

胚胎发育早期头部特异表达的小鼠新基因 BSG5 的克隆及功能*

唐小静¹ 李昱华¹ 陈微² 陈荣³ 程睿³ 李亦平^{1,2,3**}

1. 北京师范大学生命科学学院, 北京 100875

2. Forsyth Institute, Harvard School of Dental Medicine, Boston, MA 02115, USA

3. 浙江赛尔生物医学研究院, 杭州 310006

摘要 Brain Specific Gene 5 (BSG5) 基因是用消减差异筛选的方法克隆到的在小鼠胚胎头部特异表达的新基因。它与人的 KIAA0628 基因在氨基酸水平上有 81.9% 的同源性。BSG5 基因长 2 487 bp, 定位在小鼠的第 15 号染色体上, 包含 2 个外显子。它编码的蛋白质全长 499 个氨基酸, 含 12 个 C2H2 型的锌指结构域。以 BSG5 基因全长编码区为探针的原位杂交结果显示 BSG5 基因在小鼠胚胎发育早期头部特异表达, 在小鼠胚胎发育稍后时期的尾部和肢芽也有表达。此外, 以鸡胚为模型研究 BSG5 基因也发现该基因在鸡胚的头部特异表达。这提示 BSG5 基因与头部发育有密切关系, 其结构与表达特征预示着它编码的是一个具 DNA 结合功能的转录调控因子 [动物学报 50 (4): 593-599, 2004]。

关键词 BSG5 基因 生物信息学 锌指结构域

Cloning and characterization of a novel mouse gene BSG5 cDNA specifically expressed in early embryo brain*

TANG Xiao-Jing¹, LI Yu-Hua¹, CHEN Wei², CHEN Rong³, CHENG Rui³, LI Yi-Ping^{1,2,3**}

1. Life Science College, Beijing Normal University, Beijing 100875, China

2. Forsyth Institute, Harvard School of Dental Medicine, Boston, MA 02115, USA

3. Zhejiang Cell Biomedical Research College, Hangzhou 310006, China

Abstract BSG5 gene was cloned by subtractive-different screening as a new gene expressed specifically in mouse brain. It is 81.9% identical to human KIAA0628 gene at the amino acid level. It is 2 487 bp long. The gene is on mouse chromosome 15 and comprises 2 exons and encodes a protein of 499 amino acid residues with 12 C2H2 type zinc fingers. BSG5 gene is specifically expressed in early mouse embryo brain by *in situ* hybridization using BSG5 full-length coding sequence as probes, and we can see it also expressed in limb bud and tail bud of the later embryos. In addition, *in situ* hybridization using chicken embryos reveals that BSG5 is specifically expressed in the brain. The structure and expression pattern implies it might be a DNA binding transcription regulator. The results indicate that the BSG5 gene may be involved in the development of brain. [Acta Zoologica Sinica 50 (4): 593-599, 2004].

Key words BSG5 gene, Bioinformatics, Zinc finger domain

锌指 (Zinc finger, ZNF) 结构是由 Miller 等人于 1985 年最初在爪蟾转录因子 TF-III_A (Petering et al., 2000) 中发现。而含有锌指结构的蛋白质则被称为锌指蛋白 (Zinc finger protein, ZNP)。锌指结构是真核生物转录因子中最普遍的 DNA 连

接基元之一, 迄今已发现 1 000 多个包含锌指结构的 cDNA 序列, C2H2 型是锌指中数量最多的一类, 约占 60%, 其主要特征是在蛋白质的一级结构上, 存在 Cys-X₂-Cys-X₁₂-His-X₃-His 的结构序列单元 (X 为任意氨基酸)。因为其中的 2 个半胱

2004-02-03 收稿, 2004-04-14 接受

* 浙江赛尔生物医学研究院资助项目 [This research was funded by a grant from Zhejiang Cell Biomedical Research College]

** 通讯作者 (Corresponding author). E-mail: YPLi@forsyth.org

© 2004 动物学报 Acta Zoologica Sinica

氨酸和 2 个组氨酸通过配位键与一个锌离子结合, 结构类似手指状, 所以这类序列单元被称为锌指结构。锌指蛋白与胚胎发育和细胞分化有密切关系, 是一类具有重要功能的转录调节因子。如小鼠 Krox20 蛋白因子是神经系统正常发育所不可缺少的 (Lania et al., 1997); 猴病毒 40 启动子蛋白 (SP1)、红细胞类 Krüppel 因子 (EKLF)、肺细胞类 Krüppel 因子 (LKLK) 等蛋白质因子是胚胎发育所不可缺少的 (Lania et al., 1997; Nuez et al., 1995; Armstrong et al., 1998)。

脑的发育是一个非常复杂的过程, 受到很多基因的调控。在发育的不同时期、不同部位、基因的表达是不同的。基因的调控与脑的发育、疾病的产生有着密切的关系。近年来大量的研究表明, 同源盒基因家族与神经系统和脑的发育密切相关, 如 Hox, PAX, LIM, POU, Emx, Dlx 等同源盒基因家族 (蔡文琴、李海标, 1999)。还有一些基因与脑的疾病相关, 如 ApoE-epsilon4 (Parasuraman et al., 2002), pADHC-9 (May et al., 1989) 等许多基因与阿尔茨海默氏病 (进行性老年痴呆) 相关。本文通过消减差异筛选的方法克隆到一个新的 C2H2 型锌指蛋白基因, 该方法曾成功地应用于组织蛋白酶 K (Li et al., 1995) 和破骨细胞特异性质子泵亚基 ATP6i (Li et al., 1996) 的克隆。

由于小鼠基因在序列和功能上都与人高度同源, 小鼠经常用作人类重要疾病基因功能研究的模型动物。利用生物信息学的方法对该基因序列特征进行了分析, 并以 BSG5 基因全长编码区为探针进行原位杂交, 对此新基因进行了克隆及初步的功能研究工作。

1 材料与方法

1.1 材料

正常雌雄和怀孕 ICR 小白鼠均购自北京大学医学部实验动物中心。总 RNA 抽提试剂 (Trizol) 和逆转录酶 (M-MLV Reverse Transcriptase) 购自 Gibco 公司。Taq 和 pfu DNA 聚合酶购自上海生工公司。引物由北京博润基因公司合成。DNA 测序由上海博亚公司完成。地高辛标记与检测试剂盒购自罗氏 (Roche) 公司。

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 μ g mRNA 为模板, 用随机引物进行反转录, 反转录

的同时掺入 32 PdCTP, 反应条件为 42 $^{\circ}$ C, 1 h; 反应产物用 0.1 mol 的 NaOH 重悬, 65 $^{\circ}$ C 孵育 20 min 使 RNA 模板水解; 用 0.1 mol 醋酸中和探针, 并用 Sephadex G-50 纯化探针。生物素标记的小鼠 9.5 d 胚胎躯干的 mRNA 将使用 Invitrogen 的消减试剂盒制备, 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 μ l 水重悬, 100 $^{\circ}$ C 加热变性 1 min。加入等体积 2 \times 杂交缓冲液 (Invitrogen), 65 $^{\circ}$ C 杂交 20 - 24 h。加入等体积 HEPES 缓冲液 (10 mmol HEPES, pH 7.5, 1 mmol EDTA), 20 μ 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将用酚氯仿抽提除去, 水相中的 cDNA 探针将用乙醇沉淀法纯化, 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 (Clontech) 中的 5 $\times 10^5$ 个克隆进行筛选。每个 145 mm 的平板上约 1 万个克隆, 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 (Li et al., 1995)。为纯化克隆, 还要在低浓度的条件下进行第二轮杂交。所得到的克隆用 ABI 377 测序仪测序。

1.2.2 生物信息学方法对 BSG5 基因序列特征进行分析 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 BSG5 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。使用生物学软件对小鼠的 BSG5 基因与人的 KIAA0628 基因的氨基酸序列进行同源性比较。在 Pfam HMM 中的蛋白家族数据库中进行查找, 进行蛋白序列保守功能区结构域分析。用 NCBI 数据库以及 Pfam (www.sanger.ac.uk/software/Pfam) 对蛋白的结构域和可能的功能进行预测。

1.2.3 原位杂交 原位杂交参照文献 (Wilkinson, 1993) 中的方法, 以 2.1 kb 的小鼠 BSG5 编码区的 cDNA 为模板, 用 DIG Nucleic Acid Labeling Kit (Roche) 制备 DIG 标记的 RNA 探针, 分别与 7.5 d、8.5 d、9.0 d、10.5 d 的小鼠胚胎, HH 10 期的鸡胚进行整体胚胎原位杂交。

1 GAGGAATCTG GCCCGGAAGT CTGTGACAGG GAGACGGCCC CGGCCCCCAT TCTTATCCTA

61 GAGTGCCTAC GGATGCCTC CGGGCGGAGC CTCGGGCAGT TTCCGGCCGG GGCCTGCGGC

121 AGCGGTGCGG GCTTTGTGGC CGCCCGGGT TACCCGCTG GATGCTGCTC GGCTTGGTTC

181 CGCGCGTCC TCTCGCCGCC GAGCTCGGA TTCTCCTGTA GGGAGCCAGT GGCAAGAAGG

M M S D E L •

241 GGATCCGGA GCCTGAATTT CTGGGATGA AAATTCACAT AAGATGATGT CAGACGAACT

• T V L E P S A S E S R E F H E D R L A Q •

301 CACAGTCTG GAGCCCTCTG CCTCTGAGTC TCGGAATTT CATGAAGACA GATTGGCGCA

• L L G N P D K Q S L E S P S S Q D G G F •

361 GCTGTTGGG AACCCAGACA AGCAGAGCTT GGAGAGTCCC TCCTCTCAGG ACGGGGGCTT

• T Q M T V T H W K I Q T G D T A Q T C S •

421 CACTCAGATG ACAGTGACCC ACTGGAAGAT TCAAACAGGA GACACAGCCC AAACGTGCAG

• K S G R N P I L N S N L L I L Q R E L I •

481 TAAGTCAGGA AGAAACCCCA TTCTCAATTC AAACCTTCTT ATACTTCAGA GAGAACTCAT

• D A E A R S C G G G G K G F P F N S D V •

541 AGATGCGGAG GCCCGTTCCT GTGGTGGTGG TGGCAAAGGC TTCCCATTTA ATTCAGATGT

• V P H Q I S H T G E K P Y K C D H C G K •

601 AGTTCCGCAT CAGATCTCGC ACACGGGGGA GAAGCCTTAC AAATGTGATC ATTGTGGGAA

• G F S Q T S L L T E H Q R V H T G D R L •

661 AGGCTTCAGT CAGACCTCCC TCCTCACTGA ACACCAGAGA GTCCACACTG GGGACAGACT

• C V C H V C G K D F V H Y A S L R E H Q •

721 CTGTGTGTC CACGTGTGCG GGAAGACTT CGTTCACTAC GCAAGCCTGA GGGAGCATCA

• H V H T G E K P F K C A Q C G K A F C H •

781 GCATGTCCAT ACTGGGAAA AGCCGTCAA GTGTGCACAG TGTGGAAAAG CATTGTGCA

• S S D L L R H Q R V H T R E R P F E C K •

841 CAGTTCAGAC CTGCTTCGAC ACCAGAGGT TCATACCAGG GAGAGGCCTT TTGAATGCAA

• E C G K G F S Q S S L L I R H Q R I H T •

901 AGAGTGTGGG AAGGGTTCA GTCAGAGCTC CTTACTCATT CGCCACAAA GAATTCACAC

• G E R P Y E C N E C G K S F I R S S S L •

961 AGGAGAAAG CCCTATGAGT GTAACGAGTG TGGGAAATCC TTCATTCCGA GTTCAAGCCT

• I R H Y Q I H T E V K Q Y E C R D C G K •

1021 TATCCGCCAT TACCAGATCC ACACAGAAGT TAAACAGTAT GAGTGCAGAG ACTGCGGCAA

• A F R H R S D L I E H Q R I H T G E R P •

1081 AGCCTTCCGA CACCGCTCGG ACCTCATCGA GCACCAGAGG ATTCACACTG GAGAGCGGCC

• F E C Q E C G K A F I R S S K L I Q H Q •

1141 CTTGAGTGT CAGGAGTGTG GGAAGGCCTT TATTCGGAGC TCGAAACTCA TTCAGACCA

• R I H T G E R P Y V C N E C G K R F S Q •

1201 GCGGATTCAC ACGGGAGAGC GGCCTTACGT GTGCAACGAG TGTGGGAAGC GTTTCAGTCA

• T S N F T Q H Q R I H S G E K L Y E C N •

1261 GACATCCAAC TTCACCCAGC ACCAGCGAAT CCACAGTGA GAAAACTCT ATGAGTGTA

• E C G K A F F L S S Y L I R H Q K I H T •

1321 CGAATGTGGG AAGGCTTCT TTCTGAGCTC ATACCTCATA CGACACCAGA AGATCCACAC

• G E R V Y E C K E C G K A F L Q K A H L •

1381 TGGTGAACGG GTGTATGAGT GTAAGGAGTG TGGGAAGGCT TTCCTGCAGA AGGCCACCT

```

      • T E H Q K I H T G D R P F E C K D C G K •
1441 CACTGAGCAT CAGAAGATCC AACTGGAGA CAGACCCTTT GAGTGCAAAG ACTGTGGAA
      • A F I Q S S K L L L H Q I V H T G E K P •
1501 AGCCTTCATC CAGAGTTCTA AGCTATTGCT GCATCAGATT GTTCATACCG GGGAGAAGCC
      • Y V C S Y C G K G F I Q R S N F L Q H Q •
1561 CTATGTGTGC AGTACTGCG GAAAGGCTT TATTCAGAGA TCCAACCTCC TTCAGCACCA
      • K M H T E E K R Y T Y S Q L G K D L T P •
1621 GAAGATGCAC ACGGAAGAGA AGCGCTACAC GTATAGCCAA CTTGGGAAA ATCTTACCCC
      • P P D L V H Q E D L S L S E T S V H L G •
1681 ACCCCAGAT CTGTCCACC AGGAGGATCT TTCCTGAGT GAGACCTCCG TTCATTTAGG
      • E R S T D R E C C G N N L
1741 AGAGAGGTCT ACAGATCGGG AGTGTGTGG AAATAATTTA TAAGACAGTC ACTCTGAGTC
1801 GGATGGTGCT TCTAACAGG CATCCTTAGT ACATATATGG CTTAGAAGAT GCGGATTGTC
1861 ACAGTTTTTC AAGTCATTGA TAGGCTGTTA CTGGTCTGTT GTCTGTCACC AAGGCAGGCA
1921 AGCCAAAGGA CTAGACAGCT GATCTTGAGC TGTTCACTGA GGGAGAACCT CTAGAGAATT
1981 TTCTAGATTT TATTTGAAGT TACAACACAC TCTCAAGTAC TTTCAAAGAA GAACTGTATT
2041 GTTACTTGAG AAATGGTTCC CTGTCTGAAA GGTTAGAGAA AGGGGCTTTT CTCATGCAAA
2101 ACTGCAAGTC GGTTCCTGT TGTCTGGGAC TCATCTCTGG CCTTTGCTTC CTCAACTGTT
2161 TGATGCCTGG TAATGCAGTT GATGCGATTG CCACCGCATC CTCCAGGCTG CTCCCCTCC
2221 TGTCCAGTG TCTCCAGTC GCTCCTGGTC ATTGTGTATC ATTTCTCAGG CGTCATGGCC
2281 AAAAGTGCTT CCCTGGTACT GAGCTGGGTT TGTGTCTTGT GCTTCTGTCG CTCAGTGTTT
2341 ACCACTGCTA CTGTACTTGC TATTAATAC CACAATCATG AACCTTCTCT TTCCTGTTTT
2401 GCTGTTTCCC CGAGGACACT AAGGATGTAT CATTGGCTCC TATGTCCCTA GTGTCTCGCA
2461 CATGACAATA AACAAAACAA AAATCCC

```

图1 小鼠 BSG5 基因 cDNA 序列

编码区序列的上方给出翻译的氨基酸序列。起始密码子 (ATG), 终止密码子 (TAA) 以及多聚腺苷酸序列 (AATAAA) 都用下划线标出。

Fig. 1 cDNA sequence of the mouse BSG5 gene

Coding sequences are shown together with the translated amino acid sequence. Initiative code (ATG) and stop code (TAA) are underlined. The polyadenylation signal sequences (AATAAA) are also underlined.

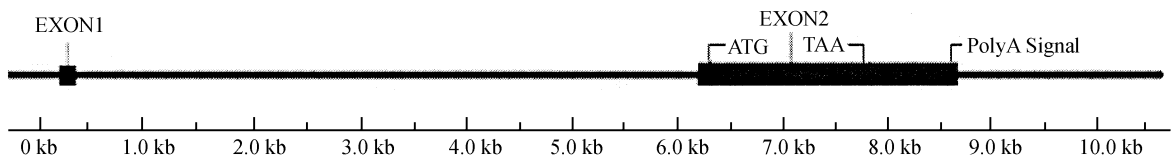


图2 BSG5 基因在染色体上的分布

小鼠 BSG5 基因的 2 个外显子 (实体方框表示) 分布在 15 号染色体上。

Fig. 2 Genomic organization

The mouse BSG5 gene is composed of 2 exons (show in box) distributed on mouse chromosome 15.

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

BSG5, 该基因长 2 487 bp, 编码一长 499 个氨基酸的蛋白 (图 1)。在 NCBI 上对该基因进行分析, 证实是一个新基因。该基因登录号为 AY610936。

2.2 BSG5 基因的生物信息学分析

2.2.1 基因组上的定位 在 Mouse Genome 数据

	1		80
KIAA0628	(1)	MILLSFVSDSNVGTGEKKVTEAWI SEDENSHRTTSDRLTVMELPSPPESEEVHEPRLGELLGNPEGGQSLGSSPSSQDRGCKQ	
<i>BSG5</i>	(1)	-----MMSDELTVLEPSASESREFHEDRLAQLLGNPDKQSLESPPSSQGGFTQ	
Consensus	(1)	SD LTVLE A ES E HE RLA LLGNPD QSL S SQD G Q	
	81		160
KIAA0628	(81)	VTVTHWKIQTGETAQVCTKSGRNHILNSDLLLQRELIEGEANPCD I CGKTFTFNSDLVVRHRI SHAGEKPYTCQCGCKGF	
<i>BSG5</i>	(49)	MTVTHWKIQTGDTAQTCSSKSGRNP I LNSNLL I LQREL I DAEARSCGGGGKGF FNSDVVPHQ I SHTGEKPYKCDHCGKGF	
Consensus	(81)	MTVTHWKIQTGDTAQ C SKSGRN ILNS LL I LQREL I DAEA C GK F FNSDLV H I SH GEKPY CD CGKGF	
	161		240
KIAA0628	(161)	QSSHLM EHQRIHTGERLYVCNVCGKDF I HYSGL I EHQRVHSGEKPFKCAQCGKAFCHSSDL I RHQRVHTRERPFECCK	
<i>BSG5</i>	(129)	SQTSLLTEHQRVHTGDRLCVCHVCGKDFVHYASLREHQHVHTGEKPFKCAQCGKAFCHSSDLLRHQRVHTRERPFECCK	
Consensus	(161)	QSS L EHQRIHTGDRL VC VCGKDF I HYA L EQV VHSGEKPFKCAQCGKAFCHSSDL I RHQRVHTRERPFECCK	
	241		320
KIAA0628	(241)	GKGFSSQSL I RHQRIHTGERPYECNECGKSF I RSSSL I RHYQ I HTEVKQYECCKGKAFRHRSDL I EHQRIHTGERPFE	
<i>BSG5</i>	(209)	GKGFSSQSL I RHQRIHTGERPYECNECGKSF I RSSSL I RHYQ I HTEVKQYECRDCGKAFRHRSDL I EHQRIHTGERPFE	
Consensus	(241)	GKGFSSQSL I RHQRIHTGERPYECNECGKSF I RSSSL I RHYQ I HTEVKQYECCKGKAFRHRSDL I EHQRIHTGERPFE	
	321		400
KIAA0628	(321)	CNECGKAF I RSSKL I QHQRIHTGERPYVCNECGKRFSQTSNFTQHQR I HTGEKLYECNECGKAFFLSSYL I RHQK I HTGE	
<i>BSG5</i>	(289)	CQCEGKAF I RSSKL I QHQRIHTGERPYVCNECGKRFSQTSNFTQHQR I HSGEKLYECNECGKAFFLSSYL I RHQK I HTGE	
Consensus	(321)	CNECGKAF I RSSKL I QHQRIHTGERPYVCNECGKRFSQTSNFTQHQR I HSGEKLYECNECGKAFFLSSYL I RHQK I HTGE	
	401		480
KIAA0628	(401)	RVYECKEKGKAF LQKAHLTEHQK I HSGDRPFECCKGKAF I QSSKLLLHQ I IHTGEKPYVCSYCGKGF I QRSNFLQHQK I	
<i>BSG5</i>	(369)	RVYECKEKGKAF LQKAHLTEHQK I HTGDRPFECCKGKAF I QSSKLLLHQ I VHTGEKPYVCSYCGKGF I QRSNFLQHQK I	
Consensus	(401)	RVYECKEKGKAF LQKAHLTEHQK I HSGDRPFECCKGKAF I QSSKLLLHQ I IHTGEKPYVCSYCGKGF I QRSNFLQHQK I	
	481		536
KIAA0628	(481)	HTEEKLYECSQYGRDFNSTTNVKNQRVHQEGLSLSKAP I HLGERSVDKGEHTGNL	
<i>BSG5</i>	(449)	HTEEKRYTYSQLGKDLTPPPDL-----VHQEDLSLSETSVHLGERSTDRECCGNL	
Consensus	(481)	HTEEK Y SQ GKD L VHQE LSLS IHLGERS DK NL	

图 3 小鼠 *BSG5* 和人的 KIAA0628 氨基酸序列同源性的比对
阴影部分为同源序列, 它们具有 81.9% 的相似性。

Fig. 3 Alignments of the amino acid sequence between KIAA0628 and mouse *BSG5* gene
They share 81.9% similarity.

库中做 BLAST 的分析得到 *BSG5* 定位在小鼠的第 15 号染色体上, 包含 2 个外显子 (图 2) 它们分别位于 291 - 395 bp、6 367 - 8 644 bp 处。对 *BSG5* 基因的编码框 (Coding sequence, CDS) 进行预测, 得到一个 1 497 bp 长的编码区序列, 位于 *BSG5* cDNA 的 284 bp - 1 780 bp 的核苷酸位置。含有起始密码子 ATG 和终止密码子 TAA, 是一个完整的编码框。并且在终止密码子后面, 找到了 Poly A 的信号 AATAAA (图 1)。

2.2.2 蛋白结构分析 对小鼠 *BSG5* 基因和人 KIAA0628 基因的氨基酸序列的同源性比较, 结果

显示两者的同源性达 81.9% (图 3)。

对小鼠 *BSG5* 基因结构域的分析显示该基因含有 12 个 C2H2 型的锌指结构域, 每个锌指结构域均含有 23 个氨基酸, 它们所在的氨基酸位置分别为: 119 - 141、147 - 169、175 - 197、203 - 225、231 - 253、259 - 281、287 - 309、315 - 337、343 - 365、371 - 393、399 - 421、427 - 449。这些结构域提示 *BSG5* 基因是一个可能的转录因子 (图 4)。

2.3 原位杂交

原位杂交结果显示, 在 7.5 dpc、8.5 dpc、

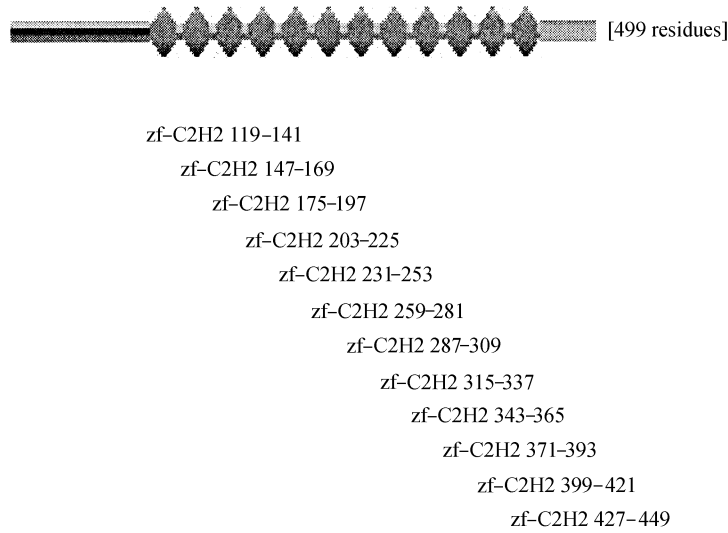


图4 BSG5 基因的结构域

显示 12 个连续的是 C2H2 型锌指结构域。

Fig. 4 The structure domains of mouse BSG5

There are consecutive 12 bars represent 12 C2H2 type zinc finger domains.



图5 BSG5 基因在不同时期小鼠中的原位杂交

BSG5 基因在小鼠早期胚胎的头部 (箭头所示) 特异表达。A: 7.5 dpc。B: 8.5 dpc。C、D: 9.0 dpc。E: 10.5 dpc。

Fig. 5 Whole-mount in situ hybridization of early mouse embryo at different stage

BSG5 gene is expressed specifically in mouse embryo head (see arrow). A: 7.5 dpc. B: 8.5 dpc. C, D: 9.0 dpc. E: 10.5 dpc.

9.0 dpc、10.5 dpc 的小鼠胚胎中, BSG5 基因主要在头部表达 (图 5)。另外发现在 9.0 dpc、10.5 dpc 的小鼠胚胎中, 尾部及肢芽也有明显表达 (图 5: C, E)。在 HH 10 期的鸡胚中 (图 6), BSG5 也主要在头部表达。

3 讨论

锌指蛋白家族是一个巨大的超基因家族, 它们广泛的存在于各种物种中, 包括酵母、果蝇、爪蟾、鼠和人。所有的锌指蛋白都有一个 28 - 30 个氨基酸长的重复序列, 这些重复序列都具有固定的保守氨基酸。其中很多成员是重要的细胞调控因子。本文利用消减差异筛选的方法克隆到一个新的锌指蛋白基因, 此基因编码的蛋白质全长 499 个氨基酸, 含 12 个 C2H2 型的锌指结构域。锌指蛋白

可以按照其所拥有的锌指结构和数目的不同分成两大类: 第一大类的锌指蛋白含有 5 个以下的锌指结构域, 如 ATZ 和 WTI (Hajra et al., 2002), 它们通常参与细胞的增殖和分化; 而第二大类的锌指蛋白则含有 5 个以上锌指结构域, 如 TFIIIA, ZFP28 (周亮等, 2002) 和本文中发现的 BSG5。虽然第二类锌指蛋白的数目多于第一类锌指蛋白, 但是除了 TFIIIA 以外, 人们对于其它此类锌指蛋白的功能还了解不多。BSG5 基因所编码的锌指蛋白具有 12 个 C2H2 型锌指结构, 属于后一类基因。它在小鼠早期胚胎、鸡胚胎脑部的特异表达情况表明此基因对于脑的早期发育有非常重要的意义。自小鼠胚胎第 7.5 d 开始, BSG5 在头部特异表达, 并且在 8.0 d、9.0 d、10.5 d 持续在头部有高表达, 在胚胎第 9.0 d 时开始在尾芽也有表达, 第 10.5 d 时

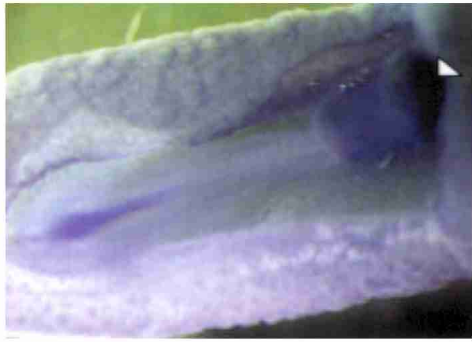


图 6 *BSG5* 基因在 HH 10 期鸡胚中的原位杂交
BSG5 基因在鸡胚头部特异表达 (白色三角所示)。

Fig. 6 Whole-mount in situ hybridization of chicken embryo at HH 10 stage

BSG5 gene is expressed specifically in chicken embryo head (show in white).

在尾芽及肢芽均有表达。胚胎在第 10.5 d 时, 很多器官正在形成 (Rugh, 1990), 此时期的胚胎很多器官尚未发育完全, *BSG5* 在不同空间时间上的表达变化提示 *BSG5* 可能在胚胎发育早期对于头部发育起至关重要的作用, 另外 *BSG5* 在尾芽及肢芽的表达表明对后期胚胎发育中四肢和尾部的形成也起重要作用。因此对 *BSG5* 基因的进一步研究不但有助于增加对于第二类锌指蛋白的功能的了解, 并且可以进一步揭示脑的早期发育的分子生物学机制。

鸡胚用于发育模式生物研究的最主要优点是易于操作和观察。鸡胚胎在 HH 5 期 (20 - 22 h) 时, 神经板已经形成, 头的雏形也基本可见。在 HH 7 期后脑形成。在 HH 10 期, 三胚层形成, 体轴也基本形成, 神经脊开始迁移。以 DIG 标记的 *BSG5* 的 RNA 为探针, 对 HH 10 期 (36 - 42 h) 的鸡胚做整体胚胎的原位杂交, 结果表明 *BSG5* 基因主要在神经系统中大量表达。

BSG5 基因在不同种属中显著的表达保守性表明 *BSG5* 基因在神经系统即脑的发育中有重要作

用。*BSG5* 基因的结构与表达特征预示着它编码的是一个具 DNA 结合功能的转录调控因子。其具体的功能及蛋白表达特点还有待进一步的研究。

参考文献 (References)

- Armstrong JA, Bieker JJ, Emerson BM, 1998. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLf *in vitro*. *Cell* 95 (1): 93 - 104.
- Cai WQ, Li HB, 1999. *Developmental Neurobiology*. Beijing: Science Press, 21 - 25 (In Chinese).
- Hajra KM, Chen D, Fearon ER, 2002. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res.* 62 (6): 1 613 - 1 618.
- Lania L, Majello B, De Luca P, 1997. Transcriptional regulation by the Sp family proteins. *Int. J. Biol.* 29 (12): 1 313 - 1 323.
- Li YP, Alexander M, Yelick P, Chen W, Stashenko P, 1995. Cloning and complete coding sequence of a novel human cathepsin expressed in giant-cells of osteoclastomas. *J. Bone Min. Res.* 10 (8): 1 197 - 1 202.
- Li YP, Chen W, Stashenko P, 1996. Molecular cloning and characterization of a putative novel human osteoclast-specific 116-kDa vacuolar proton pump subunit. *Biochem. Biophys. Res. Comm.* 218 - 813.
- May PC, Johnson SA, Poirier J, Lampert-Etchells M, Finch CE, 1989. Altered gene expression in Alzheimer's disease brain tissue. *Can. J. Neurol. Sci.* 16 (4 Suppl.): 473 - 476.
- Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F, 1995. Defective hematopoiesis in fetal liver resulting from inactivation of the EKLf gene. *Nature* 375 (6529): 316 - 318.
- Parasuraman R, Greenwood PM, Sunderland T, 2002. The apolipoprotein E gene, attention, and brain function. *Neuropsychology* 16 (2): 254 - 274.
- Petering DH, Huang M, Moteki S, Shaw CF, 2000. Cadmium and lead interactions with transcription factor A from *Xenopus laevis*: a model for zinc finger protein reactions with toxic metal ions and metallothionein. *Mar. Environ. Res.* 50 (1 - 5): 89 - 92.
- Rugh R, 1990. *The Mouse*. New York: Oxford University Press.
- Wilkinson DG, 1993. *In situ hybridization*. In: Stern CD, Holland PWH ed. *Essential Development Biology: A Practical Approach*. Oxford, UK: IRL Press, 257 - 274.
- Zhou L, Yuan WZ, Wang YQ, Pi HL, Zhu CB, Wu XS, 2002. Studies on cloning and expression of a human novel zinc finger protein ZNF28. *J. Human Norm. Uni. Nat. Sci.* 25 (3): 72 - 76 (In Chinese).
- 蔡文琴, 李海标, 1999. *发育神经生物学*. 北京: 科学出版社, 21 - 52.
- 周亮, 袁葵周, 王跃群, 皮花亮, 朱传炳, 吴秀山, 2002. 一个人类锌指蛋白家族新基因 *ZFP28* 的克隆及表达. *湖南师范大学学报 (自然科学版)* 25 (3): 72 - 76.