

人类心脏发育候选基因 ZNF569的生物信息学分析

黄欣琼¹, 唐显庆¹, 张晓红¹, 石金凤¹, 李美香¹, 吴秀山²

(1. 南华大学医学院组织胚胎学教研室, 湖南省衡阳市 421001;

2. 湖南师范大学生命科学院心脏发育研究中心, 湖南省长沙市 410081)

[关键词] ZNF569基因; 心脏发育; KRAB/C2H2

[摘要] 目的 分析人类心脏发育候选基因 ZNF569的生物信息学, 对其功能研究提供指导意见。方法 利用生物信息学网站及生物软件对人类心脏发育候选基因 ZNF569进行生物信息学分析。结果 ZNF569是一个新的人类 KRAB/C2H2型锌指蛋白基因。此基因在基因组 DNA 上长约 56 28 kb 定位于 19q13.12。开放阅读框全长 2061个碱基, 含 6个外显子, 5个内含子, 编码一个长 686个氨基酸残基的蛋白质。该蛋白包含一个 KRAB结构域和 18个锌指结构域, 进化分析表明各种真核生物中都存在 ZNF569的同源蛋白, 而 ZNF569蛋白与鼠 ZNF74蛋白的相似度达到 94%, 它们之间的蛋白质序列保守度更高。ZNF569本身 18个锌指结构域蛋白质序列也高度保守。结论 ZNF569及其同源蛋白构成了一个在进化中高度保守的新的 KRAB-ZF基因, 而它的功能尚未见报道。

[中图分类号] Q785

[文献标识码] A

Biinformatical Analysis on Heart Developmental Candidate Gene ZNF569

HUANG Xin-Qiong¹, TANG Xian-Qing¹, ZHANG Xiaohong¹, SHI Jin-Feng¹, LIM Erxiang¹, and WU Xiushan²

(1. Department of Histology & Embryology, University of South China, Hengyang 421001, China; 2. Center for Heart Development, College of Life Sciences, Hunan Normal University, Changsha 410081, China)

[KEY WORDS] ZNF569 Gene, Heart Development, KRAB/C2H2

[ABSTRACT] **Aim** To analyse biinformatics on heart developmental candidate gene ZNF569 providing guiding opinion for functional study of ZNF569. **Methods** Biinformatical analysis on ZNF569 was made by utilizing Biinformatical Web and software. **Results** Protein containing KRAB/C2H2 is an important transcriptional factor. A alignment between the cDNA sequence and human genome indicates that ZNF569 is identical on chromosome 19q13.12, spanning approximately 56.28 kb in reverse manner on the genome and organized into 6 exons. ZNF569 encodes a putative protein of 686 amino acids. The predicted amino acid sequence of ZNF569 contains an N-terminal Krüppel-associated box (KRAB) domain and a series of 18 C2H2 zinc finger motifs that extend to the end of the protein sequence. A search of published DNA databases for sequences similar to that of ZNF569 demonstrated that they had varying degrees of similarity to a number of previously identified C2H2-type zinc-fingers protein. Sequence analysis and database comparison indicate that the predicted protein contains 18 different C2H2 zinc-finger domains in tandem arrays, characteristic of transcription factor proteins of this family. **Conclusion** A search of analyzed evolutionary relationship among different zinc-finger proteins with phylogenetic tree analysis show that ZNF569 is one of the conserved proteins during evolution.

在心血管系统疾病的研究中已发现, 某些基因在时空上的表达紊乱可导致先天性心脏病, 而对部分心血管系统病症高发人群的研究也证明部分后天形成的心脏病与遗传因素有关。研究现有信号途径对于心脏发育的关系, 发现新的参与了心脏发育过程的信号途径调控蛋白就成为一项极有意义的工作。锌指结构是很多转录因子用来接触 DNA 的一种结构。人类基因组估计有 300~700个锌指蛋白,

它们都具有锌指结构域, 而 C2H2型占绝大多数, 此种锌指蛋白含有大约 28个氨基酸的重复锌指序列, 并由 2个半胱氨酸残基和 2个组氨酸残基整合一个锌离子构成, 此结构域能结合到特异的 DNA 识别位点, 激活或抑制基因的表达^[1]。有 1/3的 C2H2型锌指蛋白同时具有一个 KRAB框, 此结构域位于锌指蛋白的氮末端非锌指区域, 由 A框和 B框组成, 折叠成 α -2螺旋参与蛋白质和蛋白质之间的作用, 这种作用大部分是通过与转录中介因子 2TF1相互作用实现的^[2]。研究表明锌指蛋白在生物发育过程中充当持家基因或调节因子的角色, 起着多方面的作用, 如 ZNF174是某些生长因子基因的负调节子^[3], AJ18参与骨的发育^[4]。许多锌指蛋白参与某些肿瘤的形成, 如 WT1^[5]。

[收稿日期] 2009-04-12 [修回日期] 2009-05-05

[基金项目] 国家自然科学基金(30270684)和湖南省教育厅一般项目资助

[作者简介] 黄欣琼, 硕士, 讲师, 主要从事心脏发育研究, E-mail 为 wame904@yahoo.com.cn。通讯作者唐显庆, 副教授, 主要从事心血管组织学研究。张晓红, 副教授, 主要从事肿瘤药理研究。

生物信息学是一门交叉科学,它包含了生物信息的获取、处理、存储、分发、分析和解释等在内的所有方面,综合运用数学、计算机科学和生物学的各种工具来阐明和理解大量数据所包含的生物学意义。目前利用生物信息学手段对未知基因的功能研究能够提供很好的前期指导。本研究对人类心脏发育候选基因 ZNF569的生物信息学进行了深入分析。

1 材料和方法

1.1 主要试剂及仪器

UNIQ10柱式 DNA 胶回收纯化试剂盒、RNaseA 和 pMD18-T 载体购自上海生工生物工程技术有限公司, Taq DNA 聚合酶、随机引物标记试剂盒、cDNA Synthesis Kits, cDNA PCR Library Kits, Ligation kit 2.0 试剂盒及限制性内切酶购自 TaKaRa 公司, SMART RACE cDNA Amplification Kit 购自 Clontech 公司,引物由博亚公司合成。移液器购自 Eppendorf 公司,微型离心机购自 Fresco 公司, SPRT001 型 PCR 仪购自 ThermoHyaid 公司, UV-2802 型紫外可见分光光度计。

1.2 cDNA 文库的构建

用传统酚-氯仿抽提的方法从人类胚胎的心脏组织中提取总 RNA 500 μg 用 Rapid mRNA TM Purification Kit (Amresco) 分离纯化 mRNA, 以此 mRNA 为模板, 用 cDNA Synthesis Kits 和 cDNA PCR Library Kits 逆转录得到第 1 条 cDNA 链, 在此基础上合成第 2 条 cDNA 链并使其平滑化, 经酚-氯仿抽提、连接、cDNA 扩增, 获得人类胚胎心脏组织的 cDNA 文库。

1.3 ZNF569 基因的克隆

以 mRNA 逆转录所得的 cDNA 文库为模板, 根据 C2H2 型锌指基因及相应未知 EST 序列设计引物, P1 引物序列 5'-AAG AGG AAA TGA CTG AGT CCC AGG G-3', P2 引物序列 5'-TCC TCT GGT GTC TAA CAA GGT GCG A-3', PCR 反应条件: 一个循环 94°C、4 min; 接下来 30 个循环, 前 29 个循环每个循环 94°C 30 s 及 52°C 30 s 一个循环 72°C 2 min; 最后一个循环 72°C 8 min; 所得 PCR 产物与克隆载体 pMD18-T 16°C 连接过夜, 转化大肠杆菌 DH5 涂板, 37°C 生长过夜, 挑单克隆菌培养, 提取质粒, 37°C 酶切鉴定后, 测序分析。

1.4 5'-RACE-PCR

采用快速扩增 cDNA 末端 (RACE) 技术在获得基因片段的基础上设计 5-RACE 基因特异性嵌套引

物 R1 和 R2 其序列分别为 5'-TCC TGA GTG AAG TCG ATA GCC A-3' 和 5'-TCC TCT GGT GTC TAA CAA GGT GCG A-3'。用心脏组织 mRNA 逆转录所得的 cDNA 文库为模板, 嵌套 PCR 引物进行 cDNA 末端扩增。用琼脂糖凝胶电泳法回收扩增 PCR 产物, 用 T 载体连接, 转化大肠杆菌 DH5。转化后经酶切鉴定和进一步测序分析, 最终获得其全长 cDNA 序列。

1.5 生物信息学分析方法

Blast 程序 <http://www.ncbi.nlm.nih.gov/blast> 用于核酸/蛋白质序列匹配查找; <http://www.ncbi.nlm.nih.gov/gorf> 用于开放阅读框 (ORF) 分析; <http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html> 用于染色体定位分析; UCSC 程序 <http://genome.ucsc.edu> 用于染色体定位分析和确定内含子/外显子交界; SMART 分析蛋白质结构域 <http://smart.embl-heidelberg.de> Motif Scan 程序分析蛋白质模体和结构域 <http://scansite.mit.edu>; DNA star 软件用于 DNA 序列、蛋白质性质及进化关系分析^[5]。

2 结果

2.1 ZNF569 基因克隆分析

利用 KRAB/C2H2 型基因设计 PCR 反应引物扩增获得部分基因片段后, 在其 5' 端设计一个嵌套 5'-RACE PCR 的引物进行 5'-RACE 扩增, 结果获得一个长度为 558 bp 的 PCR 产物 (图 1)。经纯化, T 载体连接, 转化, 测序, 从而获得完整的基因 cDNA 序列。该基因全长为 2061 bp 含有 6 个外显子, 全基因在基因组序列上长达 56.28 kb 其 3' 端有一个加尾信号 AATAAA。此基因共编码 686 个氨基酸, 蛋白质大小约 79581.85 Da (约 80 kDa)。起始密码子开始于 559 终止密码子位于 2620。此基因编码的蛋白质在氮末端有 1 个 KRAB 框, 18 个 C2H2 型锌指 (ZF), 属于锌指蛋白家族的 KRABII/C2H2 亚家族。根据核苷酸和蛋白质的特征, 经国际基因命名委员会批准, 命名为 ZNF569。

2.2 ZNF569 的进化保守性分析

利用 ZNF569 的氨基酸序列在 NCBI 数据库中进行 BLASTp 搜索, 发现各种真核生物中都存在 ZNF569 的同源蛋白, ZNF470 (H. sapiens)、ZNF585B (H. sapiens)、ZNF84 (H. sapiens)、ZNF43 (H. sapiens)、ZNF624 (H. sapiens)、ZNF74 (M. musculus) 及 ZNF184 (M. musculus) 等的 KRAB 框与 ZNF569 的

KRAB框蛋白质序列高度保守,其中第 5位氨基酸的 D,第 10位氨基酸的 E,第 15位氨基酸的 E,第 16位氨基酸的 Q,第 22位氨基酸的 Q,第 25位氨基酸的 L完全相同(图 2)。ZNF569蛋白与鼠 ZNF74蛋白的相似度达到 94%,它们之间的蛋白质序列保守度更高(图 3)。ZNF569基因 18个锌指结构蛋白质序列也高度保守,这 18个锌指的第 3位氨基酸的 C,第 6位氨基酸的 C,第 8位氨基酸的 K,第 10位氨基酸的 F,第 19位氨基酸的 H,第 23位氨基酸的 H完全相同(图 4)。进化树分析说明,ZNF569及其同源蛋白构成了一个在进化中高度保守的新的 KRAB-ZF 基因,而它们的功能尚未见报道。

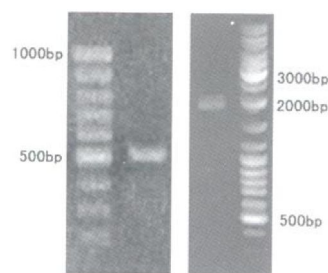


图 1. ZNF569 5'-RACE-PCR 扩增结果 左图为 558 bp 的 5'-RACE-PCR 产物纯化结果,右图为 ZNF569 ORF PCR 产物纯化结果。

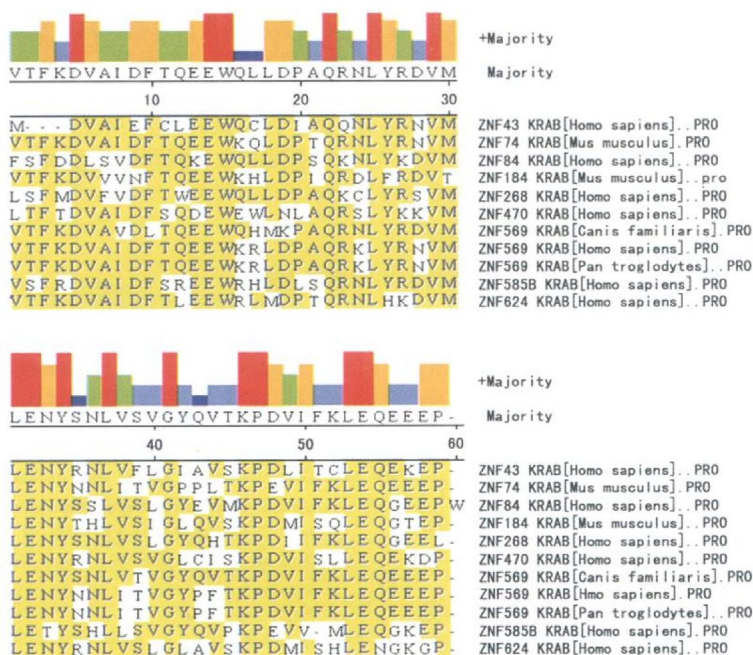


图 2. ZNF569 的 KRAB 框与其它锌指基因 KRAB 框的多序列比对图

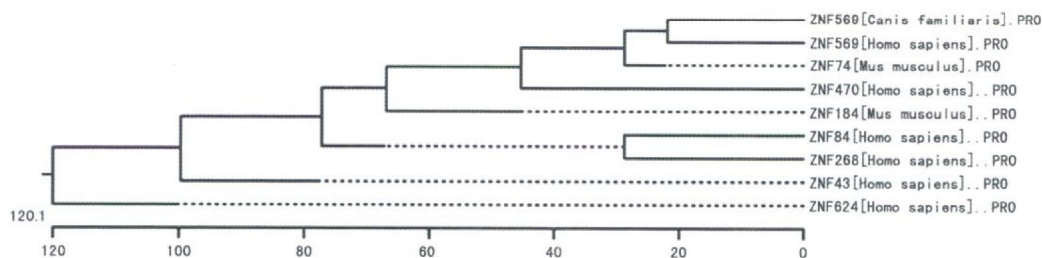


图 3. ZNF569 的 Un-rooted 进化树分析

3 讨论

锌指蛋白基因家族成员在心脏发育过程中起重要作用,包括 Nkx2-5, GATA-4, GATA-5及 GATA-6等。本实验室利用基因的同源性和 5'-RACE,克隆了一个新的 KRAB/C2H2型的锌指蛋白基因

ZNF569,它编码的蛋白含 1个 KRAB 框和 18个 C2H2型锌指结构。表达模式是研究基因功能的一条重要线索,ZNF569 mRNA 有选择性地心脏组织和骨骼肌表达。在胚胎时期 ZNF569可在心脏、大脑中检测到强的表达,表明 ZNF569很可能在早

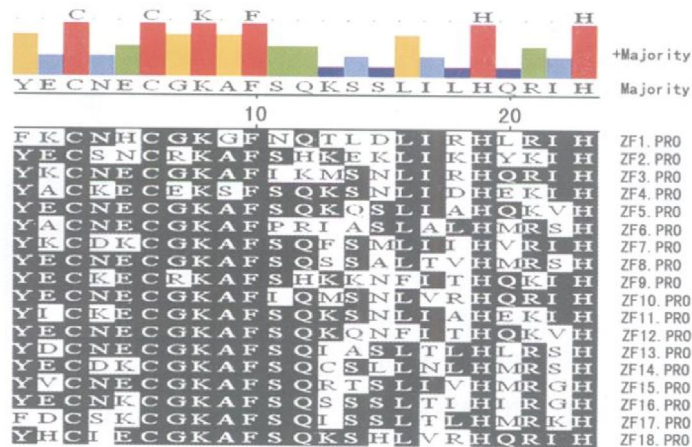


图 4 ZNF569蛋白质 18个锌指多序列比对图

期胚胎分化和胚胎的形成以及心脏的晚期发育或心脏功能的发挥中起着重要的作用。ZNF569在心脏发育和心脏发病机制中的重要性有待进一步研究。生物信息学中序列比对的主要意义有从相互重叠的序列片段中重构 DNA 的完整序列。ZNF569正是通过这样的方式得到的电子全序列, 然后经实验证明的。序列对比的另一个意义就是在各种试验条件下从探测数据中决定物理和基因图存贮, 遍历和比较数据库中的 DNA 序列比较两个或多个序列的相似性在数据库中搜索相关序列和子序列寻找核苷酸的连续产生模式找出蛋白质和 DNA 序列中的信息成分。序列比对考虑了 DNA 序列的生物学特性。经过多序列比对分析说明各种真核生物中都存在 ZNF569的同源蛋白, 而 ZNF569蛋白质与鼠 ZNF74蛋白的相似度达到 94%, 它们之间的蛋白质序列保守度更高。ZNF569中 18个锌指结构蛋白质序列也高度保守。此外分子进化和比较基因组学也是生物

学主要内容。进化树分析说明, ZNF569及其同源蛋白构成了一个在进化中高度保守的新的 KRAB-ZF 基因, 而它们的功能尚未见报道。这些生物信息学分析结果对 ZNF569 功能研究提供重要的指导意见。

[参考文献]

- [1] Tian C. KRAB-type zinc-finger protein A pak specifically regulates p53-dependent apoptosis [J]. *Nat Cell Biol* 2008 **11** (5): 580-591
- [2] Thomas H, Emerson RO. Evolution of C2H2-zinc finger genes revisited [J]. *BMC Evol Biol* 2009 **4** (9): 51-58
- [3] Hirasawa R, Feil R. A KRAB domain zinc finger protein in imprinting and disease [J]. *Dev Cell* 2008 **15** (4): 487-488
- [4] Garcia-Garcia M J, Shibata M, Anderson KV. Chatg, a KRAB zinc-finger protein, regulates convergent extension in the mouse embryo [J]. *Development* 2008 **135** (18): 3 053-062
- [5] Xinqiong Huang. ZNF569, a novel KRAB-containing zinc finger protein, suppresses MAPK signaling pathway [J]. *Biochem Biophys Res Commun*, 2006 **346**: 621-628

(此文编辑 文玉珊)