

\* 学术论文 \*

小鼠脑特异表达基因 *Bsg4* 的克隆及表达 \*张 晶<sup>1</sup> 王 峦<sup>1,3</sup> 陈 微<sup>2</sup> 薛 鹏<sup>1</sup> 陈 荣<sup>3</sup> 慈宏亮<sup>1,3</sup> 李亦平<sup>1,2,3\*\*</sup>1. 北京师范大学生命科学院生物医学研究所, 北京 100875; 2. The Forsyth Institute, Boston 021152 USA;  
3. 浙江赛尔生物医学研究院, 杭州 310006

**摘要** 利用消减差异筛选的方法获得在小鼠头部特异表达的基因 *Bsg4* (Brain Specific Gene 4), 基因全长 4308bp. 应用生物信息学方法分析基因结构, 表明该基因定位于小鼠的第 3 号染色体, 包含 26 个外显子, 25 个内含子, 编码区长 3639bp. 用 RT-PCR 方法扩增编码区, 并构建到 pBlue-script SK(-)载体上, 然后采用原位杂交技术研究了 *Bsg4* 基因在小鼠胚胎发育中的表达情况. 以 *Bsg4* 基因编码区为探针的原位杂交结果显示 *Bsg4* 基因在小鼠胚胎头部和成年小鼠脑的海马中有特异表达, 显示 *Bsg4* 基因与头部发育和海马的功能有密切关系. 对 *Bsg4* 基因的时间和空间表达模式的研究将有助于我们进一步深入地揭示它在脑发育中的作用.

**关键词** 脑特异基因 生物信息学 原位杂交

脑发育过程是一个由多基因参与调控的复杂过程. 许多基因仅在脑中特异表达, 参与神经系统的形成或调控. 如 *Hox*, *PAX*, *LIM*, *POU*, *Emx*, *Dlx* 等同源盒基因家族<sup>[1]</sup>, 都参与了脑和神经系统的发育. 在鼠中 *Hox* 家族中的 *Hoxc-8*, *Hoxb-4*, *Hoxa-2*, *Hoxd-3*, *Hoxd-13* 等与神经系统的形成有关. 鼠的 *Pax1-Pax9* 基因在发育过程中参与诱导特殊细胞形式的分化以及神经系统的形成. *LIM* 家族同源盒基因可能参与有关神经元细胞的生命过程和神经发生. 缺失 *lim-1*, *otx-2* 会使小鼠不产生前脑; 而 *Caspase-3* 的缺失会造成小鼠前脑细胞数目的增生<sup>[2]</sup>. 基因 *cerberus* 编码一个可分泌蛋白, 其 mRNA 注入卵母细胞可诱导脑结构发生异位<sup>[3]</sup>. *En-1*, 2 在中脑的发生中起重要作用. 缺失 *En-1*, 2 基因会抑制小鼠顶盖和后脑的前半部分的形成. *En-1*, 2 的表达也是中脑出现极性的标志. *Pax* 也与中脑的发生有关, *Pax-5* 缺失的小鼠, 其脑峡不能发生.

一些基因表达的异常也与多种脑疾病发生有着密切的关系, 如 *ApoE-epsilon4*<sup>[4]</sup>, *pADHC-9*<sup>[5]</sup> 等许多基因与老年痴呆症相关. 因此研究脑的特异基因将为我们了解脑的早期发育、神经系统的构建奠定基础.

## 1 材料和方法

### 1.1 材料

ICR 小鼠购自北京大学医学部实验动物中心. 逆转录酶、消减试剂盒和杂交缓冲液购自 Invitrogen. Sephadex G-50 购自 Amersham Biosciences, QI-Aquick Gel Extraction Kit 购自 QIAGEN 公司, 小鼠胚胎的 cDNA 文库购自 Clontech, pBluescript SK(-) 载体购自 Stratagene, *Sal I* 和 *Kpn I* 购自 New England Biolabs, 高保真的 DNA 聚合酶购自宝生物工程大连有限公司. 地高辛标记与检测试剂盒购自 Roche 公司. 其他试剂为国产分析纯试剂.

2004-04-20 收稿, 2004-06-09 收修改稿

\* 浙江赛尔生物医学研究院资助项目

\*\* 通讯作者, E-mail: YPLi@Forsyth.org

## 1.2 方法

**1.2.1 消减差异筛选脑特异基因** 收集 9.5 d 小鼠胚胎的头部, 用常规方法制备 mRNA. 取 5 $\mu$ g mRNA 为模板, 逆转录采用随机引物, 并掺入 <sup>32</sup>P-dCTP, 42 $^{\circ}$ C 反应 1h; 0.1 mol/L NaOH 重悬反应产物, 65 $^{\circ}$ C 孵育 20 min 使 RNA 模板水解; 0.1 mol/L 醋酸中和探针, 探针用 Sephadex G-50 纯化. 生物素标记的小鼠 9.5 d 胚胎躯干的 mRNA 使用消减试剂盒制备, 实验步骤按其说明书进行.

将 <sup>32</sup>P 标记的 cDNA 探针与 10 倍过量的生物素标记的 mRNA 进行杂交. 离心沉淀探针混合物并用 20 $\mu$ L 水重悬, 100 $^{\circ}$ C 热变性 1 min. 加入等体积 2 $\times$  杂交缓冲液, 65 $^{\circ}$ C 杂交 20~24 h. 加入等体积 HEPES 缓冲液 (10 mmol/L HEPES, pH 7.5, 1 mmol/L EDTA), 20 $\mu$ g 链霉亲和素, 冰上孵育 10 min. 生物素化的 RNA 和 RNA-cDNA 与链霉亲和素形成的复合物用酚-氯仿抽提除去, 水相中的 cDNA 探针用乙醇沉淀法纯化, 在相同条件下进行第二轮消减杂交.

我们用 7.5 d 小鼠胚胎的 cDNA 文库中的 5 $\times$  10<sup>5</sup> 个克隆进行筛选. 每个 145 mm 的平板上约 10<sup>4</sup> 个克隆, 每板制两张膜用于与第二轮消减杂交得到的 cDNA 探针杂交. 杂交、洗膜和曝光将按照常规的方法进行<sup>[6]</sup>. 为纯化克隆, 还要在低浓度的条件下进行第二轮杂交. 所得到的克隆进行测序验证.

**1.2.2 *Bsg4* 基因序列特征分析** 将所得序列在 NCBI (www.ncbi.nlm.nih.gov) 上用 BLAST (basic local alignment search tool) 方法与已知序列进行比对, 以确证所得序列为一新序列. 用 ORF (open reading frame) finder 预测该基因的编码框. 使用 Mouse Genome 数据库做 BLAST 分析, 找到 *Bsg4* 基因在小鼠基因组上的定位及分布, 对其外显子和内含子进行分析. 使用 Vector NTI 8.0 软件对小鼠的 *Bsg4* 基因与人的 KIAA0476<sup>[7]</sup> 基因的氨基酸序列进行同源性比较. 用 NCBI 数据库的 CCD (conserved domain database) 对蛋白的结构域和可能的功能进行预测.

**1.2.3 RT-PCR 克隆 *Bsg4* 的全长编码区序列** 根据 *Bsg4* 的 cDNA 序列用软件设计一对引物, 在 5' 端引物加入 *Sal* I 酶切位点, 在 3' 端引物加入 *Kpn* I 酶切位点. 两引物序列为:

P1: 5' TCCGGTCGACTG AACCTATTACAGA

TGTGGCA 3'

P2: 5' ATTAGGTACCGCATTCCAGAGGTG-CTCCAAA 3'.

提取小鼠胚胎的总 RNA 为模板, 经 M-MLV 逆转录酶合成 cDNA 第一条链. 使用高保真的 DNA 聚合酶扩增 *Bsg4* cDNA 序列. 采用 100 $\mu$ L 的 PCR 反应体系, 反应条件为: 94 $^{\circ}$ C, 1 min; 55 $^{\circ}$ C, 1 min; 72 $^{\circ}$ C, 2 min; 经过 30 次循环后, 于 72 $^{\circ}$ C 延伸 7 min. PCR 产物用 1% 琼脂糖凝胶进行分离, 并用 QIA quick Gel Extraction Kit 纯化和回收后将 PCR 产物克隆至 pBluescript SK(-) 载体上. 连接产物转化感受态大肠杆菌 DH5 $\alpha$ , 筛选重组子, 制备质粒 DNA, 经测序验证为 *Bsg4* 全长编码区序列.

**1.2.4 原位杂交** 按照参考文献[8]中的方法进行原位杂交. 收集 8.5, 9.5 d 小鼠胚胎, 用 4% 多聚甲醛 4 $^{\circ}$ C 固定过夜. 之后用 PBS 洗涤胚胎, 甲醇系列脱水 (25%, 50%, 75%, 100%), 甲醇洗涤 2 次, 再反向系列甲醇进行水合. 依次用 PBS 洗涤胚胎, 6% 的 H2O2 胚胎漂白 1h, PBS 洗涤胚胎, 10 $\mu$ g/mL 蛋白酶 K 处理 15 min, 2 mg/mL 甘氨酸处理 5 min, 0.2% 戊二醛/4% 多聚甲醛再固定 20 min, 60 $^{\circ}$ C 预杂交 8 h. 成年小鼠脑切片用常规方法进行石蜡包埋和切片, 二甲苯脱蜡, 梯度乙醇进行水合, 经 0.2 mol/L HCl 和蛋白酶 K (15 $\mu$ g/mL) 处理 4% 多聚甲醛固定, 室温预杂交 2 h.

以 3.6 kb 的小鼠 *Bsg4* 编码区的 cDNA 为模板, 根据 DIG Nucleic Acid Labeling Kit 制备 DIG 标记的 RNA 探针, 并把探针加入预杂交液, 与胚胎和切片在 60~65 $^{\circ}$ C 下进行杂交. 杂交过夜. 杂交后经多次洗涤和 RNase 处理后, 加入 BCIP/NBT 显色. 小鼠胚胎用 4% 多聚甲醛固定, 脑切片用 50%, 70%, 95%, 100% 浓度梯度乙醇进行脱水. 二甲苯处理后, 树脂封片, 显微镜下拍照.

## 2 结果

### 2.1 消减差异筛选脑特异基因

利用消减差异筛选的方法对 7.5 d 小鼠胚胎的 cDNA 文库进行筛选, 克隆到脑特异表达的基因 *Bsg4* (GenBank 登录号: AY512564). 测序结果显示, *Bsg4* 长 4308 bp. 把 *Bsg4* 在 NCBI 上进行 BLAST 分析, 结果证实 *Bsg4* 是一新基因.

## 2.2 *Bsg4* 序列的生物信息学分析

用 ORF finder 对 *Bsg4* 基因的编码框进行预测, 得到一个 3639 bp 长的编码区序列, 位于 *Bsg4* cDNA 的 607 ~ 4245 核苷酸位置. 含有起始密码子 ATG 和终止密码子 TAG, 是一个完整的编码框 (图 1). 在 Mouse Genome 数据库中做 BLAST 的分析时将 *Bsg4* 定位在小鼠的第 3 号染色体上, 包含

26 个外显子, 25 个内含子 (图 2). 小鼠 *Bsg4* 基因和人 KIAA0476 基因的氨基酸序列的同源性比较结果显示两者的同源性为 81.2% (图 3). 对 *Bsg4* 基因结构域的分析显示该基因从 N 末端 39 ~ 223 氨基酸含有一个 DENN 结构域, 287 ~ 361 氨基酸为一个 dDENN 结构域. 这些结构域提示 *Bsg4* 基因可能具有调节信号传递的作用 (图 4).

```

1      tggtggaattactttgtggctgggcttgcaggaatggagcgcctata
51     cctgaggagaagtgggttcccgagcctaccgggcccctgcgccctcccc
101    gccagctgaacctattacagatgtggcagtcacctcctaggcactggggc
151    aggaggtgccccaggcctatacatgcattcagactctcgggtggccac
201    cccttggagctcagtgctgggctcctgggtggaactcaacagtcactctg
251    ctaccgagggggcgtgacaagccccctcgttgagctgggggtttat
301    atgaagggaaggacgaccaagccttgggttccaagcttgacacgaca
351    ccgtatagccactcagccaatctggccctccagcccgggaccccccc
401    cactacctcatgtatcggcggcagcagaaggggcagggtcgtcatgcc
451    tgggcatcactgacctctgctggttctgcccagcaagggtgagggtact
501    cctcactactgcccgtaccocgcaacctcaacctggcatgtgggg
551    cccagcagtgacctgtgctacaaggtaggcctggccaagcccaacac
601    tgggtatgaggcagttctatgaagcgttcccaggggccaggctgcaga
      M R Q F Y E A F P R A R L S E
651    gaggcaggcacggcactggcctgatgagtgccgtggaacggggcggg
      R Q A R A L G L M S A V E R G R
701    cgtgggagggcagctgtgcccagccgacgagccatagctgtgctgccc
      A L G G R A V R S R R A I A V L S
751    cgtggcctgcccctccagccttccgagccttctcacccttcttaccg
      R W P A F P A F R A F L T F L Y R
801    atactcgtctcaggccccatcgtctgccccttggagcgcacatctccc
      Y S V S G P H R L P L E A H I S
851    acttcattcacaacctgccattcccacccacagagaccacgcatccta
      H F I H N V P F P S P Q R P R I L
901    gtgcagatgtctccctatgacaacctgctcctctgcaacctgtctctc
      V Q M S P Y D N L L L L C Q P V S S
951    acccctgcccctcagtggtgctagctttttgcagctgctgcagaacctg
      P L P L S G A S F L Q L L Q N L
1001   gcccagaactggctattacattgctgctgctgctcaccagagcacaaa
      G P E L A I T L L L A V L T E H K
1051   ctgctggtccactcgtcggccagatctgctcaccagcgtgtgcaaggc
      L L V H S L R P D L L L T S V C E A
1101   ccttctcctgatgatcttccctctgcatggcagtgccctacatcccct
      L V S M I F P L H W Q C P Y I P
1151   tgtgcccgctagtgctggcggatgtgctgagtgccccctgccccttcat
      L C P L V L A D V L S A P V P P F I
1201   gtgggtatccactccagttatttcgatctgcatgacccacctgctgact
      V G I H S S Y F D L H D P P A D V
1251   catctgtgttgatcttgataccaacagcctcttccaggagaaaagaagc
      I C V D L D T N T L F Q E E K K
1301   cctctccgctaggaccctgcccgaagaccatacaagctttttattggcc
      P L S A R T L P R R P Y K L L L L A
1351   accctgacaagtctgtaccagcagctggaccagacttacctggaccaga
      T L T S L Y Q Q L D Q T Y T G P E
1401   ggaagggcctccttggaaattcctgctgacagactatgagcctgtgtgtg
      E E A S L E F L L T O Y E A V G
1451   gcccgaggactcgcctggaacgggaagtccaggagcccttctcgccttc
      G R R T R L E R E V Q G A F L R F
1501   atggcctgtctgctcaagggttaccggaacttccctcctcctcaccgca
      M A C L L K G Y R N F L R P L T Q
1551   ggccccctctgaggggtctcagatgttgacaaccttttttaccctcagg
      A P S E G S R D V D N L F Y L Q
1601   gtttctcaagtcccgggaacgctcaagccacaagctgtactctcagctg
      G F L K S R E R S S H K L Y S Q L
1651   ctgcacacagatgcttccagacttccaggaatgctcttttggctc
      L H T Q M F S Q F I E E C S F G S
1701   cggccggcagcagccctggaattctttgactcttgtgttgacaaggttc
      A R H A A L E F F D S C V D K V
1751   acccagcaggaggaagccgaaccacaccttagtggagctggaggaa
      H P E Q E K P E P T P L V E L E E
1801   ctatcagggaagtgagctcactgtctttatcacacccccagggaccacc
      L S G S E L T V F I T P P E E P P
1851   agtactggaagcagtgagctactcctcagctactgttatgatgggtttc
      V L E G S E S T P Q Y G Y D G F
1901   cagagctaaggctgagctatttggctcctcaggacaacagggggcc
      P E L K A E L F E S P Q E Q Q G A
1951   ctccctgtgccaggccatcccggagtgctcccagcagctcctgacctgc
      L P V P G P S R S A P S S P A P R
2001   ccgaaccaaacaggagatgaaggttgacagcgaatggcacagaagttag
      R T K Q E M K V A Q R M A Q K S
2051   ccactgtccctgagctgtgggcccgggtgctgctgggtcactgctacgga
      A T V P E L W A R C L L G H C Y G
    
```

```

2101  cttgtgttcttatgacctgcccgcctacgtggttccgtaccctcccgggt
    L W F L G L P A Y V R S V P S R V
2151  gcgtgattacataccgctaccacgtgttacgggaaatggagaaccgca
    R A L H T A Y H V L R E M E N R
2201  aattgttactcctgatgaggtgtgtaccgctgctgatgacgtctgc
    K V V L P O E V C Y R V L M Q L C
2251  tcacactatggcgaccgcttctgtctgtgcgagtcagctagagatcg
    S H Y G Q P V L S V R V M L E M R
2301  accggcagcattgtgcccaacaccatcactacggttactacaacaagg
    R A G I V P N T I T Y G Y Y N K
2351  cagtgtggaagcaaatggccatctggtactccgggtggcagcctacgc
    A V L E S K W P S G T P G G R L R
2401  tggccaagctccgcaatgtttgctcctgggggctgctcagttccgcaacc
    W A K L R N V V L G A A Q F R Q P
2451  cttaaaggaccgaaggcagcaacagcagcagcagcagcagcagcagcagc
    L K D R R R Q Q Q Q Q Q Q Q Q Q Q
2501  agaagcaacagcagaccctatctggagcgtccctcccacacgcccct
    Q K Q Q T E P Y L E R P S P T R P
2551  cttcagcccagactacttggctggcgaagtctacgggaaccatcttc
    L Q R Q T T W A G R S L R E P S S
2601  acccatggggccctggtcaagagtgtagcctaggcagtgccctggga
    P M B R L V K S G S L G S A R G
2651  cacagccacagtgaggctgtgtgctcacatgatagagcctctgggg
    T G P T V E A G V A H M I E A L G
2701  gtccctggaaccccggatcacctgtgcttggcaggaccggaagcctctc
    V L E P R G S P V P W Q D G S L S
2751  agacctgagctgacaggggaaagaaatggaccctggaggcagccctggg
    D L S L T G E E M A P G G S P G
2801  gctcaggtccgctgagtgccagtcactgagggccctagaaggata
    G S G S A L S A Q S T E A L E G I
2851  agtggcgggttccaaactagtgagatcaggagaggtgggcaactcc
    S G R G S K T S G C Q E E V G T P
2901  cagaaaaggctgggtgcccctccaacagctgctcacccttcccggc
    R K G L G A R L Q Q L L T P S R
2951  gcgctctgcttccgcctccccacctgagtgccctctgacctcctcct
    R A S A S R I P P P E L P S D L P
3001  cctgcccggccgaagctccatgtagccttctgtggcccggagaacg
    P A A R R S P M D S L L W P R E R
3051  ccccgatcactgcttctgagagctctctctctggcagtgagtggg
    P G S T A S E S S A S L G S E W
3101  atatctcagagcttccctcagcagcctgagcttccgcttccctcagag
    D I S E S S L S S L S L R R S S E
3151  cgtctcagtgacaccctggggctttccagccacctctctggaattct
    R L S D T P G A F Q P P S L E I L
3201  gatgtctagctgttctgtgtctatgctgctgattccctgtgtacagk
    M S S C S L C H A C D S L V Y D
3251  aggaatcactgctgctggcaccgagactctaaccttaatacaact
    E E I M A G W A P D D S N L N T T
3301  tgtcccttctgcccctgctcactttgtcccctgctcagtgccagactct
    C P F C A C H F V P L L S V Q T L
3351  tgattctgaccagtgaccctagtcacaagcttcccttgcggcgcca
    D S R P S A P S P K S S L A G A
3401  gtgctgcaagatgctcctgcccctggggccctggcctgtgctcagt
    S G C K D A P A P G G P G P V L S
3451  gaccggagttctgcttccctggatcagcccagctctgcaatggca
    D R R F C L A L D Q P O L C N G H
3501  tatgggagtgcttcccggcagttgagaatggggcatgggctaacgta
    M G S A S R R R V E N G A W A Y L
3551  gccctctggtctgctgtaaggaactggagctcctggtaggaatgaggc
    S P L V L R K E L E S L V E N E G
3601  agcagggtgctggcttacctgaactgctgctgcccaccctatcattt
    S E V L A L P E L P A A H P I I F
3651  ctggaaacctctgtgtatttccaacggctacgcctgctcagtggtctac
    W N L L W Y F Q R L R L P S V L
3701  caggcctgtgttggcctcctgtaatggcccaccagccagctctcc
    P G L V L A S C N G P P P S Q L S
3751  caggaccatctccatggtaaccctgatccagctctctgctatgca
    Q G P S P W L T P D P A S V H V H
3801  tctgctgggatgctcctgaccctgatcccaacagctgcccacctctct
    L L W D V L T P D P N S C P P L
3851  acgtgcttgagggtccacagccagatcccaacagcgggtgtgtgcca
    Y V L W R V H S Q I P Q R V V W P
3901  gcccagtcacctcagcttagcttggcattactgagtcagtgctgg
    G P V P S C L S L A L L E S V L R
3951  ccagctcggactcaacgagttcataaagctgttggccttctctggaga
    H V G L N E V H K A V G L L L E
4001  cgctaggcccccctcccaccgctgaccctacagaggggcaatcagct
    T L G P P P T G L H L Q R G I Y R
4051  gagatcttattcctgaccatggctgctctggcgaaggaccctggat
    E I L F L T M A A L G K D H V D I
4101  agtggccttcgacaagaagtacaagctgctttaaacaagctggccagca
    V A F D K K Y K S A F N K L A S
4151  gcattggcaaggagaaactgaggcagcggcggcagacagatgccacctcc
    S M B K E E L R Q R R A Q M P T P
4201  aaggccatcgaactgccaaggtttttggacacctctggaatgctaggg
    K A I D C R K C F B A P L E C *
4251  acccttaagctcccctcaccgcttgagtgagaggtgaggagaggtg
4301  attctaga

```

图 1 *Bsg4* cDNA 全序列及预测的氨基酸序列

起始密码子(ATG)和终止密码子(TAG)用下划线标出

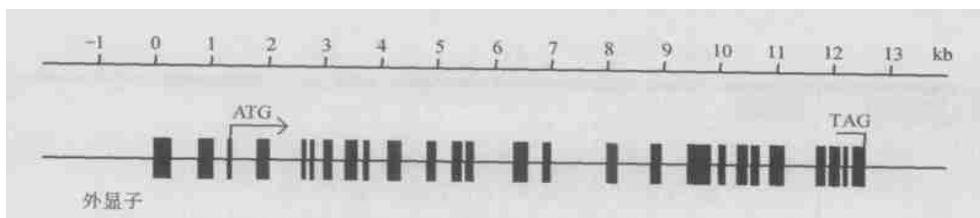
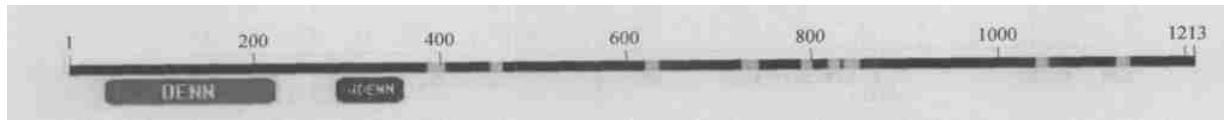


图 2 小鼠 *Bsg4* 基因在染色体上的分布  
26 个外显子和 25 个内含子分布在染色体上 13 kb 的范围。  
实体方块部分代表外显子。起始密码子 ATG，终止密码子 TAG

KIAA0476	(1)	1	MNQSLSGHWARKCPRAHARLLLGATPWNVSLDFWVSLNPSSATAGAVTSPPSLSWGCNMRGRVVPVSLASKCLTRHPATQGTWPLQAPGTPAFTSLTG	100
<i>Bsg4</i>	(1)			
Consensus	(1)			
KIAA0476	(101)	101	QQRGGGCMFWASLTSAWCCPYRALLILTAGCPATSTLACGAQDCTCAIRWWRPRTNWTDFYEAFFPRARLSEQRARLGLLSAVERGRALGGRAV	200
<i>Bsg4</i>	(1)		MRDFYEAFFPRARLSEQRARLGLMSAVERGRALGGRAV	
Consensus	(101)		RFYEAFFPRARLSEQRARLGLLSAVERGRALGGRAV	
KIAA0476	(201)	201	RSRRAIAVLSRWPAFAFRAFLTFLYRYSVSGPHRLPLEAHSIFHNVPFSPQRPRIIVQMSPYDNILLQQPVSSPLPLSGASFLLQLLQSPLELAIT	300
<i>Bsg4</i>	(39)		RSRRAIAVLSRWPAFAFRAFLTFLYRYSVSGPHRLPLEAHSIFHNVPFSPQRPRIIVQMSPYDNILLQQPVSSPLPLSGASFLLQLLQSPLELAIT	
Consensus	(201)		RSRRAIAVLSRWPAFAFRAFLTFLYRYSVSGPHRLPLEAHSIFHNVPFSPQRPRIIVQMSPYDNILLQQPVSSPLPLSGASFLLQLLQSPLELAIT	
KIAA0476	(301)	301	LLLAVLTHEKLLVHSLRPDLLTSVCEALVSMIFPLHWQCPYIPLQPLVLAADVLSAPVVFIVGIIHSSYFDLHDPADVIQVLDLNTLFTQTEKKLLSPRT	400
<i>Bsg4</i>	(139)		LLLAVLTHEKLLVHSLRPDLLTSVCEALVSMIFPLHWQCPYIPLQPLVLAADVLSAPVVFIVGIIHSSYFDLHDPADVIQVLDLNTLFTQTEKKLLSPRT	
Consensus	(301)		LLLAVLTHEKLLVHSLRPDLLTSVCEALVSMIFPLHWQCPYIPLQPLVLAADVLSAPVVFIVGIIHSSYFDLHDPADVIQVLDLNTLFTQTEKKLLSPRT	
KIAA0476	(401)	401	LPRRPYKLLATLTLNLYQQLDQTYTQPEEASLEFLLDYEAVQGRRLREREVOGAFLRFMAQLLKGYNRFLRPLTOAPSEGARDVDMFLQGLKSR	500
<i>Bsg4</i>	(238)		LPRRPYKLLATLTLNLYQQLDQTYTQPEEASLEFLLDYEAVQGRRLREREVOGAFLRFMAQLLKGYNRFLRPLTOAPSEGARDVDMFLQGLKSR	
Consensus	(401)		LPRRPYKLLATLTLNLYQQLDQTYTQPEEASLEFLLDYEAVQGRRLREREVOGAFLRFMAQLLKGYNRFLRPLTOAPSEGARDVDMFLQGLKSR	
KIAA0476	(501)	501	ERSHKLVSQLLHTQMFQSFIEECFSGARHAALFFDSCVQKHPDEKPEPTPLVELELQSGELTVFITPPEEPALPEGESTPQYCYDGFPELRAE	600
<i>Bsg4</i>	(336)		ERSHKLVSQLLHTQMFQSFIEECFSGARHAALFFDSCVQKHPDEKPEPTPLVELELQSGELTVFITPPEEPALPEGESTPQYCYDGFPELRAE	
Consensus	(501)		ERSHKLVSQLLHTQMFQSFIEECFSGARHAALFFDSCVQKHPDEKPEPTPLVELELQSGELTVFITPPEEPALPEGESTPQYCYDGFPELRAE	
KIAA0476	(601)	601	LFESLQEGPGLVPGPSRSAPSPAPRRTKQEMKVAQRMAQKSAVPELWAROLLGHOYGLWFLCLPAYVRSAPSRVQALHTAYHVLROMESGKVVLPD	700
<i>Bsg4</i>	(436)		LFESLQEGPGLVPGPSRSAPSPAPRRTKQEMKVAQRMAQKSAVPELWAROLLGHOYGLWFLCLPAYVRSAPSRVQALHTAYHVLROMESGKVVLPD	
Consensus	(601)		LFESLQEGPGLVPGPSRSAPSPAPRRTKQEMKVAQRMAQKSAVPELWAROLLGHOYGLWFLCLPAYVRSAPSRVQALHTAYHVLROMESGKVVLPD	
KIAA0476	(701)	701	EVCYRVLMLQCSHYQDPVLSVRVMLEMRAGIVPNTITYGYNKAVLESKWPSTGGRRLRWAKLRNVVLGAAGFRQPLKDRR0000000000000000	800
<i>Bsg4</i>	(538)		EVCYRVLMLQCSHYQDPVLSVRVMLEMRAGIVPNTITYGYNKAVLESKWPSTGGRRLRWAKLRNVVLGAAGFRQPLKDRR0000000000000000	
Consensus	(701)		EVCYRVLMLQCSHYQDPVLSVRVMLEMRAGIVPNTITYGYNKAVLESKWPSTGGRRLRWAKLRNVVLGAAGFRQPLKDRR0000000000000000	
KIAA0476	(801)	801	EDVSAHQEGSSQAEPYLERPSPTRPLQRQTTWAGRSLRDPASPPGRVYKSGSLDSARGOPTVEAGVAHMI EALGVLEPRGSPVPMHDSLSLSTGE	900
<i>Bsg4</i>	(636)		TEPYLERPSPTRPLQRQTTWAGRSLRDPASPPGRVYKSGSLDSARGOPTVEAGVAHMI EALGVLEPRGSPVPMHDSLSLSTGE	
Consensus	(801)		EPYLERPSPTRPLQRQTTWAGRSLRDPASPPGRVYKSGSLDSARGOPTVEAGVAHMI EALGVLEPRGSPVPMHDSLSLSTGE	
KIAA0476	(901)	901	EPLPQSGSGSALSAGSTEALQSLDQPKAQRQDEAGTPRRGLGARLQQLTPSRHSPASRIPPELPPDPPARRSPMDSLLHPRERPGSTASE	1000
<i>Bsg4</i>	(723)		EMAPGSGSGSALSAGSTEALQSLDQPKAQRQDEAGTPRRGLGARLQQLTPSRHSPASRIPPELPPDPPARRSPMDSLLHPRERPGSTASE	
Consensus	(901)		EPLPQSGSGSALSAGSTEALQSLDQPKAQRQDEAGTPRRGLGARLQQLTPSRHSPASRIPPELPPDPPARRSPMDSLLHPRERPGSTASE	
KIAA0476	(1001)	1001	SSASLGSSENDLSESSLNLSLRRSSERLSDTPGAFQPPSLEILLSSCSLACRACDLSYDEEIMAGWAPDSDNLTTCPFACGHPVLLSVQTLDSRPSVP	1100
<i>Bsg4</i>	(823)		SSASLGSSENDLSESSLNLSLRRSSERLSDTPGAFQPPSLEILLSSCSLACRACDLSYDEEIMAGWAPDSDNLTTCPFACGHPVLLSVQTLDSRPSVP	
Consensus	(1001)		SSASLGSSENDLSESSLNLSLRRSSERLSDTPGAFQPPSLEILLSSCSLACRACDLSYDEEIMAGWAPDSDNLTTCPFACGHPVLLSVQTLDSRPSVP	
KIAA0476	(1101)	1101	SPKS--AGASGKADAPVGGPGVPLSDRRCLALDEPQLONGHMGGASRRVSEGAWAYLSPLVRLKELESLEVENEGSEVLALPELPAHP IIFNLLIYF	1200
<i>Bsg4</i>	(923)		SPKSLAGASGKADAPVGGPGVPLSDRRCLALDEPQLONGHMGGASRRVSEGAWAYLSPLVRLKELESLEVENEGSEVLALPELPAHP IIFNLLIYF	
Consensus	(1101)		SPKS AGASG KADAP VGGPGVPLSDRR CLALD PQLONGHMG ASRRVE GAWAYLSPLVRLKELESLEVENEGSEVLALPELPAHP IIFNLLIYF	
KIAA0476	(1199)	1199	QRLRLPSILPGLVLAASDDP--SHSQAPSPWLTDPDASVQVRLLDVLLTPDNPSCPLVYLWRVHSQIPQRVWVPGVPVPSLALLLESVLRHVGLNEV	1300
<i>Bsg4</i>	(1023)		QRLRLPSILPGLVLAASDDP--SHSQAPSPWLTDPDASVQVRLLDVLLTPDNPSCPLVYLWRVHSQIPQRVWVPGVPVPSLALLLESVLRHVGLNEV	
Consensus	(1201)		QRLRLPSILPGLVLAASDDP--SHSQAPSPWLTDPDASV VLLLDVLLTPDNPSCPLVYLWRVHSQIPQRVWVPGVPVPA LSLALLLESVLRHVGLNEV	
KIAA0476	(1296)	1296	HKAVGLLLETLPPTGLHLGRG IYREILFLTMAALGKDHVD I VAFDKKYSAFNKLASSMGKEELRRRAQMPTPKA I DCRKCFGAPPEC	1391
<i>Bsg4</i>	(1123)		HKAVGLLLETLPPTGLHLGRG IYREILFLTMAALGKDHVD I VAFDKKYSAFNKLASSMGKEELRRRAQMPTPKA I DCRKCFGAPLEC	
Consensus	(1301)		HKAVGLLLETLPPTGLHLGRG IYREILFLTMAALGKDHVD I VAFDKKYSAFNKLASSMGKEELRRRAQMPTPKA I DCRKCFGAP EC	

图 3 小鼠 *Bsg4* 和人的 KIAA0476 氨基酸序列同源性的比对  
阴影部分为同源序列，它们具有 81.2% 的相似性

图 4 *Bsg4* 基因的预测结构域

第一个是 DENN 结构域, 第二个是 dDENN 结构域

### 2.3 *Bsg4* 基因全长编码区的克隆

通过 RT-PCR 我们获得了 *Bsg4* 基因全长编码区的序列, 并将它克隆至 pBluescript SK(-)载体上. 通过测序验证, 所得到的全长编码区的序列用于原位杂交探针的制备.

### 2.4 原位杂交

原位杂交结果显示, 在 8.5 d 的小鼠胚胎中, *Bsg4* 基因主要在头部的神经褶和神经管中表达(图版 I A). 在 9.5 d 小鼠胚胎, 脑中的表达信号最强. 成年小鼠脑石蜡切片的原位杂交结果显示, *Bsg4* 主要在小脑皮层和海马结构中表达(图版 I B). 这些结果暗示 *Bsg4* 在神经系统的发育和执行脑功能中发挥重要作用.

## 3 讨论

脑是生物体中最复杂, 调控最为精密的器官, 与其他器官相比脑具有独特的基因表达谱. 很多基因是在脑中特异表达的, 由于这些基因的表达丰度都很低, 用常规方法研究比较困难, 所以人们对这些基因的表达情况知之甚少. 我们将消减差异筛选的方法应用于小鼠脑中特异基因的寻找, 克隆到了一个在小鼠脑中特异表达, 并且在海马中表达的新基因 *Bsg4*. 该基因定位在小鼠的第 3 号染色体上, 含有一个 3639 bp 的完整阅读框, 编码一个包含 1213 个氨基酸的蛋白质. 与人 KIAA0476 基因的氨基酸序列同源性比较显示两者同源性达到 81.2%, 说明在不同种属中 *Bsg4* 具有一定的保守性, 暗示 *Bsg4* 在生物体内可能发挥着重要的作用.

我们利用生物信息学的方法预测了 *Bsg4* 编码的蛋白质, 它包含了一个 DENN 结构域和一个 dDENN 结构域. 现在已知 DENN 结构域存在于调节 Rab 介导过程以及调节 MAPKs 通路的某些蛋白质中. 对 Rab6IP1 进行的结构分析, 以及后来的研究显示出 DENN 结构域在调节 GDP/GTP 转换的活

性中发挥了作用<sup>[9, 10]</sup>, 但现在对于 DENN 的确切功能了解得仍然不清楚<sup>[11]</sup>. dDENN 结构域是一个伴随 DENN 结构域而存在的结构. 在 Rab6IP1 和 Rab3GEP 中进一步分析相关的序列发现 DENN 结构域旁边都会出现 dDENN 结构域, 但其功能还不清楚, 不过可以看出它对于 DENN 结构域的功能非常重要<sup>[12]</sup>. *Bsg4* 基因上的这些重要的结构域提示 *Bsg4* 可能对信号传递具有调节作用, 从而参与基因的表达调控, 并在胚胎发育过程发挥重要作用.

胚胎的整体原位杂交和切片的原位杂交结果显示, *Bsg4* 基因在胚胎发育中主要在头部表达. 在胚胎发育早期, 神经系统发育还不完全, 8.5 d 胚胎刚刚形成神经管, 到了 9.5 d 胚胎脑腔形成, 之后直到出生整个神经系统逐渐发育成熟, 脑的结构才发育完整<sup>[13]</sup>. 这提示 *Bsg4* 基因可能在胚胎发育中与头部的发育有关, 尤其是在脑中的强烈表达显示 *Bsg4* 基因可能对于脑的发育具有重要的信号调节作用. 此外, *Bsg4* 基因还在成年小鼠的海马结构和小脑皮层中有强烈表达. 这表示到神经系统发育成熟后, *Bsg4* 基因的表达集中在海马和小脑皮层中. 海马是学习和记忆的重要物质基础, 与情绪活动及内脏反应也有着密切关系, 并且是免疫调节的高级中枢之一, 与许多疾病有着密切的关系, 切除双侧海马区后, 患者会产生严重的“新近记忆”紊乱<sup>[14]</sup>. 小脑是负责维持身体平衡和运动的协调的重要中枢, 小脑病变与许多运动性疾病有密切关系, 小脑皮层的萎缩会造成“散发性退化性运动失调”. 我们的 *Bsg4* 基因表达的结果说明在成年小鼠中 *Bsg4* 也在发挥作用. *Bsg4* 很可能是一个在脑和神经系统的发育, 海马和小脑皮层结构功能的发挥上起到重要作用的基因. 对 *Bsg4* 基因的研究将有助于我们进一步揭示脑和神经系统的发育, 以及相关疾病的分子机制.

## 参 考 文 献

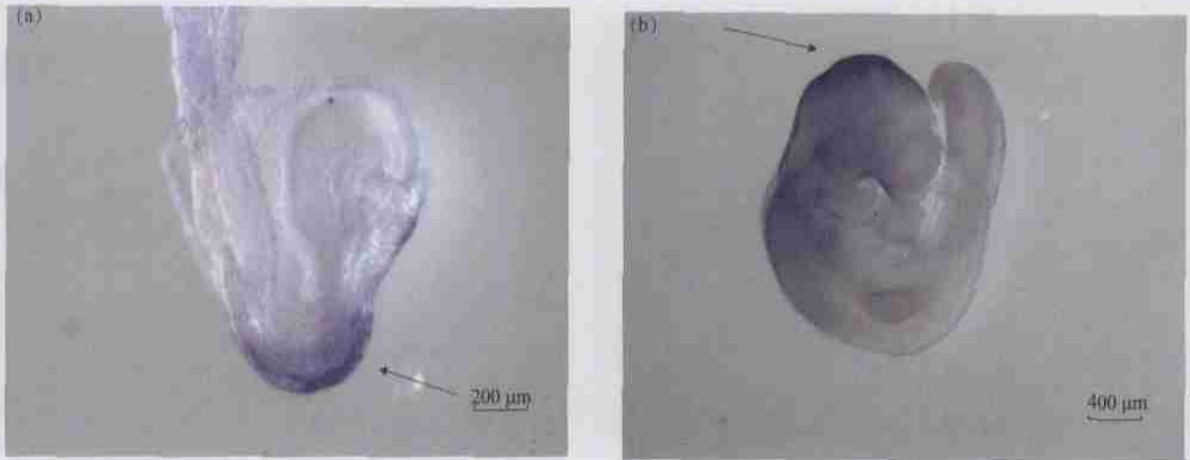
- 1 蔡文琴, 等. 发育神经生物学. 北京: 科学出版社, 1999

- 2 Momoi T, et al. Strain-specific caspase 3-dependent programmed cell death in the early developing mouse forebrain. *Neuroreport*, 2003, 14(1): 111
- 3 Lupo G, et al. Induction and patterning of the telencephalon in *Xenopus laevis*. *Development*, 2002, 129: 5421
- 4 Parasuraman R, et al. The apolipoprotein E gene: attention and brain function. *Neuropsychology*, 2002, 16(2): 254
- 5 May P C, et al. Altered gene expression in Alzheimer's disease brain tissue. *Can J Neurol Sci*, 1989, 16(4 Suppl): 473
- 6 Li Y P, et al. Cloning and complete coding sequence of a novel human cathepsin expressed in giant cells of osteoclastomas. *J Bone Min Res*, 1995, 10: 1197
- 7 Seki N, et al. Characterization of cDNA clones in size-fractionated cDNA libraries from human brain. *DNA Res*, 1997, 4(5): 345
- 8 Wilkinson D G, *In situ* hybridization. In: *Essential Development Biology: A Practical Approach*. Oxford: IRL Press, 1993
- 9 Janoueix-Louisey I, et al. Two-hybrid system screen with the small GTP-binding protein Rab6. Identification of a novel mouse GDP dissociation inhibitor isoform and two other potential partners of Rab6. *J Biol Chem*, 1995, 270(24): 14801
- 10 Tamaka M, et al. Role of Rab3 GDP/GTP exchange protein in synaptic vesicle trafficking at the mouse neuromuscular junction. *Mol Biol Cell*, 2001, 12(5): 1421
- 11 Wada M, et al. Isolation and characterization of a GDP/GTP exchange protein specific for the Rab3 subfamily small G proteins. *J Biol Chem*, 1997, 272(7): 3875
- 12 Levivier E, et al. uDENN, DENN, and dDENN: Indissociable domains in Rab and MAP kinases signaling pathways. *Biochem Biophys Res Commun*, 2001, 287(3): 688
- 13 Nagy A, et al. *Manipulating the Mouse Embryo: A Laboratory Manual*. 3rd edition, New York: Cold Spring Harbor Laboratory, 2002
- 14 Scoville W B, et al. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatr*, 1957, 20: 11

微生物学 1993—2003 年被引频率位列前 20 位的期刊

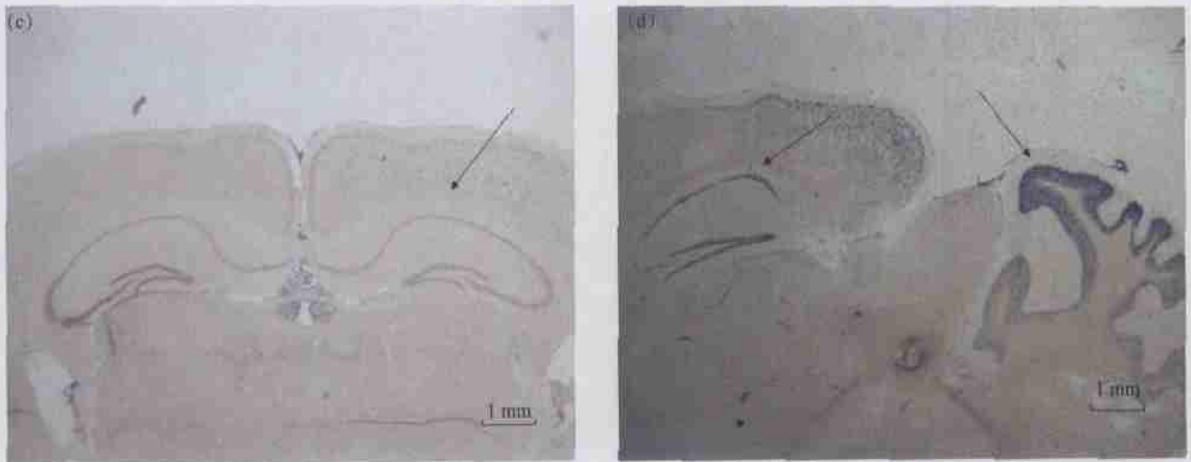
排序	期刊名称	论文数	总被引	平均被引
1	Microbiol Rev	120	21817	181.81
2	Science	283	34789	122.93
3	Nature	203	24405	120.22
4	Microbiol Mol Biol Rev	214	14569	68.08
5	Annu Rev Microbiol	289	17401	60.21
6	Clin Microbiol Rev	376	19778	52.60
7	Proc Nat Acad Sci USA	1698	75440	44.43
8	Fems Microbiol Rev	536	14658	27.35
9	Mol Microbiol	5100	126784	24.86
10	J Virol	13552	330217	24.37
11	Int J Syst Bact	1214	27103	22.33
12	Parasitol Today	683	14144	20.71
13	Trends Microbiol	879	17271	19.65
14	J Bacteriol	10981	198841	18.11
15	Appl Environ Microbiol	9003	151513	16.83
16	Antimicrob Agents Chemother	6597	110017	16.68
17	Virology	6413	105484	16.45
18	Microbiology-UK	2705	40374	14.93
19	J Gen Virol	4178	60968	14.59
20	Mol Biochem Parasitol	2298	29814	12.97

(摘自《英语科技论文撰写与投稿》)



A 小鼠胚胎的原位杂交

(a) 基因在 8.5d 胚胎神经管处(箭头所示)特异表达; (b) 基因在 9.5d 胚胎头部(箭头所示)特异表达



B 小鼠脑切片的原位杂交

基因在小鼠海马((a)中箭头和(b)左边箭头所示)和小脑皮层((b)右边箭头所示)表达