

小鼠胚胎脑特异表达的新基因 bsg3 的克隆及初步研究*

姚芝桑¹⁾ 付佳琪¹⁾ 陈微³⁾ 杨玮¹⁾
邹星¹⁾ 李宁⁴⁾ 李亦平^{1,2,3,4)ñ}

(1) 北京师范大学生命科学院生物医学研究所; 2) 北京师范大学生命科学院细胞增殖与调控生物学教育部重点实验室; 100875, 北京; 3) 美国哈佛大学 Forsyth 研究所, 021152, 波士顿, 美国; 4) 浙江赛尔生物医学研究院, 310006, 杭州 第一作者 25 岁, 女, 硕士生)

摘要 通过消减差异筛选的方法克隆到一个在小鼠胚胎脑特异表达的基因 bsg3 (brain specific gene 3). 它编码的蛋白与人的 KIAA0961 具有 80.3% 的同源性. bsg3 基因包含一个 1 644 bp 的完整阅读框, 编码一个含 548 个氨基酸, 具有 1 个 KRAB 结构域 (kruppel-associated box) 和 13 个 C2H2 型锌指结构域的蛋白. 该基因定位在小鼠第 7 号染色体上, 包含 5 个外显子, 4 个内含子. 以 bsg3 基因全长编码区为探针的原位杂交结果显示, bsg3 在小鼠胚胎脑及鸡胚脑特异表达. 半定量反转录聚合酶链式反应 (RT-PCR) 的结果表明, 在成体小鼠的组织中, bsg3 基因脑中表达也较强. 这提示 bsg3 基因可能在小鼠脑发育中起着重要的作用. 对 bsg3 基因时间和空间的表达模式的分析将有助于进一步揭示它在脑发育及功能维持中的作用.

关键词 bsg3 基因; KRAB 结构域; 脑; 小鼠

分类号 Q 785; Q 786

脑的发育是一个非常复杂的过程, 受很多基因的调控. 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, 为防治脑发育异常引起的疾病提供理论基础.

锌指蛋白 (zinc finger protein, ZFP) 是真核细胞中最普遍存在的一类 DNA 结合蛋白. 根据这些 ZFP 锌指结构上的特征, 可分为 C2H2 型、C3H4 型和 C4H4 型等, 其中 C2H2 型 ZFP 构成最大的转录因子家族. 约有 1/3 的 C2H2 型 ZFP 的 N 端含有保守的 KRAB 区域, 称为 Kruppe I 型 C2H2 ZFP. KRAB 是由 75 个氨基酸残基构成的蛋白质结构域, 依功能不同一般被划分为 A、B 两个亚区. 在哺乳动物细胞中, 含 KRAB 结构域的 ZFP 是最常见的转录因子. 一些实验证明 KRAB 结构域可以对基因的转录起抑制作用^[1]. 研究发现, 在 KRAB 中的 A 亚区的 45 个氨基酸的残基序列对转录抑制是必需而有效的, B 亚区自身并不具有转录抑制效应, 但具有增强 A 亚区的抑制功能^[2-3]. 尽管 KRAB ZFP 的功能大部分是未知的, 但它们已在细胞分化和发育中显示了重要的作用.

我们应用消减差异筛选的方法寻找到了一个在小鼠脑中特异表达的基因 bsg3 (brain specific gene 3), 对它的序列特征进行了生物信息学分析, 用胚胎原位杂交和半定量反转录聚合酶链式反应 (RT-PCR) 对它的表达情况进行了研究, 并对它的可能功能进行了初步预测.

* 浙江赛尔生物医学研究院资助项目; 高等学校重点实验室访问学者资金资助项目

ñ 联系人

收稿日期: 2003-12-29

1 材料与amp;方法

1.1 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 μg mRNA 为模板,逆转录采用随机引物,并掺入 ^{32}P dCTP,42 $^{\circ}\text{C}$ 反应 1 h;0.1 mol L^{-1} NaOH 重悬反应产物,65 $^{\circ}\text{C}$ 孵育 20 min 使 RNA 模板水解;用 0.1 mol L^{-1} 醋酸中和探针,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(Invitrogen 公司)制备,实验步骤按其说明书进行.

将 ^{32}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 μL 水重悬,100 $^{\circ}\text{C}$ 热变性 1 min,加入等体积 2 倍杂交缓冲液,65 $^{\circ}\text{C}$ 杂交 20 h.加入等体积 HEPES 缓冲液(10 mmol L^{-1} HEPES,1 mmol L^{-1} EDTA,pH 7.5),20 μg 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 cDNA 探针用乙醇沉淀纯化,在相同条件下进行第 2 轮消减杂交.

我们用 7.5 d 小鼠胚胎 cDNA 文库中的 5×10^5 个克隆进行筛选,每个直径 145 mm 的圆平板上约有 1 万个克隆,每板制 2 张膜用于与第 2 轮消减杂交得到的 cDNA 探针杂交.杂交,洗膜和曝光按照常规方法进行^[4].为纯化克隆在低浓度条件下进行第 2 轮杂交,所得到的克隆用 ABI377 测序仪测序.

1.2 用生物信息学方法分析 bsg3 基因序列特征 将所得序列在 NCBI(www.ncbi.nlm.nih.gov)上用 BLAST(basic local alignment search tool)方法与已知序列及小鼠基因组进行比对,对其外显子和内含子进行分析.使用 Vector NTI 软件对 BSG3 与人的 KIAA0961 的氨基酸序列进行同源性比较.用 NCBI 和 Pfam 数据库对蛋白的结构域和可能的功能进行预测.

1.3 RT-PCR 克隆 bsg3 的全长编码区序列 根据 bsg3 的 cDNA 序列用软件设计一对引物,在 5' 端引物加入 Hind 酶切位点,在 3' 端引物加入 Kpn 酶切位点.两引物序列为:

P1 (sense): 5' GCAAAGCTTATGATCCTCGA GGAA GAACAA GCC 3';

P2 (antisense): 5' ATGGTACCTTTGCAA GGCAA GGATTCGT 3'.

提取小鼠胚胎的总 RNA 为模板,经 M-MLV 逆转录酶(Invitrogen 公司)合成 cDNA 第 1 条链.使用高保真的 DNA 聚合酶(大连宝生物工程有限公司)扩增 bsg3 cDNA 序列.采用 100 μL 的聚合酶链式反应(PCR)反应体系,反应条件为 94 $^{\circ}\text{C}$ 1 min,58 $^{\circ}\text{C}$ 1 min,72 $^{\circ}\text{C}$ 2 min,循环 30 次,72 $^{\circ}\text{C}$ 延伸 7 min. PCR 产物用 $w = 1\%$ 琼脂糖凝胶进行分离,将产物所在的胶的条带切下,用 QIAquick Gel Extraction Kit 试剂盒(QIAGEN 公司)纯化和回收.分别将纯化的 PCR 产物和 pBluescript SK(-) 载体(Stratagene 公司)进行 Hind 和 Kpn 的双酶切(New England Biolabs 公司)反应,经过酚与氯仿按体积比 1:1 纯化后,用 T4 DNA 连接酶进行连接,将 PCR 产物克隆至 pBluescript SK(-) 载体上.连接后的产物转化感受态大肠杆菌 DH5,筛选重组子,制备质粒 DNA,经测序验证 bsg3 的全长编码区序列.

1.4 原位杂交 原位杂交参照文献[5]的方法,以 bsg3 编码区的 cDNA(1.6 kb)为模板,用 DIG Nucleic Acid Labeling Kit(Roche 公司)制备 DIG 标记的 RNA 探针,分别与 8.5,9.5,10.5 d 的小鼠胚胎和第 7,10 天的鸡胚进行整体的原位杂交.

1.5 半定量 RT-PCR 设计 bsg3 引物为:sense,5' TTCATCAGAATCACGCA GAG 3';antisense,5' GGAA TATCACAATGTCGCA GT 3',扩增产物为 439 bp.内参对照 galdp 引物为:sense,5'

ACCACA GTCCA TGCCA TCAC 3'; antisense, 5' TCCACCACCCTGTTGCTGTA 3', 扩增产物为 452 bp. 取小鼠成体各组织 100 mg, 使用 Trizol 试剂 (Invitrogeng 公司), 按其说明书一步法提取总 RNA. 使用逆转录酶 M-MLV, 用 Oligo (dT) 合成 cDNA 第 1 条链, 然后以 cDNA 第 1 条链为模版进行扩增. PCR 扩增条件为: 94 2 min, 94 1 min, 55 0.5 min, 72 1 min, 循环 30 次, 72 延伸 7 min. 取 PCR 产物 6 μ L, 经 $w = 1.5\%$ 琼脂糖凝胶电泳, 照相.

2 结果

2.1 脑特异基因 bsg3 序列 利用消减差异筛选的方法克隆到脑特异表达的基因 bsg3, 长 2 604 bp, 序列如下:

```

1 GAAGAAGGAAAAGAAGAGAAAAGAAAAGAAAACAAGGCCTGGAGTGTCTGCCAGTCTTCAG
60 AAGAGACCCTAGAAGCCAGCAGCTACTACTTCTGTTTAAAACAGAGATTATGTCTCTAGC
120 CTC TGT GAA CGG CTT CAC CAG GAC ATG ATC CTC GAG GAA GAA CAA GCC AGT TCT TGC GAT
      M I L E E E Q A S S C D
180 TCT GGA ATC ATG GCT CAT AGT ACG GTG ATG TTC AGA GAT GTG GCT GTG GGC TTT TCT CAG
      S G I M A H S T V M F R D V A V G F S Q
240 GAA GAG TGG GAG TGC CTG AGT GCC TAC GAG AGG GAT CTG TAC AGA GAT GTG ATG CTA GAG
      E E W E C L S A Y E R D L Y R D V M L E
300 AAC TAC AGC CAC TTG GTG TCA CTG GCA GGA TGT TCC ATT TCT AAG CCA GAT GTG ATC ACC
      N Y S H L V S L A G C S I S K P D V I T
360 CTG TTG GAA CAA GGG AAG GAG CCG TGG ATG ATT GTG AGG GCT GAG AAG AGA AGG TGG AGC
      L L E Q G K E P W M I V R A E K R R W S
420 CGA GAT CTG GAA TCC AGA TAC AGC AGT AAC GGG TTA CTT CCA GAA AAG AAT ACT TAC GAA
      R D L E S R Y S S N G L L P E K N T Y E
480 ATT AAT CTA TCT CCG TGG GAA ATA ATG GGA AGA ATT CAA AGG CGT GGC CCC GAA GAC TCC
      I N L S P W E I M G R I Q R R G P E D S
540 CTT TTG GGG AAA GAC TTT GAA TAT AAA ATA TAT GAA GAA CAA GAG AAC TCT CAT AGG GTA
      L L G K D F E Y K I Y E E Q E N S H R V
600 TAT TTC AGA CAC GTG ATA AAA ACC ACC TCT GGA AAA AGA CCT CGG TAC AGA AAA CGC ACA
      Y F R H V I K T T S G K R P R Y R K R T
660 CCT GTT AGT CTC TAT CAG AAA ACT CCT AAC GGA GAA AAG CCC TAC GAA TGT GGG GAG TGT
      P V S L Y Q K T P N G E K P Y E C G E C
720 GGG AAG GCT TTC AAA GTC CGC CAA CAA CTT ACT TTC CAT CAG AGG ATC CAC ACC GGC GAG
      G K A F K V R Q Q L T F H Q R I H T G E
780 AAG CCC TAC GAA TGC AAG GAA TGT GGC AAA GCC TTC CGA CAG TGC GCC CAC CTC AGT CGC
      K P Y E C K E C G K A F R Q C A H L S R
840 CAT CAG AGG ATC CAC GCC TCT GAC AAG CTC TAT GAA TGT AAG AAG TGT GCA AAG ATC TTC
      H Q R I H A S D K L Y E C K K C A K I F
900 ACG TGT AGC TCG GAC CTT CGA GGG CAT CAG AGG TCT CAC GTG GGC GAG AAG CCC TAC GAT
      T C S S D L R G H Q R S H V G E K P Y D
960 TGT AAA GAG TGC GGG AAG GCC TTT CGA GTC CGT GGA CAG CTC ATG CTC CAC CAG AGG ATC
      C K E C G K A F R V R G Q L M L H Q R I
1 020 CAC ACC GGT GAG AAG CCC TAC GCA TGC ACG GAG TGC GGC AAG AGC TTC CGG CAG GTA GCG
      H T G E K P Y A C T E C G K S F R Q V A
1 080 CAC CTC ACT CGG CAT CAG AGG CTC AAC AGT GGC AGC TCA CGC CAC GAG TGT AAG GAG TGC
      H L T R H Q R L N S G S S R H E C K E C
1 140 GGC CGG GCT TTC CTG TGC AGC ACG GGC CTC CGG CTG CAT CAC AAA CTG CAC ACG GGA GAG
      G R A F L C S T G L R L H H K L H T G E

```

```

1 200 AAA CCC TAC GAC TGC AAG GAG TGT GGG AAG GCG TTC CGG GTG AGG CAG CAG CTC ACC CTG
      K P Y D C K E C G K A F R V R Q Q L T L
1 260 CAC GAG AGG ATC CAC ACC GGC GAG AAG CCC TTC GAC TGC AAG GAG TGC GGG AAG ACC TTC
      H E R I H T G E K P F D C K E C G K T F
1 320 AGC CGC GGC TAC CAT CTC ACC CTC CAC CAG AGG ATC CAC ACT GGG GAG AAG CCT TAC GAG
      S R G Y H L T L H Q R I H T G E K P Y E
1 380 TGT AAG GAG TGC CGC AAG TTC TTC CGC CGG TAC TCG GAG CTC ATT TCC CAC CAG GGC ATT
      C K E C R K F F R R Y S E L I S H Q G I
1 440 CAC ATC GGG GAG AAA CCC TAT GAC TGC AAG GAG TGC GGA AAG GCT TTT AGG CTG TTC TCC
      H I G E K P Y D C K E C G K A F R L F S
1 500 CAG CTG ACT CAG CAC CAG AGC ATT CAT TTT GGT GAG AAG CCC TAT AAA TGT ATG GAA TGT
      Q L T Q H Q S I H F G E K P Y K C M E C
1 560 GAG AAG ACT TTC AGA CTC CTC TCA CAA CTG ACG CAG CAT CAG AGC ATT CAC ACG GGC GAG
      E K T F R L L S Q L T Q H Q S I H T G E
1 620 AAG CCT TAC GAT TGT AAG GAG TGC GGG AAA GCC TTT CGC CTC CAC TCG TCT CTC ATT CAG
      K P Y D C K E C G K A F R L H S S L I Q
1 680 CAC CAG AGA ATC CAC TCT GGC GAG AAG CCG TAC AAA TGT ACG GAG TGT AAG AAG GCC TTC
      H Q R I H S G E K P Y K C T E C K K A F
1 740 CGT CAA CAC TCA CAC CTC ACC TAC CAT CAG CGA ATT CAT AAG AAT TTA TAA GGA CCC TTC
      R Q H S H L T Y H Q R I H K N L
1 801 TACGAATCCTTGCCTTGCAAATGTAATAAGGAGCAGTAGGAAATCAGAAAATTCTAGAAG
1 860 AAAAAAAGGTTTGAATCCTTATTTTGTCTGTGTTTCATAAGTAACAGCATAAAAGGTCAAAG
1 920 GAAAAAATGGTAAACTTTTGGAAATTCCTTTCGTAAATGTATAAAAGGTATTTTGTCTTA
1 980 CTGTACACCTACCTAAGGGCATAGTCAGGGTACCAACTCATTTTGGAAAATTAATAAATAA
2 040 AAAGAATTGTCGTTTAGCCTGCATAAACTCTCCGAGTGACTGAAAGTGGTTGAGGTACAC
2 100 AGTTCTTAGAGTGTTCATCAGAATCACGCAGAGTCATATGGAATTAAGAAGTTGTATCA
2 160 GGAAAACCATTGTGTAATAACATACAGTGGAGCAGCGATTCTCAGCCTGTGGGTCCCGAC
2 220 CCTCTGAGGGTAGAGCCACCTTTCTCAGGAGTCACCTAGGACCATCGAGACACAGATTC
2 280 GCATTATGATCCATTACAGTAGCAAAATTACAGTTATGAAGTAGCAACACGTGATTTAAT
2 340 GGTGGGAGTCACACAGCATGAGGAAGTGTATTAAGGGTCACAGCGTTAGGGAGGTTGA
2 400 GAACCACTGCTCTAGAATTCAGCCATGATAGACTGGCTTGTACTTGGTAGTGAATAGGTC
2 460 AGGCTGCCAGGATGTGAACCTACTGATCCGTCTTGCCACCAGTGTAGCACAGTAAAGCGT
2 520 CCCGTGTCTTCCACTGCGACATTGTGATATTCCCACCTAACTCAACGCACCAAACCTCTA
2 580 AATCAATGAGACAACCACCAACATT

```

编码区序列的下方给出翻译的氨基酸序列。起始密码子(ATG)、终止密码子(TAA)都用下划线标出。对该基因在 NCBI 上做分析, 结果发现该基因为一功能未知的新基因。向 NCBI 提交该序列, GeneBank 登录号为 AY605122。

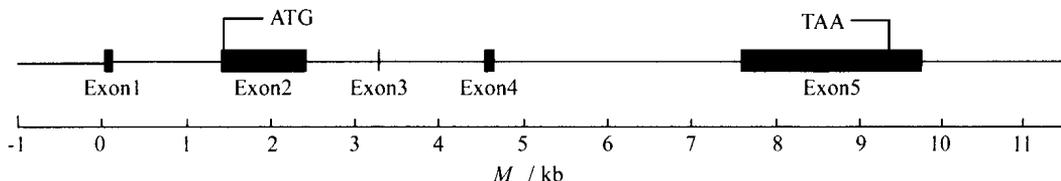


图 1 bsg3 在染色体上的分布

黑块部分代表外显子; 两黑块间为内含子; ATG 为起始密码子; TAA 为终止密码子。

2.2 bsg3 序列的生物信息 在小鼠基因组数据库中做比对分析得知 bsg3 定位在小鼠的第 7 号染色体上, 包含 5 个外显子, 4 个内含子, 见图 1。对 BSG3 结构域的分析显示该蛋白 N 末端

含有一个 KRAB 结构域, C 末端含有 13 个相连的 C2H2 型锌指结构域(zf-C2H2) (图 2, 见封三中). 这些结构域提示 bsg3 基因可能是一个转录因子. 小鼠的 BSG3 和人的 KIAA0961^[6] 氨基酸序列的同源性比较结果显示两者的同源性达 80.3%, 见图 3.

KIAA0961	(1)	MA RDL VMFRDVAV D ESQEEWECL NS Y Q R N LYRDV I
BSG3	(1)	MILEEEQASSCDSGI MA HST VMFRDVAV G ESQEEWECL SA Y E R D LYRDV M
Consensus	(1)	MA VMFRDVAV ESQEEWECL Y R LYRDV
KIAA0961	(36)	LENYS N LVSLAGCSISKPDVITLLEQGKEPWM V VR D EKRRW TL DLESRY D
BSG3	(51)	LENYS H LVSLAGCSISKPDVITLLEQGKEPWM I VR A EKRRW SR DLESRY S
Consensus	(51)	LENYS LVSLAGCSISKPDVITLLEQGKEPWM VR EKRRW DLESRY
KIAA0961	(86)	TKK L FQG K DI YE M NLS Q W KV M E RI KSC G L EEQE SP H
BSG3	(101)	SNG L LPE K NT YE I NLS P W E I M G RI QRR G PEDSLLGKDFEYKIV EEQE NS H
Consensus	(101)	L K YE NLS W M RI G EEQE H
KIAA0961	(122)	E V C FR Q V T KTIS E K M P T YRK L T SLP LYQK SH N R EKPYECGECGKAF R VRQ
BSG3	(151)	R V Y FR H V I KTIS G K R P R YRK R T PVS LYQK TP N G EKPYECGECGKAF K VRQ
Consensus	(151)	V FR V KTIS K P YRK T LYQK N EKPYECGECGKAF VRQ
KIAA0961	(172)	QLTFHQRIHTGEKPYECKE CGKAFRQCAHLRHRQRIH T SDKLYECKKC G K
BSG3	(201)	QLTFHQRIHTGEKPYECKE CGKAFRQCAHLRHRQRIH A SDKLYECKKC A K
Consensus	(201)	QLTFHQRIHTGEKPYECKE CGKAFRQCAHLRHRQRIH SDKLYECKKC K
KIAA0961	(222)	IFTC G SDLR V HQR I H I GEKPY E CKECGKAFRVRGQL N LHQRIHTGEKPY E
BSG3	(251)	IFTC S SDLR G HQR S H V GEKPY D CKECGKAFRVRGQL M LHQRIHTGEKPY A
Consensus	(251)	IFTC SDLR HQR H GEKPY CKECGKAFRVRGQL LHQRIHTGEKPY
KIAA0961	(272)	C K ECGK A FRQ Y AHLTRHQRLN IAEKCY ECKE CG Q AFLCSTGLRLRHHKLHT
BSG3	(301)	C T ECGK S FRQ V AHLTRHQRLN SGSSRH ECKE CG R AFLCSTGLRLRHHKLHT
Consensus	(301)	C ECGK FRQ AHLTRHQRLN ECKE CG AFLCSTGLRLRHHKLHT
KIAA0961	(322)	GEKPY E CKECGKAFRVRQQLTLH Q RIHTGEKPY D DCKE CGKTFSRGYHLTL
BSG3	(351)	GEKPY D CKECGKAFRVRQQLTLH E RIHTGEKPY F DCKE CGKTFSRGYHLTL
Consensus	(351)	GEKPY CKECGKAFRVRQQLTLH RIHTGEKPY DCKE CGKTFSRGYHLTL
KIAA0961	(372)	HQRIHTGEKPYECKE C Q KFFRRYSELISHQGIHIGEKPY E CKECGKAFRL
BSG3	(401)	HQRIHTGEKPYECKE C R KFFRRYSELISHQGIHIGEKPY D CKECGKAFRL
Consensus	(401)	HQRIHTGEKPYECKE C KFFRRYSELISHQGIHIGEKPY CKECGKAFRL
KIAA0961	(422)	FSQLTQHQSIIHFGEKPY F KC K ECEKTFRLLSQLTQHQSIIHTGEKPYDCKE C
BSG3	(451)	FSQLTQHQSIIHFGEKPY Y KC M ECEKTFRLLSQLTQHQSIIHTGEKPYDCKE C
Consensus	(451)	FSQLTQHQSIIHFGEKPY KC ECEKTFRLLSQLTQHQSIIHTGEKPYDCKE C
KIAA0961	(472)	GKAFRLHSSLIQHQRHS GEKPYKC K ECKKA FRQHSHTLYHQRIH NVT
BSG3	(501)	GKAFRLHSSLIQHQRHS GEKPYKC T ECKKA FRQHSHTLYHQRIH KNL
Consensus	(501)	GKAFRLHSSLIQHQRHS GEKPYKC ECKKA FRQHSHTLYHQRIH

图 3 小鼠 BSG3 和人的 KIAA0961 氨基酸序列同源性的比对

阴影部分及 Consensus 为同源序列, 它们具有 80.3% 的相似性.

2.3 原位杂交结果 在 8.5, 9.5, 10.5 d 小鼠胚胎中, bsg3 基因主要在脑表达, 在第 期鸡胚中, bsg3 基因也主要在脑表达 (如图 4 箭头所示, 见封三下).

2.4 bsg3 半定量 RT-PCR 检测结果 在成体小鼠 10 种组织中的 bsg3 表达结果表明, 在脾、肾、脑、骨骼肌、心脏中都有明显的扩增条带, 其中以在脑中表达水平最高, 而在肝、肠、睾丸、肺

中有较弱的表达水平(图 5).
这提示 *bsg3* 在上述组织中广泛表达.

3 讨论

脑是生物体中最复杂的器官,与其他器官相比脑具有独特的基因表达谱.脑特有的基因占有很大的比例,

但是由于这些基因的表达丰度都很低,用常规的方法研究比较困难,所以人们对脑中特异表达的基因知之甚少.我们首次将消减差异筛选的方法应用于小鼠脑中特异基因的寻找.通过这个方法,我们筛选到一个在小鼠胚胎脑特异表达的基因 *bsg3*.该基因定位在小鼠第 7 号染色体上,含有一个 1 647 bp 的完整阅读框,编码一个含有 548 个氨基酸的蛋白.

通过生物信息学的方法预测到,在 *BSG3* 蛋白的 N 端含有 ZFP 结构域 KRAB, C 端含有 13 个相连的 *zf-C2H2*. Hennemann 等^[7]报道, KRAB 蛋白 *Krim-1* 和 *c-Myc* 结合,可抑制 *c-Myc* 的作用,从而抑制了 *c-Myc* 的致癌活性. KRAB 蛋白抑制作用的机理可能是:它们是 RNA 聚合酶,和启动子及 RNA 结合和剪切的抑制剂^[8].作为哺乳动物最常见的转录因子,这个家族的蛋白参与了核仁维持、细胞分化、细胞增殖、细胞凋亡以及致癌性转化等一序列细胞活动.*zf-C2H2* 是一类结合核酸的重要结构域. *BSG3* 这些结构域提示 *bsg3* 可能作为一个重要的转录因子,参与了基因表达和胚胎发育等重要过程.

小鼠胚胎及鸡胚原位杂交的结果显示 *bsg3* 在脑特异表达,而半定量 RT-PCR 的结果也显示 *bsg3* 在成体小鼠脑中表达水平较高,这提示 *bsg3* 可能在胚胎发育过程中和脑的发育有关.脑和神经系统的发育非常复杂,受很多基因的调控.如 *hox*, *pax*, *lim*, *pou*, *emx*, *dlx* 等同源盒基因家族^[9],都参与了脑和神经系统的发育.鼠的 *hoxc-8*, *hoxb-4*, *hoxa-2*, *hoxd-2*, *hoxd-13* 等与神经系统的格局化有关. *cnbp* 基因敲除的小鼠前脑缺失,面部畸形,而基因敲除的小鼠转基因后则能挽救这种缺失^[10]. *En-1, 2* 缺失的小鼠顶盖和后脑的前半部分不能发生,而 *pax-5* 缺失的小鼠,其脑峡也不能发生.还有一些基因与疾病相关,如 *Apo E-epsilon4* 等基因与进行性阿尔茨海默病相关^[11].同样 *bsg3* 也很可能是一个对脑的发育及功能维持起重要作用的基因,对 *bsg3* 基因的研究将有助于进一步揭示脑的发育以及相关疾病的分子机制.

4 参考文献

- [1] Abrink M, Ortiz J A, Mark C, et al. Conserved interaction between distinct Kruppel-Associated box domains and the transcriptional intermediary factor 1 [J]. Proc Natl Acad Sci, 2001, 98:1422
- [2] Margolin J F, Friedman J R, Meyer W K, et al. Kruppel-associated boxes are potent transcriptional repression domains[J]. Proc Natl Acad Sci, 1994, 91(10):4509
- [3] Witzgall R, O'Leary E, Leaf A, et al. The kruppel-associated Box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression[J]. Proc Natl Acad Sci, 1994, 91(10):4514
- [4] Li Yiping, Alexander M, Yelick P, et al. Cloning and complete cDNA sequence of a novel human Cathepsin expressed in osteoclast[J]. J Bone Min Res, 1995, 10:1197

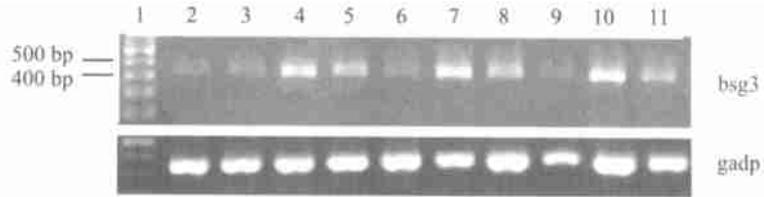


图 5 RT-PCR显示 *bsg3* 在成体小鼠组织中的表达情况

1. DNA 相对分子质量标准; 2-11. 成体鼠的肝、肠、脾、肾、睾丸、脑、骨骼肌、肺、心脏、子宫; *gapd* 做阳性对照.

- [5] Wilkinson D G. In situ hybridization Stern CD, Holland PWH essential development biology: a practical approach[M]. Oxford: IRL Press, 1993: 257-274
- [6] Ishikawa K, Nagase T, Nakajima D, et al. Prediction of the coding sequences of unidentified human genes: 78 new cDNA clones from brain which code for large proteins in vitro[J]. DNA Res, 1997, 4(5): 307
- [7] Hennemann H, Vassen L, Geisen C, et al. Identification of a novel Kruppel-associated box domain protein, Krim-1, that interacts with c-Myc and inhibits its oncogenic activity[J]. J Biol Chem, 2003, 278(31): 28799
- [8] Rutia R. KRAB-containing zinc-finger repressor proteins[J]. Genome Biol, 2003, 4(10): 231
- [9] 蔡文琴, 李海标. 发育神经生物学[M]. 北京: 科学出版社, 1999: 21-25
- [10] Chen Wei, Liang Yuqiong, Deng Wenjie, et al. The zinc-finger protein CNBP is required for forebrain formation in the mouse[J]. Development, 2003, 130(7): 1367
- [11] Parasuraman R, Greenwood P M, Sunderland T. The apolipoprotein E gene, attention, and brain function[J]. Neuropsychology, 2002, 16(2): 254

CLONING AND CHARACTERIZATION OF A NOVEL MOUSE GENE *bsg3* cDNA SPECIFICALLY EXPRESSED IN BRAIN

Yao Yisang¹⁾ Fu Jiaqi¹⁾ Chen Wei³⁾ Yang Wei¹⁾
Zou Xing¹⁾ Li Ning⁴⁾ Li Yiping^{1,2,3,4)}

(1) Biomedical Research Institute, (2) Key Laboratory of cell proliferation and regulation Biology of Ministry of Education, College of Life Science, Beijing Normal University;
3) The Forsyth Institute, Harvard University, 02115, Boston, USA;
4) Zhejiang Cell Biomedical Research College, 310006, Hangzhou, China)

Abstract Bsg3 gene was cloned by subtractive-different screening as a gene expressed specifically in developing mouse brain. It is 80.3% homology to human KIAA0961 gene at the amino acid level. Bsg3 contains a 1644 bp long complete reading frame and encoding protein of 548 amino acid residues, which contains a KRAB motif and 13 ZF-C2H2 motifs. The gene is on mouse chromosome 7 and comprises 5 exons and 4 introns. It is specifically expressed in developing mouse and chicken brain by whole mount in situ hybridization using bsg3 full-length coding sequence as probes. In addition, RT-PCR shows that bsg3 has a higher expression in brain than in other adult tissues. These results indicate that bsg3 gene may be involved in the development of brain. The analysis of expression profile of bsg3 gene may provide valuable insights into the function of bsg3 in the development and function of brain.

Key words bsg3; KRAB; brain; mouse



图1 在芽增殖培养基 MS₂ 上生长的美洲商陆芽丛

图2 在生根培养基 R₃ 上已生根的美洲商陆植株

图3 移栽至土壤的美洲商陆生根植株



图2 Bsg3 基因的预测结构域
第1个是KRAB 结构域，后面是13个C2H2型锌指结构域。

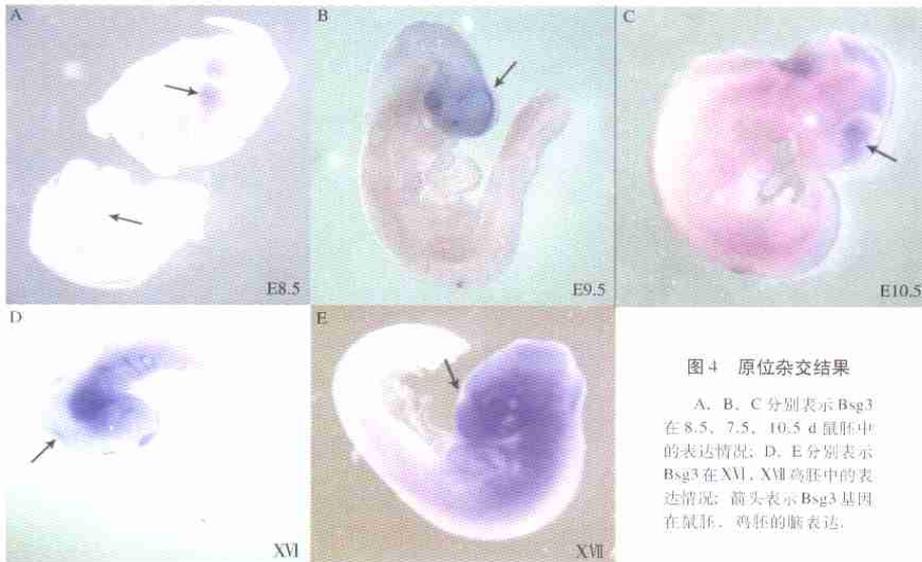


图4 原位杂交结果

A, B, C 分别表示 Bsg3 在 8.5、7.5、10.5 d 鼠胚中的表达情况；D, E 分别表示 Bsg3 在 XVI, XVII 鸡胚中的表达情况；箭头表示 Bsg3 基因在鼠胚、鸡胚的脑表达。