

一个血清抑制基因的克隆*

谭德勇** 赖建华 钱伟 余敏 陈瑶 何婕

(云南大学生命科学学院, 昆明 650091)

摘要 比较血清培养细胞和血清饥饿细胞的基因表达差异, 获得了一段血清饥饿细胞中特异表达的 cDNA 序列, 以此序列出发, 通过搜索表达序列标签 (EST), 拼接出完整的基因序列, 通过 PCR 分段克隆获得全长 cDNA 序列. 该基因全长 5 429 bp, 编码框预测有 791 个氨基酸残基. GenBank 搜索, 该基因与已有的细胞周期调控基因没有同源性. 所以, 该基因是一个新的与细胞周期有关的基因 (GenBank 接受号: AY050169). 由于该基因最初发现在无血清培养条件下表达, 故叫血清抑制基因 (serum inhibit gene, Si-1 基因).

关键词 血清饥饿, 基因, 克隆

学科分类号 Q78

血清是细胞生长的必需成分, 其中含有细胞生长的必需因子, 在没有血清时细胞生长停止在 G0 期. Gupta 等^[1]报道血清饥饿导致真核生物细胞翻译起始因子的表达减弱; Helbing 等^[2]报道血清饥饿可以提高肿瘤抑制基因 p33ING1 的表达; Ahn 等^[3]报道在血清饥饿诱导白介素 6 和白介素 8 的表达, 而 Fas 相关的死亡结构域蛋白 (Fas-associated death protein, FADD) 和 caspase-8 基因的表达减弱. 本实验室的研究结果表明在血清饥饿时, 细胞生长正调控基因表达减弱, 细胞生长负调控基因表达增强^[4]. 这些结果表明, 血清饥饿导致细胞内基因表达的变化. 所以, 通过比较血清培养和血清饥饿细胞中基因的表达差异, 可以找到控制细胞生长有关的基因, 本研究采用这一方法克隆了一个血清抑制基因.

1 材料与方法

1.1 材料

人脑胶质瘤细胞株 U251 由重庆医科大学范维珂教授提供; 细胞培养基购于 Gibco 公司; PCR 扩增试剂购于 Promega 公司, RNA 分离纯化试剂和凝胶回收试剂盒由上海华舜公司供货; cDNA 合成试剂盒和 DNA 连接试剂盒由大连宝生物公司提供.

1.2 方法

1.2.1 引物合成: 由上海生工生物工程技术服务公司合成.

1.2.2 细胞培养: a. 培养条件: U251 细胞用含 10% 血清的 RPMI1640 培养基, 5% CO₂, 37℃ 培

养; b. 血清饥饿处理: 在常规培养条件下培养的细胞, 经 0.3% 胰酶消化后, 以 50% 的密度 (细胞贴壁后占培养面 50% 的面积) 铺满, 培养 24 h 后, 换无血清培养基, 培养 48 h, 收获细胞作为血清饥饿培养细胞; 继续培养 24 h 收获细胞作为血清培养的对照细胞; c. 恢复血清培养细胞: 当血清饥饿处理 48 h 后, 换 10% 血清培养基培养 48 h 的细胞; d. 原位杂交培养细胞处理: 将盖玻片放入指形培养瓶中, 并接种细胞, 使细胞长在盖玻片上. 生长在盖玻片上的细胞置于 100% 的乙醇中固定 30 min, 用于细胞原位杂交.

1.2.3 RNA 抽提: 按试剂盒的操作程序提取总 RNA.

1.2.4 cDNA 合成: 取前述提取的 RNA 样品 5 μl (约 1 μg), 加入 10 μmol/L Oligo (dT)₂₀ 1 μl, 72℃ 处理 2 min, 冰上骤冷 1 min, 加入 5 × 含 DTT 的缓冲液 2 μl, 10 mmol/L dNTP 1 μl, 逆转录酶 (RAV-2) 1 μl (200 U), 总体积 10 μl, 42℃ 保温 1 h, 冰上冷却 2 min, 4℃ 贮存备用.

1.2.5 基因转录表达的差异显示: 与常规的 DD-PCR 方法相似^[5,6], 略有改变. 5' 随机引物为 cttcattgcc, 3' 端分别用 12 条锚定引物: (dT)₁₇N₁N₂ (N₁ 为 A, G, C, N₂ 为 A, T, C, G). 反应体系: 引物各 2.5 μl, 10 mmol/L dNTP 1 μl, 逆转录产物

* 国家自然科学基金 (39960030) 和云南省应用基础研究重点项目 (1999C002Z) 资助.

** 通讯联系人.

Tel: 0871-5032061, E-mail: dytan3@public.km.yn.cn

收稿日期: 2002-03-27, 接受日期: 2002-05-28

0.5 μ l, 10 $\times$ 缓冲液 2.5 μ l, Taq DNA 聚合酶 1 U, 加水至 25 μ l. 循环程序: 95 $^{\circ}$ C 2 min; 95 $^{\circ}$ C 30 s, 42 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 2 min, 35 个循环; 72 $^{\circ}$ C 10 min. PCR 产物 7 μ l 在 8% 非变性聚丙烯酰胺凝胶上电泳, 当二甲苯青至胶底时结束电泳, 以 EB 染色显带. 从凝胶上切下差异带后, 置于干净的 500 μ l 管中, 加四蒸水 100 μ l, 沸水煮 5 min, 离心, 取上清液 5 μ l 进行二次 PCR 扩增, 反应条件同第一次 PCR 扩增. 琼脂糖凝胶电泳分离纯化二次 PCR 扩增产物, 利用柱式胶回收试剂盒回收第二次 PCR 扩增产物.

1.2.6 测序: 将回收的二次 PCR 扩增产物直接克隆到 T-vector 载体上, 用 T7 启动子序列进行测序. 测序由宝生物公司完成.

1.2.7 GenBank 同源比较: 利用美国国立生物信息中心网站在线工具 Blast, 在 GenBank 数据库中进行同源比较, 并检索相关信息^[7].

1.2.8 细胞原位杂交: 按 DIG DNA 标记和检测试剂盒的操作说明操作 (见 Roche 公司的产品说明书).

2 结 果

2.1 DD-PCR 寻找差异片段

比较血清饥饿细胞和血清培养细胞的 DD-PCR 的部分结果见图 1, 图 1 中一部分条带在血清饥饿细胞中特有, 一部分条带在血清条件下特有. 其中箭头所示的片段在血清饥饿条件下表达较强而在血清培养条件下表达很弱.

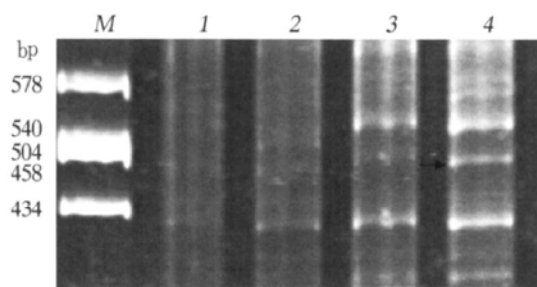


Fig. 1 a part of DD-PCR map which shows the gene expression difference between normal cells and serum starvation cells

M: pBR322/Hae III marker; 1 and 3: normal cells, 2 and 4: serum starvation cells; 1, 2 show the products amplified by same primer; 3, 4 show the products amplified by same primer.

2.2 原位杂交结果

将图 1 中箭头所指片段连凝胶挖出, 用同样的引物进行第二次扩增, 并用地高辛标记, 进行细胞原位杂交, 结果见图 2. 血清培养细胞中无任何表达信号 (图 2a), 在血清饥饿细胞中有强阳性信号 (图 2b), 在血清恢复后, 阳性信号消失, 但在细胞密集的地方, 还有阳性信号 (图 2c), 这些生长密集的细胞是多层生长细胞, 其生长状态与血清饥饿状态相当. 这一结果表明, 该序列是在无血清条件下表达, 血清培养时表达被抑制, 故命名为血清抑制基因, 因这类基因有很多个, 将该基因命名为血清抑制基因-1 (serum inhibit-1, Si-1).

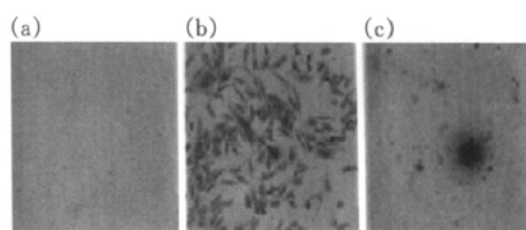


Fig. 2 The *in situ* hybridization of cells with Si-1 gene fragment

(a) cells cultured by 10% serum medium; (b) serum starvation cells; (c) recover serum cultured cells after serum starvation.

2.3 全序列的克隆及结构分析

将该 cDNA 片段进行重组克隆后, 测序, 有效序列长为 421 bp (位于图 3 序列中的 2 991~3 411), 但没有阅读框. 测序结果用 GenBank 数据库中的 Blast 程序搜索表达序列标签 (EST), 进行计算机步移拼接, 获得 5 kb 左右的 cDNA 序列. 以这一序列为基础, 设计 PCR 引物, 从血清饥饿的 U251 细胞来源的 cDNA 序列为模板, 分段扩增, 获得全长基因序列与人类基因组数据比较, 该基因定位在 2 号染色体, 全基因覆盖 50 kb 左右的基因组 DNA 序列, 共有 5 249 bp (图 3). 将这一序列用 GenBank 数据库中的 ORF 程序预测编码框, 编码框从 437 bp 到 2 812 bp, 共编码 791 个氨基酸残基 (图 3), 有 21 个外显子 (基因数据库搜寻结果未显示). 将 cDNA 序列进行同源性比较, 该序列除了与人的 DNA 序列有同源性外, 与其他已测定的模式生物 DNA 序列没有同源性. 氨基酸序列的同源性比较, 除人类序列外, 也未发现明显的与该序列同源的模式生物蛋白质序列 (结果未显示).

```

1  aggcgttaacgcgcacgcgcttagggatccggccgtggccgagcgcgcggccgtaagaccgcggtgactagcatgcagatacccatgct
91  ctgactttctgccctccactgacatggccacccgggggggagaggagacttccagacttcagctcagcgcagcagcagcagcagcagc
181  ttccagctttcagtgcatgaacgggagctggactgtttttctggatcatagctatgccaaaccttggagtgcccccacccagatgccagt
271  agtgcgcgcgcaccccgcatgctctttgtcactcccgccggcagcagcaagatcatatgattgttaaatacctgtgaagtatgtgtg
361  gcccggttctctatttctcactgtttacagtgaaatcagacgtcccaatagatgtggagacggtcacatcaacgcctatgccactcatga
M P L Y D
451  caatcagaagcagcagcagcgtgatgaatgagtgacggcgtatcttttccaggactgatgcagatgcccttccaccacagga
N Q K A R S V M N E C E R H V I F A R T D A D A P P P P E D
541  ctgggagagcagtgatcaacaggactggctggacaatggccagacaagctattcaacaagatccctaaagccctgagcttgcacggct
W E E H V N R T G W T M A Q N K L F N K I L K A L Q S D R L
631  tgcgcgtttggcacaacgaagggcttgaatgagcagctgcgcgctgttctgtggacaagtgctcaaggagagtgccgagcgtct
A R L A N E G A C N E P V L R R V A V D K C A R R V R Q A L
721  ggcaagtgtagctgggatacgaagctgacccagtgagctgcacacaccccttggagacgttgagtgccctcagcagcagcagc
A S V S W D T K L I Q W L H T T L V E T L S L P M L A A Y L
811  ggatgctttgcagacgtgaagggaagatcccaaccttgattgaccggatgctgtgtcatccaacaagactgggctcagggagc
D A L Q T L K G K I P T L I D R M L V S S N T K T G A A G A
901  tgaggccttctctctctactgaagagggcctgggacccctgtgtgggtgtgtttctcatacaaaacgaacaaactccctcgtctcc
E A L S L L K R P W D P A V G V L S H N K P S G S P
991  gtgatttctcatgctcctctctgttccctccagctctgttcttccacttcacgcccacccgcttctggcaatctcagcttctctgctt
L I L I A S S S S V F P T S R R H R F W Q L S C L
1081  ggcaagtgatccctgtagccacccatctgctgaacatggcagtgagggttctctacagtgctcagcagcagcagcagcagcagc
G K V I P V A T H L L N N G S G V G V L Q C L E H M I G A V
1171  gagaagcaagctcggagattcacagcatttccacacaaacccattatctgattgctggaacacagagccttgggtgctgtca
R S K V L E I H S H F P H K P I I L I G W N T G A L V A C H
1261  tgttcagtaattgagtgatgtcactgcagttgtctgttctctgttctgttctgttctgttctgttctgttctgttctgttctgttct
V S V M E Y V T A V C L G F P L L T V D G P R G D V D D P
1351  cctctgttgatgaagactcagctcctcttctcattgttcagaaatcccttcaatgtcaccctgaagcctgaagccttccggagaa
L L D M K T P V L F V I G Q N S L Q C H P E A M E A D F R E K
1441  gattcagctcagaacagcttgggtgtgttggggagctgatgacaatctcagaataagcaagcaagaatcagaagcgttgac
I R A E N S L V V V G G A D D N L R I S K A K K A S E G L T
1531  tcagagctgtggcagatgtattcaggatgagattgtgacttctgacttgagtgctcactcgtcgtcagggctcagctcagctcga
Q S M V D R C I Q D E I V D F L T G V L T R A E G H M G S E
1621  acctcgggacagcagctgctgagaagaagaagcccgccgctgtggccgcagagacttggccttgaagctccctgaagcggcgagc
P R D Q D A E K K K P R D V A R R D L A F E V P E R G S R
1711  acctcctccacagctgccagctgcccgctcaccctcagcctcagcagcagcagcagcagcagcagcagcagcagcagcagc
P A S P A A K L P A S P S G S E D L S S V S S S P T S S P K
1801  gaccaaaagtaccacagtgacctctgccagaagtcagctcagattggaagttctcagctgtgaagacatgtgcagcggacagaagc
T K V T T V T S A Q K S S Q I G S S Q L L K R H V Q R T E A
1891  tgtgtcagccacaacagctcaagttccatttcatcagaaccacagaggaagagagaagagagagagagagagagagagagagagag
V L T H K Q A Q V P I S S E P P E E G E K E D L R V Q L K R
1981  acacccctcagctccttctgtcagtaagacctccaaacgacgaagatcaagtgctccttctcctcagcagcagcagcagcagc
H H P S S P L P G S K T S K R P K I K V S L I S Q G D T A G
2071  agggccttctgtccttccaaagagtgctccagaagctcagcttgggaagcccatcaccatgacatggggcagcgttcagcagggc
G P C A P S Q G S A P E A A G G K P I T M T L G Q A S A G A
2161  caaggcctcagcagcttctcaccacagcgaagtcagcttctcaggttgagctcagcagcagcagcagcagcagcagcagcagc
K E L T G L L T T A K S S S E G G V S A S P V S V S S
2251  cagcactgcaccagctgcttgcacacactgcagagccgctgtgtgccacatctcctgcagctcctcagcagcagcagcagcagc
S T A P S A L H T L Q S R L V A T S P G S S L P G A T S A
2341  cagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc
S L L Q G L S F S L Q D I S S K T S G L P A N P S P G P A P
2431  acaggcagcagctgagtgagtgccaccccatgcagagcctgggtcccatcaccacagcagcagcagcagcagcagcagcagcagc
Q A T S V K L P T P M Q S L G A I T T G T S T I V R T I P V
2521  ggccacacactctctccttgggtgccaactctgtgggaagccacagccatccacagctgtgaccaatgggggctcgtcagctt
A T T S L G A T P G G K P T A I H Q L L T N G G L A K L
2611  ggcaagcagcctcctgctgctcagatctctaaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc
S S L P G L A Q I S N Q A S G L K V P T T I T L R G G
2701  gccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc
P S R I T T L S P M G S G A A P S E E S S S Q V L P S S Q
2791  gcgcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc
v L P P A P
2881  gtgtggtcctgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc
2971  aggttaggtgatacgggtgcagcaaggggaagcaccatcgtccaggatctgcaaatctgggtcctgggaacccagactcctcagcagatc
3061  tggctgtacatggatcagaaccacttcttccccgcttaagctgtgtgttgaccacagggcagcagatagggactgctgtgcatttaa
3151  tgaaggtgtttccttttgaagctgtgtcaccctctgcgcctagttgggagagacatccatctggtctgggatttcgggagttagaat
3241  ggaaagctctttgtctaaagactggagtcacctggcctgccaactgtgtgttcagagccggagcgggctgttttggacatcactgttgc
3331  ttactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc
3421  tctgtgttttgttagtatacataatgaaggcaagagagactgctcattcttccagccttctgacttactctagattgtggagcag
3511  tgggtgccagtgaaacctgggccccttcttcttccacatgcttgggcagttctgtatccagagagtcgcaggttcagataagct
3601  gaagaagagtaatagaacagcaaggaagtggttgaagatgtgtagtaagccctgtgtgttgtgtttaggtctctgtctgtctaccc
3691  aaggaaactgtgtgttcagctggagataaaaaagaagaatttgccaaagtcagagaagaaccccaaccccgaaatacctctgtctccagtc
3781  tctggaggtgaagcagggacaataagctaaagtagtatcttggccatccaggaacacttgtggcattaggacagatgaaggccatgttca
3871  gtgttttctgttctatctatgagacttttcttctcctgttacaagtgggaagatgattgacagtgactctactatgcaggcgtgttg
3961  gtaccaacctgagccctataggtggcagtcctcggagaagtggtcacagaagatggagctctgacccctgcttacccttccacaact
4051  tgtgtgcaagatagtttagatttggtttagaagctatctccagaacaggtcccatcactagaatgttcttagttaaggttaataaat
4141  taggcaaccccaagtgtagctcactcaagtgtcctttctgttaggcaggaagggcccaacatgtgcttaaaatgtagtcattgtgtgt
4231  gccacagtagcagtggtgtatctatcaccagtgacactgtgttcaatctgggagccttctggccagctgtagtggaaggggagct
4321  catagtgctcaggtgtcagggcaaggttaggtcaggttactgtggcaggggggagtggtgtgtaggagagaggtaccctaaacccata
4411  ccttccctcctgacctgaaagctgactcacaacagggttcacacagaattaggtgtgttttgcattagctgttaggtgactttctc
4501  aaattcttaattcagaaggtatttagtaacttgaggaaggtatgaattctggaggagcagcagcagcagcagcagcagcagcagc
4591  acaggttaggtatcaccacacaaagagcagtggaacaataagactatatgagctatatgaagcttttaggaatcatttaggacagacag
4681  agccctaaacacccattcatgacttaagttgtgtgtgtcagtgatgtgtgggacaaagaaaaactaacaagccgacctgcctttatgat
4771  aaattcttaattcagaaggtatttagtaacttgaggaaggtatgaattctggaggagcagcagcagcagcagcagcagcagcagc
4861  acgtggattgttctacaacttgacagggtaacagaggaagtggtcagggcctagagtcacgttttccagttcccttgcacaaatatt
4951  tcttggaaacgcgaaggaagctttacatttctatagaagacctggaatccataacctcagaaggcaatattttagtagaaaaatgtgga
5041  aggatcaggaagttcttagattcttggatgacagatgcatgttgatccctatggagatgtccttgtgttttaggtcactgaggttagga
5131  agacctgtctactcttggtttccacactagaacagcttctgggctggatgggttatagagctgagcggctgtgtgtgttctgttttcat
5221  taacaaaaacaattaaaaacacaaaaaac

```

Fig. 3 The full cDNA of Sf-1 gene

The sequence in box was get by DD-PCR

3 讨 论

根据人类基因组草图资料分析, 本研究克隆的基因位于第二号染色体, 分子标记在 Flj10081 处, 序列被包含在人类测序工程的 NT_26972.1 序列之中. 根据染色体序列定位于 89904K 至 89949.5K 之间, 覆盖 45.5K 的 DNA 序列, 由于人类基因组草图序列存在一些 gap 序列, 不能准确地确定其 DNA 序列的长度. 但根据 Blast 搜寻, Sfr1 基因序列位于基因数据库中另一序列 AC-079754 RP11-67L23 内, 该序列是一个完整的 contig (人类基因组测序时, 几个序列或克隆通过交替重叠而拼接起来的一段序列), Sfr1 基因位于该序列的 24 847~ 70 019 bp 之间, 由此推算该基因占据 DNA 序列为 45.173 kb. 两个 contig 序列资料的分析可以确定该序列在基因组中的长度为 45 kb 左右.

本序列除了在人类基因数据库中有同源序列外, 在其他模式生物中没有找到核苷酸序列的同源序列, 蛋白质序列的同源性比较, 只发现 N 端的 390 氨基酸残基与果蝇中 2R 染色体上 CG8233 基因 (AE003815) 的 N 端序列有 38% 的同源性, $\text{expect} = 1e^{-67}$; 164~ 376 之间有 213 个氨基酸残基与拟南芥中一个基因有 30% 的同源性, $\text{expect} = 2e^{-14}$; 222~ 369 之间有 148 个残基与线虫的一个推测蛋白有 27% 的同源性, $\text{expect} = 9e^{-11}$. 由此显示, 该基因可能是人类特有的基因.

本实验室对 p16, p21 和 p53 基因在血清培养

和血清饥饿细胞中的表达进行过研究^[4], 发现这几个基因在血清饥饿细胞中较血清培养细胞中表达较强, Sfr1 基因也有同样的表达特征, 表明该基因也有可能是一个细胞生长抑制基因, 甚至是肿瘤抑制基因. 这有待进一步研究.

本序列在 GenBank 中的登记号为 AY050169.

参 考 文 献

- 1 Gupta S, Bose A, Chatterjee N, *et al.* p67 Transcription regulates translation in serum-starved and mitogen-activated KRC-7. *Cell*, 1997, **272** (19): 12699~ 12704
- 2 Helbing C C, Veillette C, Riabowol K, *et al.* A novel candidate tumor suppressor, ING1, is involved in the regulation of apoptosis. *Cancer Res*, 1997, **57** (7): 1255~ 1258
- 3 Ahn J H, Park S M, Cho H S, *et al.* Non-apoptotic signaling pathways activated by soluble fas ligand in serum-starved human fibroblasts. Mitogen-activated protein kinases and NF- κ B-dependent gene expression. *J Biol Chem*, 2001, **276**: 47100~ 47106
- 4 罗兰, 何云刚, 舒昆贤, 等. 血清对胶质瘤细胞株 U251 基因的表达式的影响. 云南大学学报 (自然科学版), 2000, **22** (2): 144~ 147
Luo L, He Y G, Shu K X, *et al.* Journal of Yunnan University (Natural Science), 2000, **22** (2): 144~ 147
- 5 Liang P, Pardee A B. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, 1992, **257** (5072): 967~ 971
- 6 Liang P, Averboukh L, Keyomarsi K, *et al.* Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelial cells. *Cancer Res*, 1992, **52** (24): 6966~ 6968
- 7 Andreas D B, Ouellette B F F. 李衍达, 孙之荣, 等译. 生物信息学. 基因和蛋白分析的实用指南. 北京: 清华大学出版社, 2000. 146~ 159
Andreas D B, Ouellette B F F. translated by Li Y D, *et al.* Bioinformatics: A Practical Guide to the Analysis of Gene and Protein. Beijing: Press of Tsinghua University, 2000. 146~ 159

Cloning of a Gene Inhibited by Serum Medium^{*}

TAN De-Yong^{**}, LAI Jian-Hua, Qian Wei, Yu Min, CHEN Yao, HE Jie

(Department of Biotechnology, Yunnan University, Kunming 650091, China)

Abstracts Comparing the gene expression difference between normal cultured cells and serum starvation cells with DD-PCR method, a cDNA fragment, which expresses increase under serum starvation, was obtained. According to this cDNA fragment, the full gene cDNA was obtained primarily by searching and splicing the EST sequences, then was cloned by PCR and sequenced. The size of this gene is 5 429 bp and it codes a peptide containing 791 amino acid residues. It is not homologous with known genes that related to cell cycle. So, this gene may be a new cell cycle-related gene (GenBank: AY050169). It was named as serum inhibition gene (Sfr1) according to its expression inhibited under serum existing.

Key words serum starvation, gene, clone

^{*} This work was supported by grants from The National Natural Sciences Foundation of China (39960030) and The Applied Basic Research Programs of Yunnan Province (1999C002Z).

^{**} Corresponding author. Tel: 86-871-5032061, E-mail: dytan3@public.km.yn.cn

Received: March 27, 2002 Accepted: May 28, 2002