



## 日本血吸虫真核生物翻译起始因子2 $\alpha$ 亚基全长cDNA的克隆与功能分析

日本血吸虫病仍然在亚洲一些国家流行,特别是在中国的一些地区及菲律宾群岛。截至2001年底,中国尚有12个省的418个血吸虫病流行县,流行区总人口约99万,患病人数居高不下[1]。在人类基因组计划的带动下,1992年美国基因组研究所(Institute for genomic research, TIGR, USA)发起了全球性的血吸虫基因组计划(Schistosoma Genome Project, SGP)[2],该计划从1994年起受到“UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR)”基金的支持[3]。“血吸虫基因发现计划”是SGP的一部分,该计划的主要目标之一是发现与鉴定日本血吸虫与曼氏血吸虫的新基因,从而寻找新的治疗药物与疫苗[4]。表达序列标签(expression sequence tag, EST)是文库中随机挑选的cDNA末端的短序列(大约300碱基)[5]。EST可以利用数据库搜索进行DNA或蛋白质的同源性分析,从而鉴定这些基因的起源[6]。我们在国内外首次构建了日本血吸虫尾蚴cDNA文库[7],在用EST策略从文库中筛选到的201个EST序列中,136个合乎要求的序列进行了同源性分析。在这136个EST中,有61.8%的序列与数据库中的蛋白或DNA序列无同源性;有38.2%的EST序列与已知基因或已知蛋白具有同源性,其中有5.9%及6.6%的EST分别与日本血吸虫及曼氏血吸虫已知基因高度同源,而有25.7%的EST与其它生物的相应基因或蛋白高度同源。我们对一个与曼氏血吸虫真核生物翻译起始因子2 $\alpha$ 亚基(eukaryotic translation initiation factor 2 alpha subunit, eIF2 $\alpha$ ) mRNA 序列高度同源的cDNA片段进行了研究,克隆了其全长开放读码框序列。

### 1 材料

#### 1.1 文库

日本血吸虫(中国大陆株)尾蚴cDNA文库由本室用CLONTECH的SMART<sup>TM</sup> cDNA文库构建试剂盒构建,一级文库库容量为 $1.8 \times 10^7$  pfu[7]。

#### 1.2 测序引物及测序

选出文库载体插入位点上游的序列设计测序引物:5'-CTCCGAGATCTGGACGAGC-3'。测序由上海基康公司完成,PERKIN ELMER公司全自动核酸分析仪(ABI PRISM<sup>TM</sup> 310 Genetic Analysis)测序。

#### 1.3 质粒

pGEM-T vector购自PROMEGA公司。

#### 1.4 工具酶和试剂

Taq E、dNTP、EcoR I、Xho I 购自广州基因公司, DNA marker: G02S 100 bp ladder购自上海博彩科技有限公司。

### 2 方法

## 2.1 噬菌体转化成质粒

按Clontech SMART™ cDNA 文库构建试剂盒 操作说明进行

## 2.2 质粒的测序和EST的获取

碱裂解法提取质粒，用 5' 端测序引物测出一个反应的序列。将所得到的原始序列进行分析编辑，除去质粒序列，留下插入的DNA片段，经过编辑后的序列称为一个EST，5' 端不具有接头序列及长度小于150 bp的EST都被舍弃。

## 2.3 EST 分析

通过NCBI数据库(<http://www.ncbi.nlm.nih.gov/blast>)的BLASTn、BLASTx程序分别在核苷酸及蛋白质水平上对每一个EST 进行同源性搜索。

## 2.4 全长cDNA序列的获取及功能分析

2.4.1 PCR扩增 以测序引物为上游引物：5'-CTCCG AGATCTGGACGAGC-3'，并根据EST序列设计与之相匹配的下游引物：5'-GCCTCGAGTCATATAATA TATTACTT-3'。以尾蚴cDNA文库为模板进行PCR扩增，琼脂糖凝胶电泳分析PCR结果。

2.4.2 目的基因的T载体连接及测序 PCR产物回收纯化后连接入pGEM-T 载体，转化大肠杆菌DH5 $\alpha$ ，并筛选出阳性克隆。质粒送往基康公司进行测序。

2.4.3 对全长cDNA序列进行结构分析及功能预测 (1)通过NCBI数据库(<http://www.ncbi.nlm.nih.gov/blast>)的BLASTn、BLASTx程序分别在核苷酸及蛋白质水平上对每一个EST 进行同源性搜索；(2)用NCBI BLAST站点的blast two sequence 程序(<http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html>)对同源性高的基因进行核苷酸及氨基酸水平的同源性比较；(3)利用Motif Scan in a Protein Sequence([http://hits.isb-sib.ch/cgi-bin/hits\\_motifscan](http://hits.isb-sib.ch/cgi-bin/hits_motifscan))对cDNA 序列所编码的蛋白质进行结构域搜索；(4)利用CD-Search(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 程序对目标蛋白进行保守区域的搜索。

# 3 结果

## 3.1 EST Blast结果

在所得到的EST序列中，JCXP0006.FXS克隆的cDNA序列(片段长543 bp，dbEST登录号为gi1147534)与曼氏血吸虫eIF2 $\alpha$ 的mRNA高度同源。

## 3.2 日本血吸虫eIF2 $\alpha$ 全长cDNA序列

日本血吸虫eIF2 $\alpha$ 全长cDNA序列与曼氏血吸虫eIF2 $\alpha$ 的mRNA(GenBank 登录号为AF376135\_1)高度同源，该序列包含一个完整的编码区，由起始密码子ATG至终止密码子TAA止，共有981 bp，编码由327个氨基酸组成的蛋白(图1)。

## 3.3 日本血吸虫(Sj)与曼氏血吸虫(Sm)eIF2 $\alpha$ 基因的同源性比较

二者同源性为87%，具体见图2。

## 3.4 日本血吸虫与曼氏血吸虫eIF2 $\alpha$ 蛋白水平的同源性比较

用NCBI BLAST站点的BLASTx程序对Sj-eIF2 $\alpha$  cDNA序列进行搜索，发现其与Sm-eIF2 $\alpha$ 氨基酸水平的同源性达79%(图3)。

## 3.5 日本血吸虫eIF2 $\alpha$ 蛋白结构域搜索结果

利用Motif Scan in a Protein Sequence对cDNA 序列所编码的蛋白质进行结构域搜索，结果表明，在上述sj-eIF2 $\alpha$ 氨基酸序列中，从第17位Asp到第86位Lys是由70个氨基酸组成的S1结构域(S1 domain)。

利用CD-Search 程序对目标蛋白进行保守区域的搜索，未发现保守区域。

```

1   ATGCCCATACAGTGCAGATATTATGAGGACTTATTCOCTGAAGTTGGGGATGTTGTAGTT
1   M P I Q C R Y Y E D L F P E V G D V V V
61  GTAACAGTTAAGGTCATCCAACCTATGGGTAGCTATCTGGAAGTCTAGAGTACAAGAAT
21  V T V R V I Q P M G S Y V E L L E Y K N
121 ATTGGAGGTATGATATTGCTCAGTGAGCTGTCTCGTGCACGTATAAGATCAATCAGCAAG
41  I G G M I L L S E L S R R R I R S I S K
181 CTTGTTGGAATCGGGTCAAATACTGAGGTGACTGTTGTACGACTCGATAGTGCCAAAGGT
61  L V R I G S N T E V T V V R V D S A K G
241 TATATTGATTTATCAAAAAGAAGAGCTTCTGCAGAGGAGATCGCAAAGTGCAAGGAAAGG
81  Y I D L S K R R A S A E E I A K C K E R
301 TTTGCAGAGCCTAAAGCGGTCAATCAAATTTTGAGGAATGTACCCGACAAGTTAGAGTAT
101 F A E A K A V N Q I L R N V A E K L E Y
361 CAGACAGATGTACAACCTCGAAGAGCTGTGCAGGAAAAGTCTTGTGATTTTGACAAAAAG
121 E T D V Q L E E L C R K T A W Y F D K K
421 ACTGGTCAAAGGCAGGGTCTACGATATTTTCAACAAGGTTGTAATTCACCAGAAATT
141 T G R K A G F Y D I F K K V V N S P E I
481 CTGGATGAGTGTGACATAGACCAGCCAACAAAAGAAATGCTTCTTACGGATATCAGGCAT
161 L D E C D I D Q P T K E M L L T D I R H
541 CGAATTAACACCAAAGGCAGTGAAAATTCGGGTGACTTTGAAGTTTCATGTTTCACTTAT
181 R L T P K A V K I R A D F E V S C F T Y
601 GATGGTATTGATGCCGTCAGAAGCGCACTTCGATCTGGACTGAAACTCAATTCAGATGCT
201 D G I D A V R S A L R S G L K L N S D A
661 CTTCCAATCCGCATCAATTTGATAGCCGCACCGCTTTATGTACTAACCACGCAAACAATG
221 L P I R I N L I A P P L Y V L T T Q T M
721 GATCCAGCTGCCGGCTTAGAACAGCTGAATGAAGTGCTAAACGTCATCCAGACATCCATT
241 D R A A G L E Q L N E V L N V I Q T S I
781 GAAAGTCAAGGTGCTTCTTTCAAATTCGACAGGCTCCTCGTGTGTTTCAGATACAGAT
261 E S Q G G S F K I R Q A P R V V S D T D
841 GACGCAGAATTACAGCGTCAAATGGATGAACTAGAAAAGCGGAATCGTGAAGTTTCTGGT
281 D A E L Q R Q M D E L E K A N R E V S G
901 GATGAGGATGATGAAGACCATGACGAGGATGAAGATGATGATGAGGAATCCAATGATGGA
301 D E D D E D D D E D E D D D E E S N D G
961 GATCAGAATGAACATAAGTAA
321 D Q N E H K *

```

图1 日本血吸虫尾蚴eIF2 $\alpha$  cDNA序列及所编码的氨基酸序列

Fig.1 cDNA and amino acid sequences of the of alpha subunit mRNA eukaryotic translation initiation factor 2 from *Schistosoma japonicum cercariae*





Sm: 781 gaaagtcaaggtgggttcattcaaaaattcaacaggctgctcgagttggttcagatacagat 840  
Sj: 781 gaaagtcaaggtgggttcattcaaaaattcaacaggctgctcgagttggttcagatacagat 840

Sm: 841 gacgcggaacttacancgccaatggatgaattagaaaaagcgaatcgtgaagtttctggt 900  
Sj: 841 gacgcggaatttacagcgtcaatggatgaacttagaaaaggcgaatcgtgaagtttctggt 900

Sm: 901 gatgaggacgatgaagacgaagatgacggggatgaggaagatgatgatgaagcatcgaat 960  
Sj: 901 gatgaggatgatgaagac—gatgacgaggat—gaagatgatgatgaggaatccaat 954

Sm: 961 gatggagaccaaaatg 976  
Sj: 955 gatggagatcagaatg 970

图2 日本血吸虫eIF2 $\alpha$  cDNA与曼氏血吸虫eIF2 $\alpha$  mRNA序列同源性比较

Fig. 2 Homology comparison between from *Schistosoma mansoni* (Sm) alpha subunit of translation initiation factor 2 mRNA and alpha subunit cDNA of *Schistosoma japonicum* (Sj) translation initiation factor 2  
-: Show heterogenous gene between Sm and Sj

eIF2在蛋白质生物合成的起始中扮演着重要角色，是蛋白质合成不可缺少的起始因子。eIF2是一个异构三聚体，由 $\alpha$ 、 $\beta$ 、 $\gamma$ 三个亚基构成。Nonato等[8]报道了人类eIF2 $\alpha$ 的部分结构，该部分由eIF2 $\alpha$ 的3~182位氨基酸残基组成，约占人类eIF2 $\alpha$ 全长N-末端的2/3。其结构主要由两个部分组成，一部分为寡聚核苷酸结合域(oligonucleotide-binding domain)，由五条 $\beta$ 链( $\beta$ 1,  $\beta$ 2,  $\beta$ 3,  $\beta$ 4,  $\beta$ 5)相互缠绕形成一个封闭的桶状结构，另一部分为螺旋状结构域(helical domains)。这两部分结构共同形成了一个S1结构域，该结构域中包含RNA结合区(RNA-binding domain)，但目前尚无实验证明eIF2 $\alpha$ 与RNA结合。eIF2 $\alpha$ 具有重要的生物学功能即参与蛋白合成起始的调控。许多实验证明，在病毒感染、热休克、重金属中毒、氨基酸极度缺乏的状态下的细胞以及一些血清的细胞中，eIF2 $\alpha$ 第51位丝氨酸(Ser-51)可被一些蛋白激酶如：HRI、RNAPKR、GCN2、PERK(PKR-like ER kinase)等磷酸化，继而翻译的起始被中断，蛋白合成被抑制，细胞的生长受到抑制[9]。这一机制与病毒感染的细胞和肿瘤细胞中的蛋白质合成密切相关，因此，许多有关eIF2 $\alpha$ 磷酸化的研究正在进行之中[10]。

我们从日本血吸虫尾蚴cDNA文库中筛到并克隆的与曼氏血吸虫eIF2 $\alpha$ 亚基基因同一性高达87%的cDNA全编码区序列，经过几种蛋白分析软件的分析后发现它具有已知的eIF2 $\alpha$ 亚基的特殊结构即S1结构域。从以往的资料可知，S1结构域是首先从核糖体蛋白S1中发现的RNA结合区，S1蛋白被证明通过其S1结构域连接RNA。S1结构域也被证明存在于其它一些物种的eIF2 $\alpha$ 序列中。

综上所述，所发现的cDNA 编码日本血吸虫的eIF2 $\alpha$ 。其全长cDNA序列已经准确克隆到pGEM-T 载体上，为进一步的功能研究奠定了基础。

Score = 490 bits (1261), Identities = 257/325 (79%)

Sj: 1 MPIQCRYEDLFPEXXXXXXXXXXIQPMGSYVELLEYKNIGGMXXXXXXXXXXXXXXXXX 60  
 MPIQCR+YEDLFPE IQ MGSYVELLEYKNIGGM

Sm: 1 MPIQCRFYEDLFPEVGDVVLVTVKVIQSMGSYVELLEYKNIGGMILHSELSRRRIRSISK 60

Sj: 61 XXXXGSNTEVTVVRVDSAKGYIDLKRRASAEI AKCKERFAEAKAVNQILRNVAEKLEY 120  
 GSNTEVTVVRVDSAKGYIDLKRRASAEI AKCKERFA+AKAVNQILRNVAEKL+Y

Sm: 61 LVRIGSNTEVTVVRVDSAKGYIDLKRRASAEI AKCKERFAKAKAVNQILRNVAEKLDY 120

Sj: 121 ETDVQLEELCRKTAWYFDKKTGRKAGFYDIFKKVVNSPEILDECDIDQPTKEMLLTDIRH 180  
 +TD QLEELCRKTAWYFD+KTGR+AG YDIFKKVVNSPEILDECDIDQ TKEMLLTDIRH

Sm: 121 KTDEQLEELCRKTAWYFDRKTGRRAGSYDIFKKVVNSPEILDECDIDQATKEMLLTDIRH 180

Sj: 181 RLTPKAVKIRADFEVSCFTYDGDIDAVRSALRSGLKLNSDALPIRINLIAPPLYVLTQT 240  
 RLTPKAVKIRADFEVSCFTYDGDIDAV+SALRSGL+LNSD+LPIRINLIAPPLYVLTQT

Sm: 181 RLTPKAVKIRADFEVSCFTYDGDIDAVKSALRSGLELNSDSLPIRINLIAPPLYVLTQT 240

Sj: 241 DRAAGLEQLNEVLNVIQTSIESQCGSFKIRQAPRVVSDTDDAELQRQMDELEKANREVS 300  
 DRAAGLEQLNEVL+VI+ SIESQCGSFKI+QA RVVSDTDDA+LQRQMDELEKANREVS

Sm: 241 DRAAGLEQLNEVLDVIKKSIESQCGSFKIQQAARVVSDTDDADLQRQMDELEKANREVS 300

Sj: 301 --XXXXXXXXXXXXXXXXXSNLGDQN 325  
 SNLGDQN

Sm: 301 DEDDEDDEDDGDEEDDEASNLGDQN 325

图3 日本血吸虫与曼氏血吸虫eIF2 $\alpha$ 蛋白序列同源性比较

Fig. 3 Homology comparison of the protein sequences of alpha subunit of the translation initiation factor 2 from *Schistosoma mansoni* and *Schistosoma japonicum*

参考文献:

- [1] 陈贤义, 姜庆五, 王立英, 等. 2001年全国血吸虫病疫情通报[J]. 中国血吸虫病防治杂志(Chin J Shchisto Control), 2002, 14(4): 241-3.
  - [2] Pena S. Third world participation in genome projects[J]. Trends Biotechnol, 1996, 14(1): 74-7.
  - [3] Johnston AD. The WHO/UNDP/World Bank Schistosoma genome initiative: current status[J]. Parasitol Today, 1997, 13(1): 45-6.
  - [4] Unnasch T. The filarial genome project[J]. Parasitol Today, 1994, 10(3), 415-6.
  - [5] Gloria R, Mark D, Bento S, et al. Identification of new *Schistosoma mansoni* genes by the EST strategy using a directional cDNA library[J]. Gene, 1995, 152(2): 141-7.
  - [6] Adams M, Dubrick M, Keriavage A, et al. Complementary cDNA sequencing : expressed sequence tags and human genome project[J]. Science, 1991, 252(5013): 1651-6.
  - [7] 陈晓光, 李华, 彭鸿娟, 等. 日本血吸虫尾蚴cDNA文库的构建及分析[J]. 中国寄生虫学与寄生虫病杂志, 2002, 20(1): 39-41.
- Chen XG, Li H, Peng HJ, et al. Construction and characterization of the cDNA library

from *Schistosoma japonicum cercariae*[J]. *Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi*, 2002, 20(1): 39-41.

[8] Nonato MC, Widom J, Clardy J. Crystal structure of the N-terminal segment of human eukaryotic translation initiation factor 2[J]. *J Biol Chem*, 2002, 10; 277(19): 17057-61.

[9] Pain VM. Initiation of protein synthesis in eukaryotic cells[J]. *Eur J Biochem*, 1996, 15; 236(3):747-71

[10] Kaufman RJ. Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls[J]. *Genes Dev*, 1999, 15; 13(10): 1211-33.

#### 参考文献:

[1] 陈贤义, 姜庆五, 王立英, 等. 2001年全国血吸虫病疫情通报[J]. *中国血吸虫病防治杂志* (*Chin J Shchisto Control*), 2002, 14(4): 241-3.

[2] Pena S. Third world participation in genome projects[J]. *Trends Biotechnol*, 1996, 14(1): 74-7.

[3] Johnston AD. The WHO/UNDP/World Bank *Schistosoma* genome initiative: current status[J]. *Parasitol Today*, 1997, 13(1): 45-6.

[4] Unnasch T. The filarial genome project[J]. *Parasitol Today*, 1994, 10(3), 415-6.

[5] Gloria R, Mark D, Bento S, et al. Identification of new *Schistosoma mansoni* genes by the EST strategy using a directional cDNA library[J]. *Gene*, 1995, 152(2): 141-7.

[6] Adams M, Dubrick M, Keriavage A, et al. Complementary cDNA sequencing : expressed sequence tags and human genome project[J]. *Science*, 1991, 252(5013): 1651-6.

[7] 陈晓光, 李华, 彭鸿娟, 等. 日本血吸虫尾蚴cDNA文库的构建及分析[J]. *中国寄生虫学与寄生虫病杂志*, 2002, 20(1): 39-41.

Chen XG, Li H, Peng HJ, et al. Construction and characterization of the cDNA library from *Schistosoma japonicum cercariae*[J]. *Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi*, 2002, 20(1): 39-41.

[8] Nonato MC, Widom J, Clardy J. Crystal structure of the N-terminal segment of human eukaryotic translation initiation factor 2[J]. *J Biol Chem*, 2002, 10; 277(19): 17057-61.

[9] Pain VM. Initiation of protein synthesis in eukaryotic cells[J]. *Eur J Biochem*, 1996, 15; 236(3):747-71

[10] Kaufman RJ. Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls[J]. *Genes Dev*, 1999, 15; 13(10): 1211-33.