

DOI:10.13350/j.cjpb.221206

• 论著 •

# 细粒棘球绦虫 TAK1 基因 ORF 的克隆及其蛋白生物信息学分析\*

刘丽英<sup>1</sup>, 李军<sup>2,3</sup>, 张文宝<sup>2,3</sup>, 吴军<sup>1\*\*</sup>, 杨梅<sup>2,4\*\*</sup>

(1. 新疆医科大学公共卫生学院劳动卫生与环境卫生教研室, 新疆乌鲁木齐 830011; 2. 新疆医科大学基础医学院生物化学与分子生物学教研室; 3. 新疆医科大学第一附属医院临床医学研究院省部共建中亚高发病成因与防治国家重点实验室; 4. 新疆医科大学新疆地方病分子生物学重点实验室)

**【摘要】** 目的 克隆细粒棘球绦虫(*Eg*)TAK1 基因的开放阅读框(ORF)全长, 预测其编码的蛋白 EgTAK1 的结构特点和生物学功能。方法 从新疆株细粒棘球绦虫中扩增 *Eg*TAK1-ORF 全长, 克隆至载体 pMD20-T 上并测序; 从 NCBI 蛋白质数据库中下载 EgTAK1 的氨基酸序列, 利用生物信息学分析软件分别对 EgTAK1 的理化性质、磷酸化位点、O-糖基化位点、N-糖基化位点、亲/疏水性、保守结构域、跨膜结构域、信号肽序列进行预测; 采用 MEGA 7.0 软件对不同物种来源的 TAK1 蛋白进行多序列比对并构建系统进化树。采用实时荧光定量 PCR 检测细粒棘球绦虫在不同发育阶段 TAK1 基因 mRNA 的相对表达量。结果 *Eg*TAK1 基因 ORF 全长 1 116 bp, 编码 371 个氨基酸, 理论等电点为 5.97, 不稳定指数为 49.59, 属于不稳定蛋白; 脂肪系数为 87.22, 总平均亲水系数为 -0.291, 为亲水性蛋白; EgTAK1 蛋白含有 64 个磷酸化位点, 3 个 O-糖基化潜在位点, 4 个 N-糖基化位点; EgTAK1 蛋白属于蛋白激酶 C 类似(PKC-like)家族; 无跨膜结构和信号肽, 亚细胞可能定位于细胞膜、细胞质及细胞核内。该蛋白含有 7 个优势 B 细胞抗原表位, 具有良好的免疫原性。二级结构包含 34.77% 的  $\alpha$ -螺旋, 5.12% 的  $\beta$ -转角, 40.97% 的无规卷曲, 19.14% 的延伸链。TAK1 蛋白同源序列比对中与多房棘球绦虫的 TAK1 同源性最高为 83%。实时荧光定量 PCR 显示, 细粒棘球绦虫在不同发育阶段 EgTAK1 表达有差异, 成虫阶段相对表达量最高。结论 基因 EgTAK1 可能是细粒棘球绦虫发育分化的重要调控基因, 预测该基因编码的蛋白含有 B 细胞抗原表位, 有可能成为包虫病免疫学预防及药物研制的靶标抗原。

**【关键词】** 细粒棘球绦虫; TAK1; 生物信息学分析; PCR

**【中图分类号】** R383.33

**【文献标识码】** A

**【文章编号】** 1673-5234(2022)12-1392-07

[Journal of Pathogen Biology. 2022 Dec.; 17(12):1392-1398.]

## Cloning of the ORF of TAK1 gene of *Echinococcus granulosus* and bioinformatic analysis of the protein

LIU Li-ying<sup>1</sup>, LI Jun<sup>2,3</sup>, ZHANG Wen-bao<sup>2,3</sup>, WU Jun<sup>1</sup>, YANG Mei<sup>2,4</sup> (1. Labour and Environmental Hygiene Department, College of Public Health, Xinjiang Medical University, Urumqi, Xinjiang 830011, China; 2. Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Xinjiang Medical University; 3. State Key Laboratory of Pathogenesis, Prevention and Treatment of High Incidence Diseases in Central Asia, Clinical Medicine Institute of First Affiliated Hospital, Xinjiang Medical University; 4. Xinjiang Key Laboratory of Molecular Biology for Endemic Diseases, Xinjiang Medical University) \*\*

**【Abstract】** **Objective** To clone the ORF of gene TAK1 of *Echinococcus granulosus*, and to predict the structural characteristics and biological functions of EgTAK1 protein. **Methods** The full length of EgTAK1-ORF was amplified from *E. granulosus* of Xinjiang strain, cloned into pMD20-T vector and sequenced. The amino acid sequence of EgTAK1 was downloaded from NCBI protein database. The physicochemical properties, phosphorylation sites, O-glycosylation sites, N-glycosylation sites, hydrophilicity / hydrophobicity, conserved domain, transmembrane domain and signal peptide sequence of EgTAK1 were predicted by bioinformatics analysis software. By MEGA 7.0 software, multi-sequence alignment of TAK1 proteins from different species was performed and phylogenetic tree was constructed. The relative expression of TAK1 mRNA in different developmental stages of *E. granulosus* was detected by real-time quantitative PCR. **Results** The ORF of EgTAK1 gene was 1 116 bp, encoding 371 amino acids. The theoretical isoelectric point was 5.97 and the instability index was 49.59, which belonged to unstable protein. The fat coefficient was 87.22, and the total average hydrophilic coefficient was -0.291, which was hydrophilic protein; EgTAK1 protein contains 64 phosphorylation sites, 3 O-gly-

\* **【基金项目】** 新疆维吾尔自治区自然科学基金项目(No. 2017D01C197)。

\*\* **【通讯作者】** 吴军, E-mail: wuj1997@sohu.com; 杨梅, E-mail: 768144163@qq.com

**【作者简介】** 刘丽英(1994-), 女, 山西人, 在读硕士研究生, 主要研究方向: 公共卫生-包虫病防控的分子生物学研究。E-mail: 405720710@qq.com

cosylation potential sites and 4 N-glycosylation sites. EgTAK1 protein belongs to PKC-like family ;without transmembrane structure and signal peptide, subcellular localization may be located in cell membrane, cytoplasm and nucleus. The protein contains seven dominant B cell epitopes and has good immunogenicity. The secondary structure contains 34.77%  $\alpha$ -helix, 5.12%  $\beta$ -turn, 40.97% random coil and 19.14% extension chain. In the homology alignment of TAK1 protein, the highest homology with TAK1 of *E. multilocularis* was 83%. qPCR showed that the expression of EgTAK1 was different at different developmental stages, and the relative expression was the highest at adult stage. **Conclusion** The gene EgTAK1 may be an important regulatory gene for the development and differentiation of *E. granulosus*. It is predicted that the protein encoded by this gene contains B cell antigen epitope, which may become the target antigen for immunological prevention and drug development of hydatid disease.

**【Key words】** *Echinococcus granulosus*; TAK1; Bioinformatic analysis; PCR

细粒棘球蚴病又称包虫病,是由细粒棘球绦虫(*Echinococcus granulosus*, *Eg*)引起的重要的人畜共患寄生虫病,呈世界性分布<sup>[1]</sup>。我国主要流行于畜牧业发达的西北地区,如新疆、西藏、青海等地,发病率较高,给人民健康以及农牧业发展造成巨大损失<sup>[2]</sup>。目前包虫病主要治疗手段有药物治疗和手术治疗,但存在药物毒副作用和术后复发的风险<sup>[3-4]</sup>。因此包虫病的免疫学预防及药物靶标的筛选研究备受关注,亟待加强对包虫病的防治及有效疫苗的研发。

细粒棘球绦虫具有双向发育的特征,胆汁酸盐是决定其向成虫发育的关键因素,杨梅等<sup>[5]</sup>通过对有/无犬胆汁(主要成分是胆汁酸盐)培养的原头蚴转录组测序及差异表达基因和 KEGG 富集通路分析发现核因子  $\kappa$ B(nuclear factor- $\kappa$ B, NF- $\kappa$ B)信号通路与 Toll 和 Imd 信号通路富集程度位居前二, *EgTAK1* 在这两条通路中而且表达明显上调。此基因表达的蛋白名称是转化生长因子  $\beta$ -激活激酶 1(TAK1), 或称丝裂原活化蛋白激酶激酶激酶 7(Mitogen activated protein kinase kinase kinase 7, MAPKKK7 或 MAP3K7), 蛋白序列号是 CDS024499579.1。TAK1 是一种关键的丝氨酸/苏氨酸蛋白激酶, 可被转化生长因子  $\beta$ (transforming growth factor- $\beta$ , TGF- $\beta$ ), 肿瘤坏死因子(tumour necrosis factor, TNF), 白细胞介素-1(interleukin-1, IL-1)以及 B 细胞受体抗原等激活<sup>[6]</sup>。TAK1 是 MAP3K 家族中唯一需要与其结合蛋白(transforming growth factor- $\beta$  activated kinase 1 binding protein, TAB)结合才能发挥激活作用的激酶<sup>[7]</sup>。当 TAK1 激活后, 可以被磷酸化、泛素化或是糖基化, 可诱导核因子  $\kappa$ B(nuclear factor- $\kappa$ B, NF- $\kappa$ B) 和丝裂原活化蛋白激酶(mitogen activated protein kinase, MAPK)等信号转导通路, 参与细胞的生长、存活、代谢等生理反应以及免疫应答反应<sup>[8]</sup>。细粒棘球绦虫等多种寄生虫疾病的发生发展都与 MAPK 信号通路相关, 通过 MAPK 信号转导通路, 促进棘球蚴细胞间信息交流, 从而影响其生长和发育<sup>[9]</sup>。TAK1 在果蝇和秀丽线虫发育中起重要作用<sup>[10]</sup>, Nemo 样激酶(nemo-

like kinase, NLK)是一种非典型 MAPK 激酶, TAK1-NLK 级联反应在秀丽隐杆线虫胚胎的内胚层中被发现, 可参与 Wnt 信号转导途径, Wnt 信号转导途径调控胚胎发育的各个方面, 与细胞增殖凋亡和胞外基质的合成有紧密关联<sup>[11]</sup>。但目前关于细粒棘球绦虫 TAK1(*EgTAK1*)蛋白结构和功能的研究鲜见报道。

本研究拟克隆 *EgTAK1* 基因 ORF 全长并测序鉴定; 然后利用生物信息学方法预测其编码蛋白的理化特性、结构特性和生物学功能; 采用实时荧光定量 PCR 检测细粒棘球绦虫在不同发育阶段 TAK1 基因的相对表达量差异; 为进一步研究 *EgTAK1* 在细粒棘球绦虫生长发育中的作用机制奠定基础。

## 材料与方 法

### 1 材 料

从新疆乌鲁木齐市某屠宰场获取自然感染 *Eg* 的绵羊肝脏, 从中获得细粒棘球绦虫的原头蚴, *E. g* 成虫是由省部共建中亚高发病成因与防治国家重点实验室提供。总 RNA 提取试剂盒与反转录试剂盒均购买于美国 Thermo Scientific 生物技术公司; PCR 扩增试剂盒、DNA 连接酶及其他生物学试剂均购自大连宝信生物公司。

### 2 方 法

**2.1 总 RNA 提取与 cDNA 逆转录** 使用 Trizol 法提取 *Eg* 生发层、原头蚴(PSC)和成虫的总 RNA, 核酸定量仪测定浓度及纯度。使用反转录试剂盒合成 cDNA, 反应体系(20  $\mu$ L): gDNA Eraser 1  $\mu$ L, 5 $\times$ gDNA Eraser Buffer 2  $\mu$ L, Total RNA 1  $\mu$ g, RNase Free dH<sub>2</sub>O 至总体积为 10  $\mu$ L, 室温 3~5 min; PrimeScript RT Enzyme Mix I 1  $\mu$ L, RT Primer Mix 1  $\mu$ L, 5 $\times$  PrimeScript Buffer 4  $\mu$ L, RNase Free dH<sub>2</sub>O 4  $\mu$ L, 2000 r/min(离心半径 3 cm)离心 2 min; 反应条件: 37  $^{\circ}$ C 15 min, 85  $^{\circ}$ C 5 s。于 -20  $^{\circ}$ C 保存备用。

**2.2 TAK1 基因的克隆与测序** 运用 NCBI 登录 Gene ID: 36346045 下载 *EgTAK1* 基因 ORF 序列, 应用 Primer 5.0 设计扩增该基因 ORF 全长序列的特异

性引物,由上海生物工程有限公司合成,Primer-F:5'-ATGCTTCGCGATAAGGGCA-3', Primer-R: 5'-TTAAATGTATTGTAGAATAATATTTGCGAG-3'。用反转录得到的PSC cDNA为模板,进行目的基因PCR扩增。PCR扩增体系(25 μL):2×PCR masterMix 12.5 μL,RNase Free dH<sub>2</sub>O 8.5 μL,模板cDNA 2.0 μL,上、下游引物各1.0 μL。反应条件:95℃预变性3 min,95℃变性30 s,56℃退火45 s,72℃延伸1 min,共34个循环;72℃延伸7 min。PCR产物经1.0%琼脂糖凝胶电泳鉴定后进行胶回收,与pMD20-T载体连接,转化入感受态DH5α大肠埃希菌,通过菌落PCR筛选阳性克隆并进行液体培养,从阳性克隆菌液中提取重组质粒,由上海生物工程有限公司测序。

**2.3 生物信息学分析** 登录NCBI,在Protein数据库(<https://www.ncbi.nlm.nih.gov/protein>)中下载EgTAK1蛋白(XP\_024345997.1)的氨基酸序列,运用在线软件ProtParam(<http://web.expasy.org/ProtParam/>)预测EgTAK1蛋白的理化性质;运用在线软件NetPhos 3.1(<http://www.cbs.dtu.dk/services/NetPhos-3.1/>)预测EgTAK1蛋白的磷酸化位点,运用在线软件NetOGlyc 4.0(<http://www.cbs.dtu.dk/services/NetOGlyc/>)和NetNGlyc 1.0(<http://www.cbs.dtu.dk/services/NetNGlyc/>)预测EgTAK1蛋白的O-糖基化位点及N-糖基化位点;运用DNASar软件中的Protean程序预测EgTAK1蛋白的抗原性,包括亲水性指数(hydrophilicity)、抗原性指数(antigenic index)、表面可及性指数(surface probability);运用在线软件ProtScale(<http://web.expasy.org/protscale/>)分析EgTAK1蛋白的亲、疏水性;运用SignalP 5.0程序(<http://www.cbs.dtu.dk/services/SignalP/index.php>)预测EgTAK1蛋白信号肽;运用Euk-mPLoc 2.0(<http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/>)对该蛋白亚细胞定位进行预测;运用在线网站CDD([ncbi.nlm.nih.gov/cdd](http://ncbi.nlm.nih.gov/cdd))和ORF finder([ncbi.nlm.nih.gov/orffinder/](http://ncbi.nlm.nih.gov/orffinder/))预测EgTAK1蛋白的保守结构域和开放阅读框;运用在线软件TMHMM 2.0(<http://www.cbs.dtu.dk/services/TMHMM-2.0/>)预测EgTAK1蛋白的跨膜结构域;运用SOPMA([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html))预测EgTAK1蛋白二级结构;运用在线软件SWISS MODEL(<http://www.swiss-model.expasy.org/>)构建EgTAK1蛋白同源模型,分析其三维结构;通过NCBI网站BLAST程序(Basic Local Alignment Search Tool,<http://www.ncbi.nlm.nih.gov/balst>)

分析蛋白序列同源性,运用MEGA7.0软件(<http://www.megasoftware.net/>)构建蛋白系统进化树。

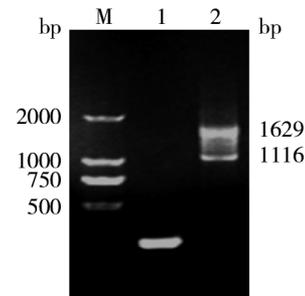
**2.4 TAK1mRNA表达水平的差异性分析** 将提取的细粒棘球绦虫不同发育阶段样本总RNA反转录为cDNA,以此为模板,以Eif为内参基因进行荧光定量PCR(qPCR)。反应体系(20 μL):SYBR Green PCR Master Mix 10 μL,ddH<sub>2</sub>O 5.8 μL,cDNA 2 μL,上、下游引物各1 μL,QuantiNova Yellow Template Dilution Buffer 0.2 μL。每个样品重复3次。反应条件:95℃预变性2 min,95℃变性5 s,58℃复性30 s,65℃延伸5 s,共39个循环。采用相对比较 $\Delta\Delta Ct$ ( $Qr = 2^{-\Delta\Delta Ct}$ )法计算目的基因在细粒棘球绦虫原头蚴及成虫中的相对转录量并进行差异性分析。

**2.5 统计学分析** 使用Graph Pad 9.0软件对数据进行t检验分析,当P<0.05为差异有统计学意义。

## 结 果

### 1 EgTAK1基因的克隆

设计引物对目的基因进行PCR扩增,用1.0%琼脂糖凝胶电泳检测,出现两条条带,较大片段长度为1 629 bp,较小片段长度为1 116 bp(图1)。长度为1 116 bp的片段经测序与NCBIGene ID:36346045的ORF序列完全一致。长度为1 629 bp的大片段经测序与GenBank:XM024499678.1的ORF序列完全一致。将大片段与较小片段序列用BioEdit比对显示两序列高度同源。



M DNA标志物(DL2000) 1,2 分别表示内参基因Egeif和目的基因EgTAK1PCR产物

图1 EgTAK1-ORF PCR扩增产物电泳结果

M DNA marker(DL2000) 1,2 PCR products of Egeif and EgTAK1

Fig. 1 Electrophoresis results of EgTAK1-ORF amplified products

### 2 EgTAK1生物信息学分析

**2.1 理化特性** EgTAK1基因编码区由1 116个碱基组成,编码371个氨基酸,分子式为C<sub>1889</sub>H<sub>2977</sub>N<sub>517</sub>O<sub>565</sub>S<sub>19</sub>,相对分子质量为42.58×10<sup>3</sup>,理论等电点(PI)为5.97,半衰期:30 h(mammalian reticulocytes)>20 h(yeast)>10 h(*Escherichia coli*),不稳定指数为49.59(>40),属于不稳定蛋白。脂肪

系数为 87.22,总平均亲水系数为-0.291,认为是亲水蛋白。EgTAK1 所含各种氨基酸的数目以及百分比,亮氨酸(leu)和异亮氨酸(Ile)数目最多,均占 7.8%;酸性带负电荷(Asp+Glu)的氨基酸残基数为 49,碱性带正电荷(Arg+Lys)的氨基酸残基数为 43。EgTAK1 蛋白含有的抗原区域主要在 2-18、21-37、41-49、52-66、71-73、76-79、81-92、98-107、114-121、124-125、127-130、150-167、172-174、176-200、207-217、219-227、237-244、246-259、262-280、292-294、295-299、304-314、317-340、342-353、355-361 位氨基酸残基附近,其呈现在表面可能性较大的区域主要在 2-9、11-16、32-26、30-32、54-61、63-65、74-82、100-104、115-117、126-130、149-153、159-163、185-194、207-216、219-223、237-238、240-242、266-277、295-296、305-313、320-325、333-339、344-350 位氨基酸附近(图 2)。

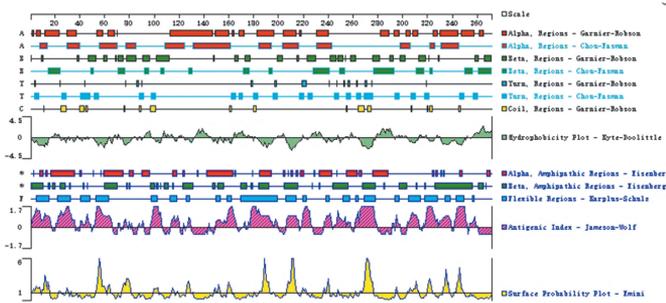


图 2 DNASTar 软件分析的 EgTAK1 的各项参数  
Fig. 2 Various parameters of EgTAK1 analysed by DNASTar

**2.2 磷酸化位点、O-糖基化位点和 N-糖基化位点**  
EgTAK1 蛋白质含有 64 个磷酸化位点,其中有 37 个丝氨酸(Ser)磷酸化位点,18 个苏氨酸(Thr)磷酸化位点,7 个酪氨酸(Tyr)磷酸化位点(图 3)。NetOGlyc4.0 在线网站分析其存在 3 个 O-糖基化潜在位点,位于第 22、26 和 108 位氨基酸处,得分分别为 0.57、0.56、0.51(均>0.5);采用 NetNGly1.0 在线软件分析发现该蛋白 N-糖基化位点(N-glycosylation)有 4 个,分别位于第 31、96、260 和 369 位氨基酸残基处(图 4)。

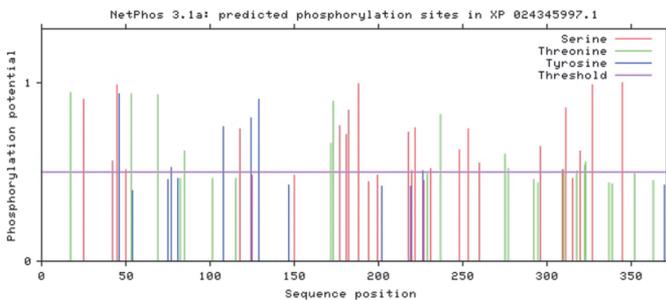


图 3 EgTAK1 蛋白磷酸化位点  
Fig. 3 Phosphorylation site of EgTAK1 protein

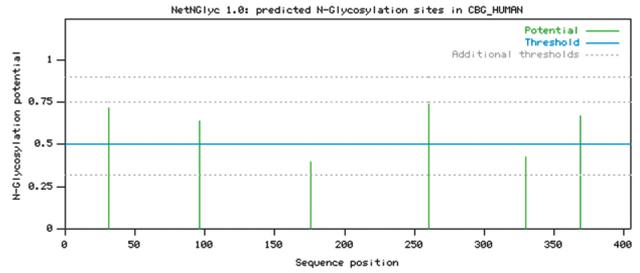
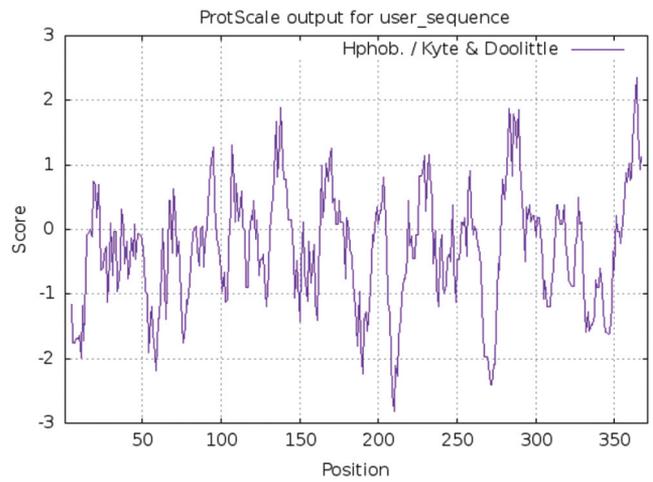


图 4 EgTAK1 蛋白 N-糖基化位点预测  
Fig. 4 predicted site of N-glycosylation of EgTAK1 proteins

**2.3 亲、疏水性** EgTAK1 蛋白序列中,第 210 位氨基酸残基具有最低值-2.811,亲水性最强;第 364 位氨基酸残基具有最高值 2.356,疏水性最强;该蛋白具有多个亲/疏水性区域分布(图 5)。



注:横轴代表 EgTAK1 的氨基酸残基位点,纵轴代表亲/疏水性得分(负值表示亲水性,正值表示疏水性)。

图 5 EgTAK1 亲/疏水性分析

Notes: The X-coordinate represents the site of amino acid residues, Y-coordinate represent score of hydrophilicity/hydrophobicity

Fig. 5 Hydrophilic/hydrophobic analysis of EgTAK1

**2.4 保守结构域** EgTAK1 蛋白具有蛋白激酶 C 样(PKC-like)超家族的保守结构域,保守区域位于 41~287 氨基酸残基之间(图 6)。



图 6 EgTAK1 蛋白的保守结构域  
Fig. 6 Conserved domains of EgTAK1 protein

**2.5 跨膜结构域、信号肽及亚细胞定位** 经预测 EgTAK1 蛋白无信号肽(图 7),可能是非分泌蛋白,且无跨膜结构域(图 8),预测 EgTAK1 蛋白可能定位于细胞膜、细胞质及核内。

**2.6 B 抗原表位** EgTAK1 蛋白有 10 个 B 细胞抗原表位(图 9),氨基酸序列见表 1,7 个可作为优势 B 细胞抗原表位。

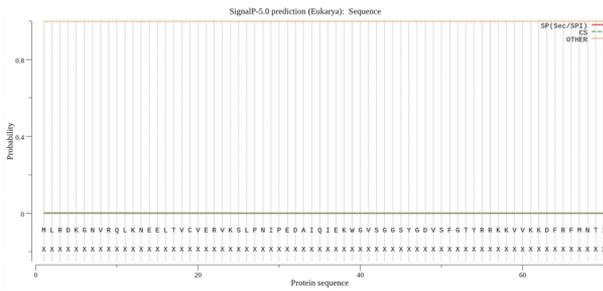


图 7 SignalP 预测 EgTAK1 信号肽  
Fig.7 SignalP predicts EgTAK1 signal peptide

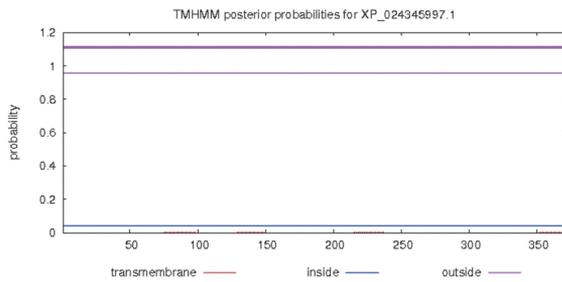


图 8 EgTAK1 蛋白的跨膜结构域  
Fig.8 Transmembrane domain of EgTAK1

表 1 IEDB 在线预测的 EgTAK1 蛋白 B 细胞抗原表位  
Table 1 prediction of the B cell epitopes of EgTAK1 protein using IEDB online

序号 No.	起始位点 Star site	结束位点 End site	氨基酸序列 Amino acid sequence	长度(bp) Length
1	5	44	KGNVRQLKNEELTVCVERVK SLPNIPEDAIQIEKRWGVS	40
2	66	75	FMNTIVERIY	10
3	103	103	M	1
4	118	120	SLQ	3
5	124	131	YSEAKYNI	8
6	185	197	FERSKEELQSVNQ	13
7	210	220	RHKDKFFASYS	11
8	240	256	LDQVVNPHSVRIRSCPP	17
9	267	284	AEDPRQRPTATQLVRLFD	18
10	303	358	ENAEAPSTSQEISITDSQTTD ALSEIPIRHEDETATIHA RES EPDIEMTVNVKDF	56

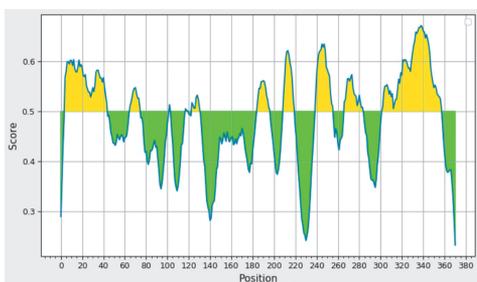


图 9 IEDB 预测的 EgTAK1 B 细胞抗原表位  
Fig.9 EgTAK1 B cell epitope predicted by IEDB

2.7 二级结构和三维结构 EgTAK1 二级结构预测结果显示： $\alpha$ -螺旋占 34.77%， $\beta$ -转角占 5.12%，无规卷

曲占 40.97%，延伸链占 19.14% (图 10)。SWISS-MODEL 模板蛋白与该蛋白的同源性为 28.35%，尚未达到同源建模要求 ( $\geq 30\%$ )。运用 Uniprot 在线软件 (<https://www.uniprot.org/uniprot>) 发现人源 TAK1 蛋白与该蛋白同源性达到 31.3%，由三维结构模型 (图 11) 可知该酶结构略显复杂，螺旋结构和无规卷曲是整个三维结构主要组成部分。此外，人源 TAK1 蛋白是由氨基末端域、DNA 结合域、铰链域和 1 个由 12 个螺旋状结构组成的配体结合域构成的，具有典型的核受体分子结构。



注：h 表示  $\alpha$ -螺旋；e 表示延伸链；c 表示无规则卷曲；t 表示  $\beta$ -转角  
图 10 SOMPA 预测的 EgTAK1 二级结构

Notes: h means  $\alpha$ -helix; e means extended connection; c means random curl; t means  $\beta$ -turn

Fig.10 The secondary structure of EgTAK1 predicted by SOMPA

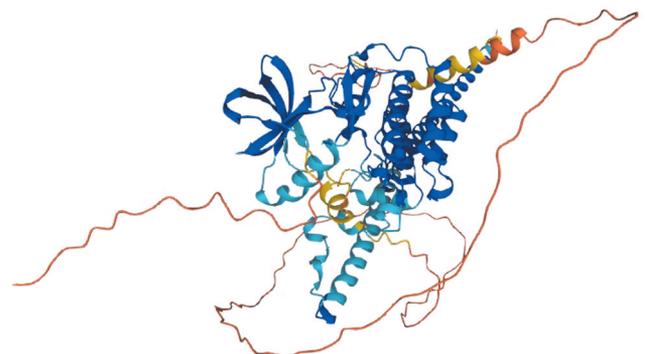


图 11 EgTAK1 的三维结构模型  
Fig.11 3D structure model of EgTAK1

2.8 氨基酸序列比对及系统进化树构建 EgTAK1 与其他物种的 TAK1 的氨基酸序列进行同源性比较，结果显示与绦虫类的多房棘球绦虫 (*Echinococcus multilocularis*, Em) 的氨基酸序列一致性是 83%，与微口膜壳绦虫 (*Hymenolepis microstoma*, Hm) 是 74%，与吸虫纲的卫氏并殖吸虫 (*Paragonimus westermani*, Pw) 是 71%，与中华枝睾吸虫 (*Clonorchis sinensis*, Cs) 是 64%，与肝片吸虫 (*Fasciola hepatica*, Fh) 是 54%，与非脊椎类的果蝇 (*Drosophila melanogaster*, Dm) 是 65%，与脊椎类的斑马鱼 (*Danio rerio*, Dr) 是 58%，与两栖类的非洲爪蟾 (*Xenopus laevis*, Xl) 是 60%，与鸟类的野鸽 (*Columba livia*, Cl) 是

56%，与哺乳类的智人 (*Homo sapiens*, *Hs*) 是 30.3%，与小鼠 (*Mus musculus*, *Mm*) 是 31.3%，与绵羊 (*Ovis aries*, *Oa*) 是 31.3%，与牛 (*Bos taurus*, *Bt*) 是 31.3%。

为明确 EgTAK1 的进化关系，将 EgTAK1 的氨基酸序列与其他物种 (见表 2) 的 TAK1 氨基酸序列 (表 2) 进行比对，采用软件 Mega7.0 完成 TAK1 氨基酸序列系统进化树 (图 12)。图 12 显示 EgTAK1 与 *Em* 以及 *Hm* 的 TAK1 在 1 个分支上，尤其 *Eg* 与 *Em* 进化关系很近。EgTAK1 与吸虫纲的 *Cs*、*Pw* 和 *Fh* 的 TAK1 虽不在 1 个分支上但相距较近。脊索动物门哺乳纲的智人、小鼠、绵羊和牛的 TAK1 在 1 个分支上，位于进化树的最顶端；往下依次为鸟纲的野鸽，两栖纲的非洲爪蟾，鱼纲的斑马鱼，节肢动物门动物，扁形动物门动物的 TAK1。

表 2 各物种 TAK1 蛋白序列登录号  
Table 2 Amino acid component of EgTAK1

物种名称 Species name	登录号 GenBank
智人 ( <i>Homo sapiens</i> )	NP_663306.1
小鼠 ( <i>Mus musculus</i> )	BAA11184.1
绵羊 ( <i>Ovis aries</i> )	XP_004011317.1
牛 ( <i>Bos Taurus</i> )	NP_001075064.1
野鸽 ( <i>Columba livia</i> )	XP_021142576.1
非洲爪蟾 ( <i>Xenopus laevis</i> )	AAC14008.1
斑马鱼 ( <i>Danio rerio</i> )	NP_001018586.1
中华蜜蜂 ( <i>Apis cerana</i> )	XP_016917872.1
温带臭虫 ( <i>Cimex lectularius</i> )	XP_014249311.1
果蝇 ( <i>Drosophila melanogaster</i> )	NP_732554.1
肝片吸虫 ( <i>Fasciola hepatica</i> )	THD20644.1
卫氏并殖吸虫 ( <i>Paragonimus westermani</i> )	KAA3673317.1
中华枝睾吸虫 ( <i>Clonorchis sinensis</i> )	GAA49925.1
微口膜壳绦虫 ( <i>Hymenolepis microstoma</i> )	CDS27658.1
多房棘球绦虫 ( <i>E. multilocularis</i> )	CDS39292.1
细粒棘球绦虫 ( <i>E. granulosus</i> )	XP_024345997.1

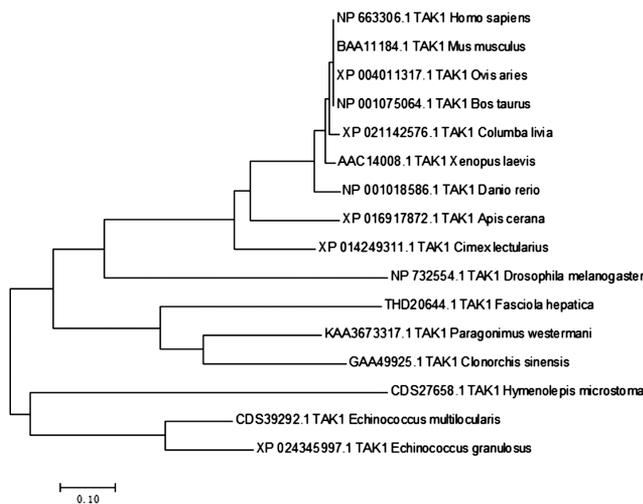


图 12 EgTAK1 系统进化树  
Fig. 12 The Phylogenetic tree of EgTAK1

2.9 EgTAK1 基因表达差异性 通过 qPCR 检测 *EgTAK1* 在细粒棘球绦虫不同发育阶段表达情况，结果显示 *EgTAK1* 在 *E. g* 的生发层，原头蚴和成虫均表达，在成虫阶段表达量最高，在成虫的表达与原头蚴相比具有差异，差异有统计学意义 ( $t = 7.935, P < 0.05$ )，在原头蚴的表达与生发层相比具有差异，差异有统计学意义 ( $t = 36.24, P < 0.05$ ) (图 13)。

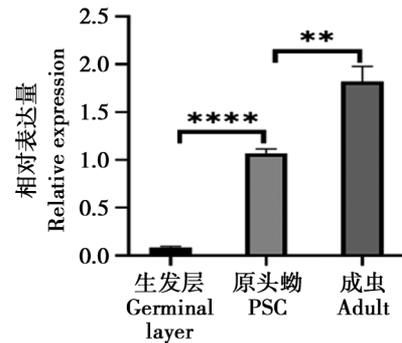


图 13 *EgTAK1* 基因在细粒棘球绦虫不同发育阶段转录水平的相对表达量  
Fig. 13 Differential transcription of TAK1 in different stages of *Eg*

## 讨论

包虫病是由感染细粒棘球绦虫的幼虫而引发的世界性公共卫生问题。中间宿主 (人或羊) 一旦误食了含有虫卵的食物，经过宿主的胃液和肠液消化后，脱壳的虫卵变成六钩蚴，六钩蚴可以穿过小肠肠壁进入血流，然后通过血液和淋巴液循环系统运输到达靶器官定居并发育成包虫囊，进而在局部浸润或生长。国内流行病学调查显示：该病新增病例高达 1.6 万/年，直接损失的经济可达 30 亿元/年<sup>[12]</sup>。目前，包虫病缺乏有效的治愈药物，这仍是 1 个亟待解决的关键问题。控制包虫病重在预防，筛选调控虫体发育的关键基因或敏感、特异的抗原是突破难题的关键。细粒棘球绦虫具有双向发育的特征，胆汁酸盐是决定其向成虫发育的关键因素<sup>[13]</sup>。杨梅等<sup>[5]</sup>通过有无犬胆汁培养的原头蚴转录组学分析发现 *EgTAK1* 基因高表达，虽然已有实验证明胆汁酸盐对原头蚴分化为成虫有重要的调控作用<sup>[14]</sup>，但具体调控机理尚不明确。

本研究从自然感染细粒棘球绦虫的绵羊肝中采集原头蚴，提取其总 RNA 并反转录出 cDNA，用此 cDNA 做模板克隆 *EgTAK1* 基因 ORF 全长，结果获得两条明显条带，其中小片段测序结果为 1 116 bp，与 NCBI 中编码 *EgTAK1* (蛋白序列号 XP-024345997.1) 的 ORF 序列大小完全一致，大片段测序结果为 1 629 bp，在 NCBI 进行序列比对发现是 *EgMAP3K* 基因的 ORF 序列 (XM-024499678.1)，后者与 *EgTAK1* 的 ORF 是同源序列 (相似度 79%)。 *EgTAK1* 的

ORF 全长为 1 116 bp, 共编码 371 个氨基酸。通过一系列在线软件对 *EgTAK1* 蛋白的氨基酸序列进行生物信息学预测, 分析得出了此蛋白的理化特性和蛋白结构特征, 发现它具有 PKC 样蛋白超家族的特征结构域; 预测该蛋白无跨膜结构和信号肽, 属于亲水性蛋白, 可能定位于细胞质、细胞膜及细胞核内。该蛋白含有许多表面抗原位点, 且含有 10 个 B 细胞抗原表位, 表明其能够刺激机体引发体液免疫应答。通过系统进化树分析发现 *EgTAK1* 蛋白与多房棘球绦虫和微口膜壳绦虫的 TAK1 进化关系较近, 图 12 显示 *EgTAK1* 与 *Em* 以及 *Hm* 的 TAK1 在 1 个分支上, 这 3 种生物都属于扁形动物门的绦虫纲, 进化关系较近; 尤其是 *Eg* 与 *Em* 进化关系很近。*EgTAK1* 与吸虫纲的 *Cs*、*Pw* 和 *Fh* 的 TAK1 虽不在 1 个分支上但相距较近, 绦虫纲与吸虫纲都属于扁形动物门, 这也符合进化关系。整个 TAK1 氨基酸序列系统进化树结构符合动物分类学的进化关系, 脊索动物门哺乳纲的智人、小鼠、绵羊和牛的 TAK1 在 1 个分支上, 位于进化树的最顶端; 往下依次为鸟纲的野鸽, 两栖纲的非洲爪蟾, 鱼纲的斑马鱼, 节肢动物门动物, 扁形动物门动物的 TAK1。

TAK1 在进化上是高度保守的激酶, 人和小鼠 TAK1 的氨基酸序列同源性可达 98%。研究表明, 人的 TAK1 基因全长 73.6 kb, 位于 6q16.1 和 q16.3 之间, 由 16 个内含子和 17 个外显子 (alternative exons, AE) 组成, 大小剪接变体 TAK1-a, b, c 和 d 是由 AE12 和 AE16 两个可变的 AE 共同决定的<sup>[15]</sup>。在秀丽隐杆线虫中 TAK1 参与的 Wnt 信号通路已被广泛研究, 特别是 TAK1 和 NLK 的同源物 mom-4 和 LIT-1, 在秀丽隐杆线虫早期的胚胎中负调控 wnt 控制的细胞命运; 由于 TAK1 可被 TGF- $\beta$  和各种细胞因子激活, 它可能为通过其他途径调控 wnt 系统提供了 1 个切入点<sup>[11,16]</sup>。

通过实时荧光定量 PCR 结果发现 *EgTAK1* 在细粒棘球绦虫生发层, 原头蚴和成虫中均表达, 在成虫阶段相对表达量最高, 在原头蚴表达量次之, 且各个差异均有统计学意义。该基因可能在细粒棘球绦虫成虫中发挥重要作用。

综上所述, *EgTAK1* 基因在细粒棘球绦虫的 3 个主要发育阶段均有表达, 其在细粒棘球绦虫成虫发育分化中的作用有待进一步研究。其具有良好的免疫原

性, 有望成为抗包虫病的靶基因。

#### 【参考文献】

- [1] Siracusano A, Delunardo F, Teggi A, et al. Host-parasite relationship in cystic echinococcosis: An evolving story[J]. Clin Develop Immunol, 2012(2012): 639362.
- [2] 李杰, 马岩, 戴庆, 等. 新疆地区包虫病致过敏性休克患者特异性免疫反应特点研究[J]. 热带医学杂志, 2021, 21(9): 1139-1142.
- [3] Loos JA, Coccimiglio M, Nicolao MC, et al. Metformin improves the therapeutic efficacy of low-dose albendazole against experimental alveolar echinococcosis[J]. Parasitology, 2022, 149(1): 138-144.
- [4] Shaprinskiy V, Verba A, Formanchuk T, et al. Surgical treatment of echinococcosis of the liver and its complications[J]. Wiadomosci Lekarskie, 2022, 75(1 pt 2): 244-250.
- [5] 杨梅, 王甜, 吴军, 等. 胆汁酸诱导的细粒棘球绦虫分化基因的转录组学分析[J]. 职业与健康, 2020, 36(10): 1316-1323.
- [6] Pathak S, Borodkin VS, Albarbarawi O, et al. Oglcnacylation of tab1 modulates tak1-mediated cytokine release[J]. EMBO J, 2012, 31(6): 1394-1404.
- [7] Xu YR, Lei CQ. Tak1-tabs complex: A central signalosome in inflammatory responses[J]. Fron Immunol, 2020(11): 608976.
- [8] Hirata Y, Takahashi M, Morishita T, et al. Post-translational modifications of the tak1-tab complex[J]. Internati J Mol Sci, 2017, 18(1): 205.
- [9] 王成华, 吕海龙, 姜玉峰, 等. 棘球蚴 mapk 信号转导通路的研究进展[J]. 中国寄生虫学与寄生虫病杂志, 2013, 31(1): 60-63.
- [10] Ohkawara B, Shirakabe K, Hyodo-Miura J, et al. Role of the tak1-nlk-stat3 pathway in tgf-beta-mediated mesoderm induction[J]. Gene Develop, 2004, 18(4): 381-386.
- [11] Behrens J. Cross-regulation of the wnt signalling pathway: A role of map kinases[J]. J Cell Sci, 2000, 113 ( Pt 6): 911-919.
- [12] Jide A, Zhang J, Chai J, et al. Comparison of the efficacy of anatomic and non-anatomic hepatectomy for hepatic alveolar echinococcosis: Clinical experience of 240 cases in a single center [J]. Fron Public Healt, 2021(9): 816704.
- [13] 司晓妹, 马俊英, 张雪飞, 等. 细粒棘球绦虫原头蚴生长发育机制研究进展[J]. 热带医学杂志, 2021, 21(12): 1622-1625, 1629.
- [14] 李军, 郭刚, 张壮志, 等. 棘球绦虫的主要生理学特点与包虫病控制方法的确立[J]. 中华预防医学杂志, 2018, 52(2): 210-214.
- [15] Tripathi V, Shin JH, Stuelten CH, et al. Tgf- $\beta$ -induced alternative splicing of tak1 promotes emt and drug resistance[J]. Oncogene, 2019, 38(17): 3185-3200.
- [16] Phillips BT, Kimble J. A new look at tcf and beta-catenin through the lens of a divergent c. Elegans wnt pathway[J]. Developmental Cell, 2009, 17(1): 27-34.

【收稿日期】 2022-08-24 【修回日期】 2022-11-01