

· 基础论著 ·

HBxAg 蛋白反式激活基因 5 的生物信息学分析及真核表达载体的构建

蓝贤勇 安纪红 成军 武会娟 张丽娟 陶明亮 袁菊 张黎颖 洪源 毛羽 陈宏

【摘要】 目的 对 HBxAg 蛋白反式激活基因 5 (XTP5) 进行生物信息学分析, 并构建其真核表达载体。**方法** 利用生物信息学软件对 XTP5 基因进行 CpG 岛、启动子活性、转录终止信号和氨基酸的一、二级结构以及信号肽、跨膜特性和功能基序进行预测。构建 XTP5 基因的真核表达载体 pcDNA3.1-XTP5 和 pGBKT7-XTP5。**结果** 生物信息学分析发现 XTP5 基因有 200 bp 左右的 CpG 岛, 有高启动子活性序列、转录终止信号位点, 二级结构以 Helix 为主, 存在信号肽的概率为 0.0000A, 无跨膜区域信号, 存在数个功能基序。成功构建了 XTP5 基因的真核表达载体 pcDNA3.1-XTP5 和 pGBKT7-XTP5。**结论** 为深入研究 XTP5 结构和功能、进一步探讨乙型肝炎病毒 (hepatitis B virus, HBV) 发病机制创造了条件。

【关键词】 乙型肝炎病毒 X 蛋白; XTP5 基因; 次要组织相容性抗原基因; 生物信息学; 表达载体

Bioinformatics prediction and construction of eukaryotic expression vectors for human gene 5 transactivated by HBxAg LAN Xian-yong*, CHENG Jun, WU Hui-juan, ZHANG Li-juan, TAO Ming-liang, YUAN Ju, ZHANG Li-ying, HONG Yuan, MAO Yu, CHEN Hong. *Institute of Infectious Diseases, Beijing Ditan Hospital, Beijing 100011, China

Corresponding author: CHENG Jun, Email: cj@genetherapy.com.cn

【Abstract】 Objective To analyze XTP5 gene (minor histocompatibility antigen HA-8) by bioinformatics and to construct eukaryotic expression vectors. **Methods** Bioinformatics software was applied to analyze XTP5 gene as followings: CpG island, promoter activity, transcription terminat signal, the first and second structure of amino acids, signal peptide, prediction for characteristics of transmembrane and functional motif. Moreover, eukaryotic expression vectors of XTP5 gene were contructed. **Results** CpG island was found with a length of 200 bp. High promoter activity and transcription terminat signal site was also found. The main second structure of amino acids is Helix. The probability of having signal peptide is 0.0000A. Several functional

作者单位: 100011 北京, 北京地坛医院传染病研究所 (蓝贤勇、成军、武会娟、张丽娟、陶明亮、袁菊、张黎颖、洪源、毛羽); 内蒙古医院感染科 (安纪红); 西北农林科技大学动物科技学院 (蓝贤勇、武会娟、张丽娟、陶明亮、袁菊、陈宏); 徐州师范大学细胞与分子生物学研究所 (陈宏)

通讯作者: 成军 Email: cj@genetherapy.com.cn

motifs were found but no transmembrane signal was found. The eukaryotic expression recombinants (pcDNA3.1-XTP5 and pGBKT7-XTP5) were constructed successfully.

Conclusions This study create conditions for deeply studying the pathogenesis of HBV.

【Key words】 Hepatitis B virus X antigen; XTP5 gene; Minor histocompatibility antigen HA-8; Bioinformatics; Expression vector

2002 年,本课题组利用抑制性消减杂交(suppression subtractive hybridization, SSH)及生物信息学(bioinformatics)技术筛选并克隆乙型肝炎病毒(HBV)X 蛋白反式激活基因 5, 并命名为 XTP5, GenBank 登录号为 AF490254^[1-3]。2006 年, XTP5 基因被诠释为:“次要组织相容性抗原 HA-8 基因”(Minor histocompatibility antigene HA-8 gene, mHA-8)^[4,6], 在造血干细胞移植中扮演主要角色。随着“结构基因组计划”的实施和计算机网络的发展,生物信息学成为功能基因组研究的重要手段。由于国内外关于 XTP5 基因的研究甚少,为进一步探讨 mHA-8 的生物学功能,本研究拟利用生物信息学方法预测其生物学功能,并运用基因工程技术构建两个真核表达载体,期望为深入研究 mHA-8、HBxAg 蛋白的调节作用和 HBV 发病机制创造条件。

材料和方法

一、材料和试剂

生物信息学分析主要通过在线网络, HepG2 细胞及感受态 *E. coli* DH5 α 为本室保存, pcDNA3.1 真核表达载体、pGBKT7 空酵母表达载体分别购自 Invitrogen 和 Clontech 公司, mRNA Purification 试剂盒购自 Amersham Pharmacia Biotech 公司, PCR-Select cDNA Subtraction 试剂盒、50 \times PCR Enzyme Mix、Advantage PCR Cloning 试剂盒购自 Clontech 公司, High Pure PCR Product Purification 试剂盒购自 Boehringer Mannheim 公司, Taq 酶、T4 DNA 连接酶、pGEM-T、IPTG 及 X-Gal 购自 Promega 公司, 玻璃奶回收试剂盒购自博大泰克公司, 引物合成及 DNA 序列测定均由 Invitrogen 公司完成, 限制性内切酶(*Eco* R I、*Bam* H I、*Sac* I、*Nde* I 和 *Pst* I) 和 Marker DL15000、DL2000 均购自 Takara 生物公司。

二、方法

1. XTP5 基因的生物信息学分析: 利用在线软件 www.ebi.ac.uk/emboss/cpg-plot/ 分析 XTP5 基因 5'-侧翼区的 CpG 岛; 利用 <http://bimas.dcrt.nih.gov/mol-bio/proscan> 和 <http://bioportal.bic.nus.edu.sg/tres/> 同时分析启动子活性; 利用 http://125.itba.mi.cnr.it/~webgene/wwwHC_polya.html 分析转录终止位点; 利用 <http://www.ncbi.nlm.nih.gov/mapview/> 分析染色体定位; 利用 <http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi> 判断该基因是否存在 Unigene 号; 利用 www.kazusa.or.jp/codon/countcodon.html 研究密码子使用频率; 利用 [ProtParamwww.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html) 预测蛋白质一级结构; 利用 www.ncbi.nlm.nih.gov/

Blastp/对氨基酸序列进行比对分析;利用 nnpredict(www.cmpharm.ucsf.edu/~nomi/nnpredict.html) 预测蛋白质二级结构;利用 SignalP 信号肽预测工具 <http://www.cbs.dtu.dk/services/SignalP/> 分析信号肽;利用 http://www.ch.embnet.org/software/TMPRED_form.html 作 Tmpred 跨膜结构分析;利用 http://myhits.isb-sib.ch/cgi-bin/protein_query 分析氨基酸序列的功能基序^[7,8]。

2. 表达载体的构建: (1) HepG2 细胞的培养、mRNA 的分离与 cDNA 的合成:复苏、培养并收获 HepG2 细胞, mRNA 纯化、反转录采用相应试剂盒,操作按说明书进行;(2) XTP5 基因的 RT-PCR 扩增:根据 XTP5 基因的完整开放阅读框和 pcDNA3.1 真核表达载体、pGBKT7 酵母表达载体设计 1 对共用引物:上游引物 Pf:5'-CGGAATTCATGTGGGAGATTTTAAGAAGA AAAG-3' (划线部分为 *Eco* R I 位点);下游引物 Pr:5'-GCGGATCCTGTGCTCAGTTTTTC AAGTAGAAT-3' (划线部分为 *Bam* H I 位点)。PCR 反应体系 25 μ l:双蒸水 17.3 μ l, 10 \times 缓冲液(20 mmol/L MgCl₂) 2.5 μ l, 2.5 mmol/L dNTP 2.5 μ l、10 μ mmol/L 上游引物 1.0 μ l、10 μ mmol/L 下游引物 1.0 μ l、cDNA 模板 0.2 μ l、Taq DNA 聚合酶(5 U/ μ l) 0.5 μ l;PCR 扩增条件:95 $^{\circ}$ C 预变性 5 min, 94 $^{\circ}$ C 变性 30 s, 55 $^{\circ}$ C 退火 45 s, 72 $^{\circ}$ C 延伸 100 s, 35 个循环后, 72 $^{\circ}$ C 保温 10 min, 4 $^{\circ}$ C 保存;最后, 1.0% 琼脂糖凝胶电泳检测;(3) pcDNA3.1-XTP5 和 pGBKT7-XTP5 重组表达质粒的构建与鉴定:用玻璃奶回收试剂盒回收 XTP5 基因的 PCR 产物, 将其与 pGEM-T 载体按 10:1 混合, 在 T4 DNA 连接酶作用下 16 $^{\circ}$ C 连接过夜, 转化到 CaCl₂ 法制备的感受态细胞 DH5 α 中, 在铺有 100 μ l IPTG 及 20 μ l X-Gal 的氨苄青霉素琼脂平板上进行蓝、白斑筛选, 挑取白色菌落于含抗生素的 LB 液体中摇菌 12 ~ 16 h, 碱裂解法提取质粒 DNA, *Eco* R I 和 *Bam* H I 双酶切克隆载体(T-XTP5)获得带双酶切位点的 XTP5 基因(酶切体系:2 μ l 10 \times K buffer, 1 μ l *Eco* R I (10 U/ μ l), 1 μ l *Bam* H I (10 U/ μ l) 酶, 10 μ l 质粒, 补水至 20 μ l);将酶切鉴定正确的克隆送 Invitrogen 测序, 测序正确的克隆酶切后的片段用 1.0% 琼脂糖凝胶电泳分离, 玻璃奶法纯化, 将带 *Eco* R I 和 *Bam* H I 酶切位点的目的片段分别克隆到有同样酶切位点的 pcDNA3.1 和 pGBKT7 线性化载体上, 从而构建重组载体 pcDNA3.1-XTP5 和 pGBKT7-XTP5。利用 *Sac* I 限制型内切酶鉴定 pcDNA3.1-XTP5 载体, 利用 *Nde* I 和 *Pst* I 鉴定 pGBKT7-XTP5 载体。

结 果

一、XTP5 基因的生物信息学分析

该基因分子式为 C₉₄₄H₁₄₆₃N₂₄₉O₂₅₆S₁₆, 共有 2928 个原子, 分子量为 57 508.0, 理论上, pI=8.96, 消光系数=44725 M⁻¹·cm⁻¹(280 nm)。在 6.0 M 盐酸胍、0.02 M 磷酸盐缓冲液、pH=6.5 条件下, 以半胱氨酸来计算, 该蛋白的半衰期为 30 h(体外哺乳类网状细胞), 大于 20 h(酵母体内), 大于 10 h(大肠埃希菌内), 不稳定系数是 42.35。对 XTP5 基因序列进行 Blastn 分析, 找到其前端 5000 bp 序列

(即 9p24.2 的 2 795 750-2 801 654), 发现 1 个 200 bp 左右 CpG 岛, 且在相同区域发现高启动子活性序列 (Score > 0.90), 同时发现该序列具有转录终止信号位点。进一步分析显示: 该基因定位于 9p24.2, 具有 Unigene 号: KIAA0020, 其 UGC、UCG 和 CGU 为低频率密码子, AAA 和 GAA 为高频率密码子。其二级结构预测以 Helix 为主, 存在信号肽的概率为 0.0000A。氨基酸序列比对结果显示: XTP5 蛋白与翻译抑制因子的 RNA 结合蛋白同源。此外, 该蛋白可能没有跨膜区域信号。

预测 XTP5 蛋白的功能基序如下: 1 个酰胺化位点, 1 个环单磷酸腺苷磷酸化位点, 6 个酪蛋白激酶 2 (CK2) 磷酸化位点, 4 个 N-端肉豆蔻基 (十四烷基) 位点, 8 个蛋白激酶 C (PKC) 磷酸化位点, 1 个酪氨酸磷酸化位点, 1 个核膜脂蛋白脂质附着位点, 1 个聚腺苷结合蛋白 (PABP) 单一结构域, 1 个 CPL 结构域, 1 个 DC-EC 重复基序, 1 个 IMP 脱氢酶 GMP 还原酶结构域。

二、XTP5 基因的 RT-PCR

收获 HepG2 细胞后分离、纯化 mRNA, 外分光光度计测定其浓度为 $2.0 \mu\text{g}/\mu\text{l}$, $A_{260}/A_{280} = 2.09$; 1.0% 琼脂糖凝胶电泳显示 mRNA 有大于 0.5 kb 清晰慧尾片状条带。电泳结果显示: RT-PCR 产物大小为 1527 bp, 与理论大小完全相符 (图 1)。

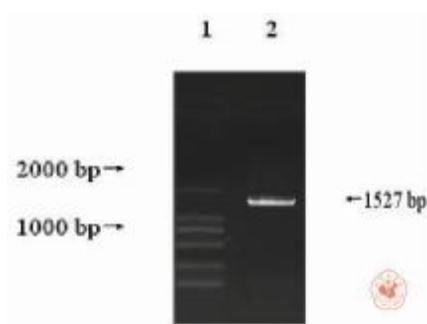


图 1 XTP5 基因的 RT-PCR 电泳
1. 标准分子量; 2. XTP5 基因的 RT-PCR

三、pcDNA3.1-XTP5 重组载体的构建与酶切鉴定

利用 Vector NTI VIII 绘制重组质粒 pcDNA3.1-XTP5 酶切图谱 (图 6)。Eco R I、Bam H I 和 Sac I 酶切该载体后实际酶切图谱与理论完全相符 (图 2~4)。

四、重组载体 pGBKT7-XTP5 的构建与酶切鉴定

利用 Vector NTI VIII 绘制重组质粒 pGBKT7-XTP5 酶切图谱 (图 5)。Eco R I、BamH I、Nde I 和 Pst I 酶切载体后实际酶切图谱与理论完全相符 (图 6、7), 提示重组载体构建成功。

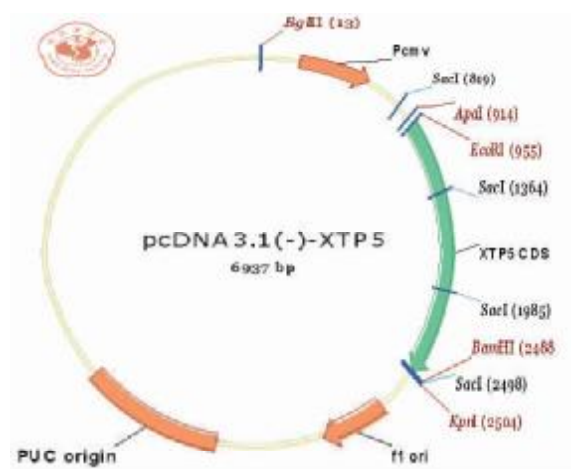


图 2 pcDNA3.1-XTP5 质粒图谱

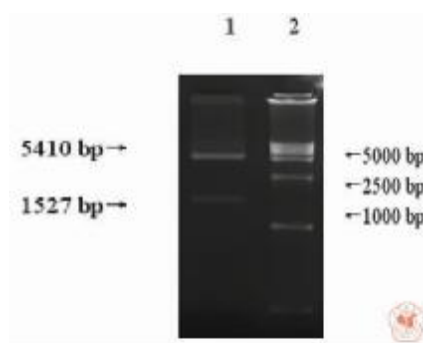


图 3 pcDNA3.1-XTP5 *EcoRI* 和 *BamHI* 双酶切

1. pcDNA3.1-XTP5 *EcoRI* 和 *BamHI* 双酶切结果; 2. 标准分子量

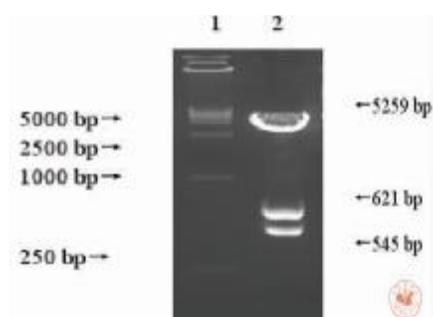


图 4 pcDNA3.1-XTP5 *SacI* 酶切

1. 标准分子量; 2. pcDNA3.1-XTP5 *SacI* 酶切结果

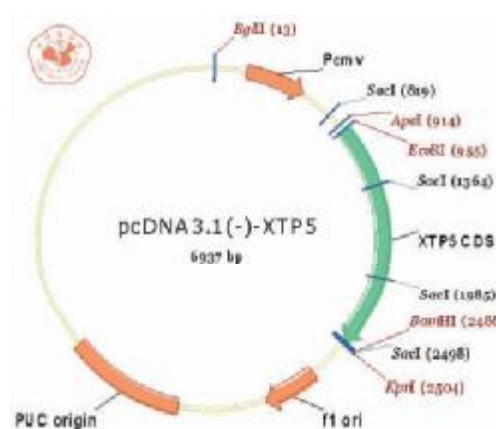


图 5 pGBKT7-XTP5 的质粒图谱

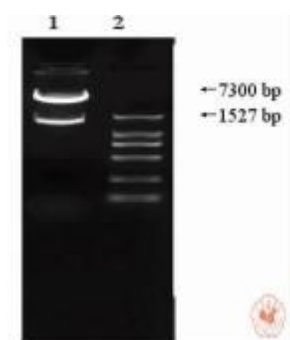


图 6 pGBKT7-XTP5 *EcoRI* 和 *BamHI* 双酶切
1. pGBKT7-XTP5 *EcoRI* 和 *BamHI* 酶切结果; 2. 标准分子量

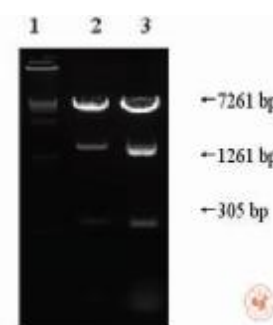


图 7 pGBKT7-XTP5 *NdeI* 和 *PstI* 双酶切
1. 标准分子量; 2、3. pGBKT7-XTP5 *NdeI* 和 *PstI* 双酶切结果

讨 论

HBV 属嗜肝 DNA 病毒科,其 X 蛋白是 HBV 致病的元凶。为阐明 HBV X 蛋白的致病机理,本课题组利用 SSH 技术构建并筛选了 HBxAg 蛋白反式激活基因差异表达的 cDNA 消减文库,发现新基因 XTP5,其 ORF 长度为 1527 bp,编码 508 个氨基酸残基。生物信息学(bioinformatics)是一门数学、统计、计算机与生物医学交叉结合的新兴学科,它已广泛地渗透到医学的各个研究领域中,成为生物医学发展不可缺少的重要工具。随着人类基因组计划的快速发展,生物信息学技术在人类疾病与功能基因的发现与识别、基因与蛋白质的表达与功能研究方面都发挥着关键的作用。

由于 CpG 岛、启动子活性和转录终止信号等是脊椎动物基因重要的基本特征,因此本研究利用生物信息学对 XTP5 进行这些预测。200 bp 的 CpG 岛、高启动子活性(Score > 0.90)和转录终止位点从理论上证实新基因 XTP5 的存在,RT-PCR 结果从实验上也证实该基因的存在。随着蛋白质组研究的推进,2006 年 XTP5 被诠释为“次要组织相容性抗原基因 HA-8(mHA-8)”^[4-6]。尽管该基因理论 pI 值、消光系数(280 nm)、不稳定系数等参数仍然需要实验证实,但在没有任何其他信息的情况下,该基因的生物信息学分析结果对今后进一步研究其结构和功能具有重要的参考价值。

XTP5 基因编码的氨基酸序列与 RNA 结合蛋白具有较高的同源性,且其功能基序包括:酪蛋白激酶 2(CK2)磷酸化位点、蛋白激酶 C(PKC)磷酸化位点、酪氨酸磷酸化位点等。这提示该蛋白与蛋白质磷酸化有关系,而蛋白质磷酸化是蛋白质翻译后修饰的最常见方式,可发生于一系列的代谢过程或信号转导途径,对体内代谢酶活性的调节、信号传导的级联、细胞运动、膜通道活性、细胞核的装配、细胞周期等起着非常重要的作用^[7]。

为了进一步研究 XTP5 基因编码蛋白的结果和功能,本研究构建了 pcDNA 3.1-XTP5 重组载体。另外,酵母双杂交系统是 1 种直接于酵母细胞内检测蛋白质之间相互作用的遗传学新方法,在细胞周期与分化、细胞凋亡、信号转导、癌基因产物功能、基因表达调控及蛋白自身特性等研究领域,越来越受青睐。由于融合蛋白必须转至核内才能活化转录,所以,膜蛋白也不适于此系统。对 XTP5 蛋白跨膜信息的生物信息学预测没有发现跨膜区段,提示该蛋白不是膜蛋白,因此,可利用酵母双杂交系统对 XTP5 基因进行深入探讨。因此,本研究也构建了酵母表达载体 pGBKT7-XTP5。

参 考 文 献

- 1 Chan DW, Ng IO. Knock-down of hepatitis B virus X protein reduces the tumorigenicity of hepatocellular carcinoma cells. *J Pathol*, 2006, 208:372-380.
- 2 刘妍, 成军, 陆荫英, 等. 乙型肝炎病毒 X 蛋白反式激活基因克隆化的研究. *中华肝脏病杂志*, 2003, 11:5-7.
- 3 Ji W, Wright MB, Cai L, et al. Efficacy of SSH PCR in isolating differentially expressed genes. *BMC Genomics*, 2002, 3:12.

- 4 Holloway PA, Kaldenhoven N, Kok-Schoemaker HM, et al. A class II-restricted cytotoxic T-cell clone recognizes a human minor histocompatibility antigen with a restricted tissue distribution. *Br J Haematol*, 2005, 128:73-81.
- 5 Araki J, Ohashi J, Muramatsu M. Application of discordant sib-pair linkage analysis for mapping minor histocompatibility antigen loci in a novel graft-vs-host-disease model. *Tissue Antigens*, 2004, 64:243-250.
- 6 Warren EH, Otterud BE, Linterman RW, et al. Feasibility of using genetic linkage analysis to identify the genes encoding T cell-defined minor histocompatibility antigens. *Tissue Antigens*, 2002, 59:293-303.
- 7 Fields S, Song O. A novel genetic system to detect protein-protein interactions. *Nature*, 1989, 340:245-246.
- 8 McCraith S, Holtzman T, Moss B, et al. Genome-wide analysis of vaccinia virus protein-protein interactions. *Proc Natl Acad Sci USA*, 2000, 97:4879-4884.

(收稿日期:2007-06-21)

(本文编辑:刘志英)