The ORF is above the line, the protein is below the line

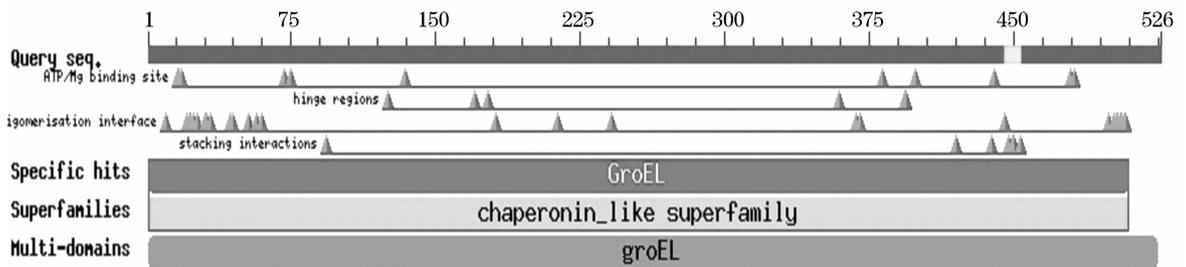


图 2 齿肋赤藓 HSP60 保守结构域分析

Fig. 2 The analysis of conserved domain prediction of *S. caninervis* HSP60

2.3 齿肋赤藓 HSP60 基因编码氨基酸的一级结构预测

经预测,齿肋赤藓 HSP60 基因编码氨基酸的一级结构如表 2。

表 2 齿肋赤藓 HSP60 的一级结构预测分析

Table 2 Primary structure analysis of *S. caninervis* HSP60

一级结构特性	预测结果
编码氨基酸数/个	526
等电点 (PI)	4.73
分子量 (MW)/Da	55469.5
负电荷残 (Asp + Glu)	76
正电荷残 (Arg + Lys)	54
分子式	C2413H4030N662O788S17
不稳定系数 (II)	29.52
平均疏水性 (GRAVY)	0.051
脂肪系数 (AI)	104.77

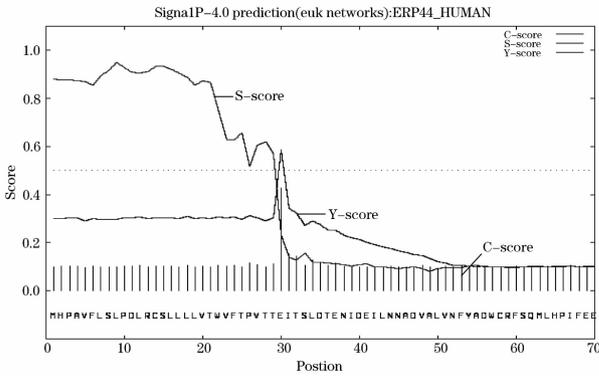


图 3 齿肋赤藓 HSP60 信号肽预测

Fig.3 The Signal P-NN prediction for HSP60 of *Syntrophomonas caninervis*

注:C score:原始剪切位点的分值;S score:信号肽的分值;Y score:综合剪切位点的分值。

Notes:C score:original splice site scores;S score:signal peptide scores;Yscore:Integrated splice site scores.

2.4 齿肋赤藓 HSP60 信号肽预测和分析

信号肽是位于蛋白质一端的一部分氨基酸序

列,可以引导目的蛋白质穿越细胞器的膜在细胞的正确位置定位。采用 SignalP3.0Server 预测齿肋赤藓热激蛋白 60 的信号肽,结果如图 3 和表 3。根据图 3 和表 3 的预测结果,齿肋赤藓 HSP60 基因所编码的蛋白在 30 和 31 位置上有一个潜在的信号肽断裂位点。并且 C 最大值与 Y 最高峰以及 S 最高值处于同一区域,因此,齿肋赤藓 HSP60 含有信号肽,该蛋白在细胞质合成后,还需要转运至正确的位置(内质网)以行使其分子伴侣等生物学功能。

表 3 齿肋赤藓热激蛋白 60 信号肽预测

Table 3 The Signal P-NN prediction for HSP60 from *S. caninervis*

指标 (Measure)	位点 (Site)	分值 (Score)	有无信号肽
max. C	30	0.427	有
max. Y	30	0.586	有
max. S	9	0.95	有

2.5 齿肋赤藓 HSP60 疏水性/亲水性分析

用 ProtScale 对齿肋赤藓 HSP60 基因编码的氨基酸序列的疏水性/亲水性进行预测(见图 4)。由图 4 可以看出,齿肋赤藓 HSP60 多肽链的第 373 位具有最低的分值,为 -2.222,亲水性最强;第 406 位具有最高的分值,为 2.156,疏水性最强。整条多肽链表现为亲水性,没有明显的疏水区。因此,推测 ScHSP60 蛋白是一种可溶性蛋白。

2.6 齿肋赤藓 HSP60 的亚细胞定位

应用 PSORT 工具预测齿肋赤藓热激蛋白的亚细胞定位,分析结果如表。根据前人研究热激蛋白 60 存在于植物细胞的某些细胞器中如内质网、线粒体、叶绿体、高尔基体等^[9]。因此,根据表 4 中齿肋赤藓 HSP60 的亚细胞定位的不同概率可知,齿肋赤藓热激蛋白定位于内质网的可能性最大,概率为 0.650。

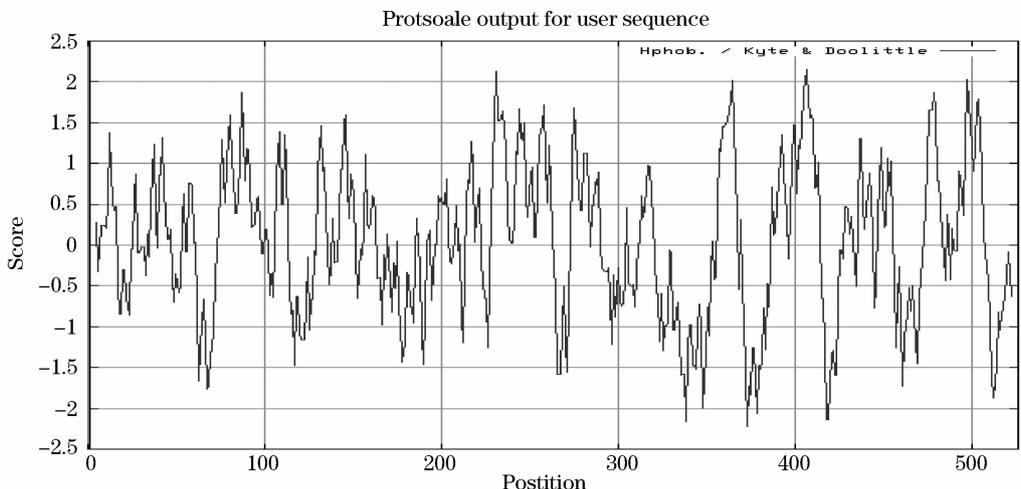


图 4 齿肋赤藓 HSP60 基因氨基酸疏水性/亲水性进行预测

Fig.4 Predicted hydrophobicity/hydrophilicity of the sequence of HSP60 deduced amino acid from *S. caninervis*

表 4 齿肋赤藓 HSP60 的亚细胞定位

Table 4 Subcellular location of *S. caninervis* HSP60

亚细胞定位	概率
内质网膜 (endoplasmic reticulum (membrane))	0.550
内质网腔 (endoplasmic reticulum (lumen))	0.100
细胞质 (cytoplasmic ground substance)	0.100
高尔基体 (Golgi body)	0.100

2.7 齿肋赤藓 HSP60 的跨膜结构域预测

利用 ExPasy 软件中的 TMHMM Server v. 2.0 工具预测齿肋赤藓 HSP60 的跨膜螺旋区,结果见图 5。结果显示:齿肋赤藓 HSP60 不是膜蛋白,无跨膜螺旋区。

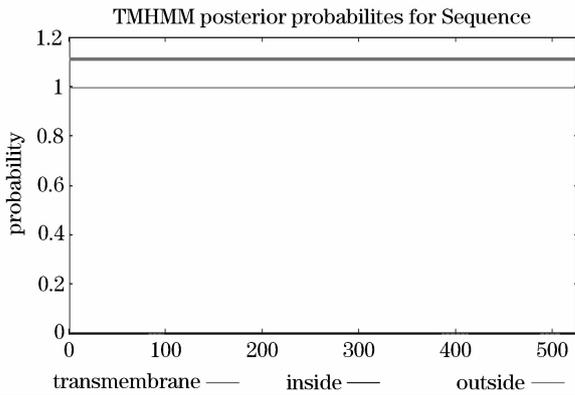


图 5 T 齿肋赤藓 HSP60 跨膜结构域预测

Fig. 5 Predicted transmembrane domain of *S. caninervis* HSP60

2.8 齿肋赤藓 HSP60 二级结构分析及预测

对齿肋赤藓 HSP60 基因编码蛋白的二级结构进行分析发现,其主要由 4 种形式组成,其中, α 螺旋 (Alpha helix) 占 54.75%, β 折叠 (Beta turn) 占 8.56%,延伸链 (Extended strand) 占 12.17%,无规则卷曲 (Random coil) 占 24.52%。二级结构中以 α 螺旋所占比例较高,且 N-末端以 α 螺旋形式存在,C 末端以无规则卷曲形式存在(见图 6 和表 5)。

2.9 齿肋赤藓 HSP60 的功能预测

用 CBS 的 Profun 软件预测齿肋赤藓 HSP60 基因编码蛋白的功能,其功能见表 6。从表 6 可知,该蛋白具有电压门控离子通道、阳离子通道、离子通道、转录功能的可能性分别为 12.682、3.174、2.965 和 1.711,这符合分子伴侣折叠转运新合成多肽以使其形成正确的构象发挥生物学功能的功能。

表 5 齿肋赤藓 HSP60 的二级结构预测分析

Table 5 The analysis of secondary structure prediction of HSP60 from *S. caninervis*

二级结构类型	氨基酸残基数目(个)	百分比(%)
α 螺旋	288	54.75%
β 折叠 Beta turn	45	8.56%
延伸链	64	12.17%
无规则卷曲	129	24.52%

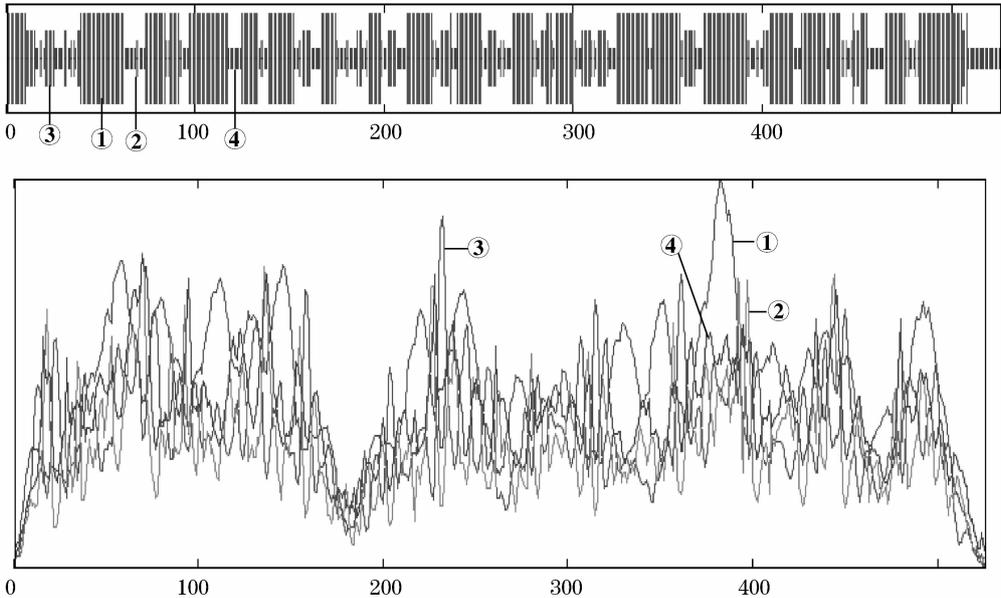


图 6 预测齿肋赤藓 HSP60 的二级结构

Fig. 6 Predicted secondary structure of HSP60 from *S. caninervis*

注:① α 螺旋; ② β 折叠; ③延伸链; ④无规卷曲。

Notes: ①Alpha helix; ②Beta turn; ③Extend strand; ④Random coil.

表6 齿肋赤藓 HSP60 功能预测

Table 6 The function prediction of *S. caninervis* HSP60

Gene Ontology category	Prob	Odds
Signal_transducer(信号传导)	0.205	0.958
Receptor(受体)	0.007	0.041
Hormone(激素)	0.001	0.154
Structural_protein(结构蛋白)	0.003	0.107
Transporter(运载体)	0.025	0.229
Ion_channel(离子通道)	0.169	2.965
Voltage-gated_ion_channel(电压门控离子通道)	= >0.279	12.682
Cation_channel(阳离子通道)	0.146	3.174
Transcription(转录)	0.219	1.711
Transcription_regulation(转录调控)	0.111	0.888
Stress_response(胁迫应答)	0.073	0.83
Immune_response(免疫应答)	0.011	0.129
Growth_factor(生长因子)	0.005	0.357
Metal_ion_transport(金属离子转移)	0.018	0.039

2.10 齿肋赤藓 HSP60 的活性位点分析

利用 NPS(Network Protein Sequence Analysis) <http://npsa-pbil.ibcp.fr> 提供的 Prosite Scan 对齿肋赤藓 HSP60 进行蛋白的活性位点分析,结果发现,齿肋赤藓 HSP60 有 6 类活性位点,分别是:N-糖基化位点,cAMP-和 cGMP-依赖的蛋白激酶磷酸化位点,蛋白激酶 C 磷酸化位点,酪蛋白激酶 II 磷酸化位点,N-豆蔻酰化位点和分子伴侣 60 信号位点,这可能与该基因发挥的分子伴侣、细胞保护等生物学功能有关(见表 7)。

表7 齿肋赤藓 HSP60 在 NPS 中的活性位点预测分析

Table 7 Scanning of *S. caninervis* HSP60 for the site/signatures against proscan database in NPS

名称 Format name	序列号 Prosite access number	膜体类型 Motif model	氨基酸序列 Amino acid sequence	概率值 Randomized probability
N-糖基化位点 N-glycosylation site	PS00001	N{P}{[ST]}{P}	66-69 NDSA 385-388 NATF	5.14E-03
cAMP 和 cGMP 依赖性蛋白激酶磷酸化位点 cAMP and cGMP dependent protein kinase phosphorylation site	PS00004	[RK](2)x[ST]	127-129 TIK 213-215 SIK 303-305 TAR 308-310 TVR 346-348 TEK 350-352 SER	1.57E-03
蛋白激酶 C 磷酸化位点 Protein kinase C phosphorylation site	PS00005	[ST]x[RK]	68-71 SAGD 165-168 SSFE 169-172 TTVD 190-193 TNNE 213-216 SIKE 280-283 TGSE 296-299 TEVD 322-325 SKDE 342-345 SVYD 369-372 TETE 371-374 TELE 416-419 TIQD 470-473 TYED	1.42E-02
酪蛋白激酶 II 磷酸化位点 Casein kinase II phosphorylation site	PS00006	[ST]x(2)[DE]	37-42 GVTIAR 72-77 GTTTAC 103-108 GIDKTV 143-148 GTMIAD 394-399 GIVPGG 398-403 GGGAAL 498-503 GMVLTT	1.48E-02
N-豆蔻酰化位点 N-myristoylation site	PS00008	G{EDRKHPFYW}x (2)[STAGCN]{P}	37-42 GVTIAR 72-77 GTTTAC 103-108 GIDKTV 143-148 GTMIAD 394-399 GIVPGG 398-403 GGGAAL 498-503 GMVLTT	1.40E-02
分子伴侣 60 信号位点 Chaperonins cpn60 signature	PS00296	A[AS]{L}- [DEQ]E{A}{Q} {R}xGG[GA]	389-400 AAIEEGIVPGGG	6.18E-08

2.11 齿肋赤藓 HSP60 同源性分析

将齿肋赤藓 HSP60 蛋白与绿藻、小立碗藓、卷柏、苜蓿、豌豆、大豆、拟南芥、葡萄等植物 HSP60 蛋白进行同源比对发现,它们在该蛋白水平上具有较高的一致性,其中,齿肋赤藓 HSP60 与小立碗藓预测的 HSP60 的同源性最高,达到 92%;与卷柏的 HSP60 的同源性次之,达到 88.83%;与拟南芥的同源性为 83.65%。借助 MEGA5.0 软件,用 NJ (Neighbor-Joining) 法构建进化树(见图 7),发现该进化树分为三大支,绿藻 HSP60 单独为一支,位于

基部;齿肋赤藓 HSP60 与小立碗藓预测的 HSP60 和卷柏 HSP60 聚为一支;其它物种(红车轴草、苜蓿、豌豆、大豆、葡萄、蓖麻、油菜、拟南芥、小麦、水稻、高粱和玉米)的 HSP60 聚为一支。由此说明,得到的基因是 HSP60 基因,并揭示了齿肋赤藓 HSP60 蛋白的结构位于简单真核生物(绿藻)HSP60 蛋白结构向高等植物(拟南芥等)HSP60 蛋白结构的过渡阶段,这符合齿肋赤藓藓类植物在整个真核植物中的进化关系。

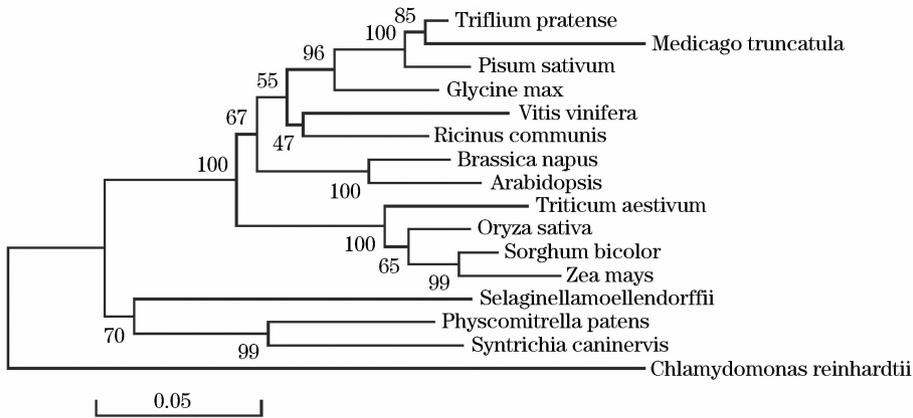


图 7 齿肋赤藓 HSP60 蛋白序列进化树

Fig. 7 Phylogenetic tree of *S. caninervis* HSP60 based on protein

3 讨论

电子克隆技术 (in silico cloning) 是基于生物学数据库中 EST (expressed sequence tag, 表达序列标签) 数据库、核酸序列数据库和蛋白质数据库等,利用生物信息学工具和计算机技术,采用序列同源比对和聚类分析、重叠区域组装和拼接等方法不断延长 EST 序列,从而快速获得功能基因的方法^[10-11]。与传统 RACE 技术及全长 cDNA 文库筛选的克隆基因方法相比,该技术具有效率高、成本低、针对性强、技术要求低等优点^[10-11]。目前,利用电子克隆获得植物新基因的报道也日益增多^[12-15]。

本研究以转录组信息为基础,运用电子克隆技术,识别获得了一个全长的齿肋赤藓 HSP60 基因,并利用生物信息学分析技术对该基因的结构特征、编码蛋白的理化性质、亚细胞定位、生物学功能等进行了预测。研究表明,齿肋赤藓 HSP60 属于热激蛋白 60 家族,该蛋白可能具有细胞外被膜、转运结合、氨基酸生物合成、生物合成的辅因子、电压门控离子通道、阳离子通道、离子通道、转录、转录调控、信号转导、胁迫应答等功能。本研究结果为深入研究该基因的表达信息、生物学功能,以及全面了解该

基因家族的信息奠定了基础,并为将来分子育种提供了理论基础。

4 结论

(1) 基于齿肋赤藓的转录组信息,用电子克隆技术获得一个全长为 1 841bp 的热激蛋白 60 基因,包含 1 581bp 的开放阅读框,编码 526 个氨基酸残基;

(2) 使用 NCBI 保守结构域数据库分析显示,编码蛋白含有 GroEL 保守域,是 chaperon-like superfamily 家族;

(3) 亚细胞定位分析和信号肽预测显示,该蛋白定位于内质网,并存在相关的信号肽;

(4) 用 ProtScale 对该基因编码的氨基酸序列进行疏水性/亲水性预测,结果显示该蛋白具有亲水性;

(5) 利用 ExPASy 的 TMHMM Server v. 2.0 工具预测了该蛋白的跨膜螺旋区,结果显示该蛋白无跨膜螺旋区;

(6) 对该蛋白二级结构的预测发现, α 螺旋所占比例较高,且 N-末端以 α 螺旋形式存在,C 末端以无规则卷曲形式存在;

(7)活化位点分析表明,编码蛋白存在6类活性位点;

(8)同源性分析表明,齿肋赤藓热激蛋白60与小立碗藓预测的HSP60同源性最高,达到92%,与卷柏的HSP60次之,同源性达88.83%。

参考文献(References)

- [1] Wunderlich M., Werr W., Schoffl F. Generation of dominant-negative effects on the heat shock response in *Arabidopsis thaliana* by transgenic expression of a chimaeric HSF1 protein fusion construct [J]. *Plant Journal*, 2003 (35): 442-451.
- [2] Waagner D., Bayley M., Marien J., Holmstrup M., Ellers J., Roelofs D. Ecological and molecular consequences of prolonged drought and subsequent rehydration in *Folsomia candida* (Collembola) [J]. *Journal Of Insect Physiology*, 2012 (58): 130-137.
- [3] Wang X. Q., Yang P. F., Gao Q., Liu X. L., Kuang T. Y., Shen S. H., He Y. K. Proteomic analysis of the response to high-salinity stress in *Physcomitrella patens* [J]. *Planta*, 2008 (228): 167-177.
- [4] Rau S., Miersch J., Neumann D., Weber E., Krauss G. J. Biochemical responses of the aquatic moss *Fontinalis antipyretica* to Cd, Cu, Pb and Zn determined by chlorophyll fluorescence and protein levels [J]. *Environmental And Experimental Botany*, 2007 (59): 299-306.
- [5] Georgopoulos C., Welch W. J. Role of the Major Heat-Shock Proteins as Molecular Chaperones [J]. *Annual Review of Cell Biology*, 1993 (9): 601-634.
- [6] Lindquist S., Craig E. A. The Heat-Shock Proteins [J]. *Annual Review Of Genetics*, 1988 (22): 631-677.
- [7] Welch W. J. The role of heat-shock proteins as molecular chaperones [J]. *Current Opinion In Cell Biology*, 1991 (3): 1033-1038.
- [8] Hendrick J. P., Hartl F. U. Molecular Chaperone Functions of Heat-Shock Proteins [J]. *Annual Review Of Biochemistry*, 1993 (62): 349-384.
- [9] 龚兴国,于红. 热休克蛋白60的研究进展 [J]. *中国病理生理杂志*, 2004 (20): 2151-2154.
- [10] 王冬冬,朱延明,李勇,李杰,柏锡. 电子克隆技术及其在植物基因工程中的应用 [J]. *东北农业大学学报*, 2006 (37): 403-408.
- [11] 胡骥,萧浪涛. 生物信息学在新基因全长cDNA电子克隆中的应用 [J]. *生物技术通报*, 2007 (4): 93-96.
- [12] 王芳平. 大肠杆菌与酵母基因组中密码对使用的比较 [J]. *内蒙古大学学报:自然科学版*, 2006 (37): 34-39.
- [13] 王俊生,范小芳,李成伟,谭光轩. 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): 34-39.
- [14] 孟鹏,李根英,陈国强,王振东. 向日葵抗坏血酸过氧化物酶的电子克隆和生物信息学分析 [J]. *山东农业科学*, 2012 (44): 1-6.
- [15] 畅丽萍,魏琦超,周岩. 马铃薯促分裂原活化蛋白激酶激酶基因的电子克隆与生物信息学分析 [J]. *广东农业科学*, 2012, (39): 159-162.