

乙型肝炎病毒前-前-S基因研究进展

张延峰, 成军 (首都医科大学传染病学研究所 北京地坛医院 传染病研究所, 北京 100015)

乙型肝炎病毒(HBV)感染呈世界性流行, 据世界卫生组织报道^[1], 全球约20亿人曾感染过HBV, 其中3.5亿人为慢性感染者, 每年约有100万人死于HBV感染所致的肝衰竭、肝硬化和原发性肝细胞癌(HCC)。中国属HBV感染高流行区, 据估计^[2]目前全国现患慢性乙型肝炎患者2000万人。尽管HBV疫苗研究取得了显著进展, 但病毒致病及感染慢性化的分子和细胞机制仍未阐明, 慢性乙型肝炎的治疗仍然未有突破。

对HBV基因组的研究是阐明病毒致病机制的基础。1979年Gelibert等^[3]测定了一株克隆在大肠埃希菌载体中的HBV(ayw血清型)全长核苷酸序列, 从而发现了第一个HBV全基因组序列, 长度为3182 nt。此后研究表明, HBV基因组含4个部分重叠的开放读码框(ORF), 即前-S/S区、前-C/C区、P区和X区。前-S/S区编码大(前-S1、前-S2及S)、中(前-S2及S)、小(S)3种包膜蛋白; 前-C/C区编码HBeAg及HBcAg; P区编码聚合酶; X区编码X蛋白^[4]。除以上4个公认的ORF外, 还发现在某些病毒株的前-S1区前存在一个前-前-S区, 这一发现引起了许多研究者的兴趣, 本文将对前-前-S区的研究进展进行简要综述。

1 前-前-S区的发现及确认

前-前-S区最早在1983年Fujiyama等^[5]克隆的adr亚型的S基因产物中被提及, 作者认为存在一身份不明的ORF。2003年, 董菁等^[6]对HBV前-前-S区进行了较为深入的研究。以2例乙型肝炎患者血清中抽提的HBV基因组为模板, 应用长距离且精确PCR方法(LA-PCR), 将HBV基因组克隆到测序载体中, 选择5个克隆进行测序。应用DNASIS

软件对上述5个克隆的ORF进行分析, 发现在前-S1区之前还存在一个ORF, 长度135 bp, 编码45 aa, 并将其命名为前-前-S。为了排除样本的选择误差, 应用了2种方法来证实前-前-S区的真实存在。首先是在GenBank中选择6株不同血清型的HBV基因组全序列, 应用DNASIS软件重新确定其ORF, 结果其中3株具有前-前-S ORF。其次是利用美国国立卫生研究院(NIH)网站的BLAST同源基因序列搜索软件, 将前-前-S区编码氨基酸序列输入后进行同源性搜索, 结果发现已经存入GenBank的序列中, 有5个克隆的氨基酸序列与该研究获得的前-前-S序列有较高的同源性, GenBank注册号分别为BAA32849、S43492、CAA25747、BAA89327和CAB38230, 氨基酸残基序列同源性分别为95%、93%、93%、77%和66%。2004年, 杨倩等^[7]对前-前-S区做了进一步研究, 再次证明了HBV基因组存在前-前-S区。从17例慢性乙型肝炎患者血清中提取HBV基因组, 应用多引物-多聚酶链反应进行HBV基因型分型, 然后将前-前-S区扩增, 克隆到测序载体后进行单克隆测序。结果测序的31个克隆都编码前-前-S多肽。2005年, 段学章等^[8]应用DNASTAR软件中的MegAlign程序分析了GenBank中的35株及3株该研究室测定的不同基因型HBV全序列中前-前-S基因分布情况, 结果发现共有24株存在前-前-S ORF; 在用Blast软件对GenBank中598株HBV全基因组序列进行前-前-S和前-X区的核苷酸相似性分析时发现90株(15.1%)HBV存在与前-前-S区核苷酸相似性高的序列。2006年, Faure^[9]通过文献检索, 共发现了241个(190个全长序列, 51个部分序列)HBV基因组序列中含有前-前-S ORF。早期研究认为前-前-S区的存在可能是由于一些位点发生变异, 是个别现象, 因而未

加以重视。近年来的研究资料表明前-前-S ORF在HBV基因组中确实是普遍存在的。

2 前-前-S区的基因型和血清型特异性以及地域分布

根据HBV全基因序列差异 $\geq 8\%$ 或S区基因序列差异 $\geq 4\%$,目前HBV分为A~H 8个基因型^[1,10]。根据HBsAg的氨基酸组成和抗原性的不同,HBV又分为4种不同亚型,即adr、adw、ayr和ayw^[11]。

我国流行的HBV基因型主要为C型和B型,血清型主要是adrq+和adw2,少数为ayw3。现有研究资料表明,前-前-X区具有基因型和血清型特异性以及明显的地域性。

杨倩等^[7]研究的15例患者中,基因型为B型者1例,B/C混合型者5例,C型9例。段学章等^[8]应用EditSeq软件对24株存在前-前-S区的HBV基因组进行了进一步分析,发现有7株C型、1株F型和3株H型都具有编码功能。Faure^[9]研究的241例存在前-前-S ORF的HBV基因组中,各基因型分布为C型(122例)、C/D混合型(5例)、B/C混合型(7例)、B型(4例)、A型(1例)。此外,在文献中102个基因组未说明基因型,根据其S基因的氨基酸序列推断,49个为C型,因此C型总计占71%。血清型分布为adr(47例)、adw(7例)、ayw(2例)、ayr(1例),在文献中184个基因组未说明血清型,根据其S基因的氨基酸序列推断,92个为adr型,adr型总计占约58%。以上资料表明除了某些例外,前-前-S区是C/adr特异性的。

前-前-S区分布具有明显的地域性。Faure^[9]的研究表明,241例含有前-前-S ORF的HBV基因组几乎都来自于亚洲或亚裔患者,其中中国内地(81例)和日本(96例)共占约73%。其余患者来自韩国(17例)、台湾(17例)、越南(5例)、印度尼西亚(4例)、泰国(3例)、欧洲(12例,均来自亚裔患者)、南非(1例,来自亚裔患者)、美国(3例)和澳大利亚(2例)。

前-前-S ORF的基因型和血清型特异性,尤其是明显的地域性特征,可能使欧美学者在最初的研究中获得的样本存在一定偏差,从而忽视了该

区域。由于存在前-前-S ORF的患者中来自我国内地的占相当高的比例,因此在我国对其进行深入而广泛的研究就具有格外重要的意义。

3 前-前-S区的序列变异

前-前-S ORF始于ATG,全长145 nt,后面紧接前-S1区的ATG,编码45个aa。杨倩等^[7]的研究根据替换突变的频度,将前-前-S多肽分为2部分,前-前-S多肽总一致率为75.6%,其中1~20 aa为高变区,8个位点发生变异,总一致率为60.0%;21~45为高保守区,仅3个位点发生变异,总一致率为88.0%。针对31个克隆的分析发现病毒氨基酸序列变异具有2个特点,其一是氨基酸序列的变异具有患者个体化突变的特点,在特定位点会出现相对特异性的变异,这种变异以替换突变为,具有较为强烈的个体化特征,即突变表现为来自某些患者的所有克隆的编码氨基酸与来自其他患者的HBV编码氨基酸不同,而来自其他患者的HBV在该位点具有一致性。这在HBV进化中的意义尚不清楚,但这种变异的指纹样特点在追踪感染链源头方面可能有意义。变异的另一个特点是在氨基酸序列中某些位点的变异表现出相对的一致性,比较突出的是第19位氨基酸位点,来自一些患者的克隆编码为R,而来自另外一些患者克隆编码L。L为非极性氨基酸,R为带正电荷的极性氨基酸,属碱性氨基酸。HBsAg血清型的分型是基于第122位氨基酸的K/R和第160位点氨基酸的K/R变异,那么前-前-S基因第19位氨基酸残基的L/R突变是否与全S基因(包括前-前-S、前-S1、前-S2和主蛋白)的血清型分型有关,还需要进一步研究。

4 前-前-S区与肝病的关系

Faure^[9]研究的241个前-前-S ORF中,69个来自肝细胞癌患者,38个来自慢性肝炎患者,28个来自急性重型肝炎患者,10个来自暴发性肝炎患者,另有96个文献中未予以说明,在已知来源的患者中,肝细胞癌患者占48%(69/145)。

以上资料表明,前-前-S区均存在于肝细胞癌、慢性肝炎、急性重型肝炎和暴发性肝炎患者,其中约半数来自肝细胞癌患者。尽管在肝细

胞癌的发病机制中HBV所起的确切作用还不清楚,但资料表明某些HBV表达的蛋白对病毒和细胞内的基因启动子都具有明显的反式激活作用,这可能与肝细胞癌的发生密切相关。尽管前-前-S区对肝细胞癌发生和发展的影响还需要进一步研究,但它可以作为评估亚洲C/adr型HBV携带者日后发生肝细胞癌风险的标志^[9]。

5 前-前-S区启动子的研究

基因组中并非所有的ORF都能表达蛋白质,作为基因序列,其上游应有启动子。杨倩等^[12]为了解前-X区上游DNA序列是否具有启动子活性,首先应用PCR技术扩增前-前-S ORF中ATG上游277 nt核苷酸序列,分析发现该序列富含TA(61%),虽无真核生物的启动子元件TATA盒,但发现含有TATA样盒。将该序列进行克隆,构建含报告基因氯霉素乙酰化酶(CAT)的重组质粒,瞬时转染HepG2细胞,酶联免疫吸附法(ELISA)检测CAT表达活性,结果发现界定的前-前-S-基因启动子序列能够激活下游CAT表达,具有启动子活性,重复实验也得到了同样的结果。Bogomolski等亦发现前-S2/S启动子不含典型的TATA盒,但有启动子活性,可与胞核提取物特异结合。在与SP1、SP2比较过程中发现,前-前-S基因与SP1部分重叠,这与SP2和前-S1的关系相似。前-前-S基因启动子完全包含在SP1中,前-前-S基因启动子与SP1的关系如何,两者之间是相互重叠,还是共用启动子,有待于进一步研究,但可以肯定的是前-前-S基因ORF上游的序列具有启动子活性。

6 前-前-S区编码多肽的研究

尽管有充分的研究资料证明前-前-S在DNA水平上是存在的,但必须从蛋白质水平证实前-X在感染HBV患者中表达。由于前-X编码蛋白是一种新型的功能未知的蛋白,不存在商品化的检测试剂盒,因而表达纯化前-X编码蛋白制备多克隆抗体应用于临床检测是必经之路。

王春花等^[13]通过PCR技术扩增HBV的前-前-S编码基因片段,将其克隆至表达载体pET32a(+),构建原核表达重组质粒,转化入大肠埃

希菌中诱导表达,经SDS-聚丙烯酰胺凝胶电泳(SDS-PAGE)和Western blot验证,成功表达了前-X编码蛋白,为进一步制备多克隆抗体奠定了基础。

7 前-前-S区作用机制的研究

对前-前-S、前-S1、前-S2和S基因的完全表达产物进行软件分析表明,前-前-S区较以往认为的大蛋白多出一个小的疏水区,其中的19(21)个疏水氨基酸在前-前-S区形成了一个小的疏水功能域,可能与蛋白的空间折叠或表面蛋白合成后分泌有关。推断的前-前-S多肽氨基酸序列中含有6个亮氨酸(L),分别位于第3、9、15、19、27、39位,前3个L是否可形成亮氨酸拉链结构,尚需要进一步证实。

7.1 前-前-S基因启动子结合蛋白基因的筛选 黄燕萍等^[14]采用HBV前-前-S基因启动子作为固相靶分子,用噬菌体表面展示的人肝cDNA文库来筛选特异的HBV前-前-S基因启动子结合蛋白,以探讨HBV前-前-S在HBV感染发病机制中的作用。根据前-前-S基因的转录起始点,选取其上游277 nt,设计并合成引物,在上下游引物的5'-端分别加上MluI和XhoI位点序列,并且以生物素标记。以固相化的HBV前-前-S启动子DNA片段作为支持分子,对肝细胞cDNA文库进行4轮“吸附—洗脱—扩增”筛选。挑选43个克隆测序,测序结果与GenBank数据库进行初步比较,筛选出20个与HBV前-前-S基因启动子特异结合蛋白,包括人类SMG-1样的磷脂酰肌醇3相关激酶激酶、28S核糖体、单倍型As2A线粒体、组氨酰-tRNA合成酶、脂肪醛脱氢酶、桥粒相关蛋白、MAX相互作用蛋白1等17个已知功能基因及3个未知功能基因,同源性为96%~100%。

核孔复合体(NPC)是一种巨大的蛋白复合物,插在核外膜的双层膜中,作为一种装配的可溶性通道发挥作用,是允许与NPC成分作用的转运受体进出的选择性通道。不同可溶性转运受体可以识别许多转入或转出细胞核的转运物质,并携带载物到达各自的目标物。所有的核转运因子可以双向通过NPC,这是通过与富含苯丙氨酸和

甘氨酸的NPC成分特定的相互作用完成的。核质转运通过NPC转移定位,具有高度选择性并非常迅速。输入蛋白或输出蛋白介导的转运所需要的能量来源只能由RanGTP周期提供,它可通过调节底物接合和释放反应来进行定向转移。由于核糖体亚基的复杂生物合成性,一直很难对核糖体从核中转出进行分析,通过系列研究表明40S转出要求RanGTP酶周期及核孔蛋白;几种核孔蛋白的突变及Ran周期的成员突变禁止60S转运;60S成熟化与核释放以及转运到细胞质有紧密联系。核孔复合体相互作用蛋白(NPIP),由免疫荧光显微术证实定位于核膜,与核孔P62共区域化。通过噬菌体展示技术筛选到的HBV前-前S启动子的结合蛋白NPIP,是否能提示HBV感染后,影响肝细胞核孔复合体功能和核糖体的成熟,从而影响肝细胞蛋白的合成及功能。

值得注意的是筛选到的HBV前-前-S启动子的另一个结合蛋白是磷脂酰肌醇3激酶相关激酶(PIKK)与信号转导密切相关。信号转导是指外部的信号通过细胞膜上的受体蛋白传到细胞内部,并激发出诸如离子通透性、细胞形状或其他细胞功能改变的应答过程。PIKK家族是细胞周期的主要调节剂。磷脂酰肌醇3激酶(PI3K)是细胞内信号转导系统的一条通道,是肌醇磷脂信使系统的重要组成部分,介于细胞受体与第二信使之间的信号转导途径,在细胞生物学功能中起到重要作用。PI3K通过将第二信使前体磷酸化而转导信息,所产生的第二信使可能参与细胞代谢、生长分化、基因表达、细胞凋亡等过程。外源性蛋白如病毒蛋白可通过PI3K途径导致细胞凋亡或癌变,有研究发现PI3K激活丝/苏氨酸激酶-Akt,并可能级联放大下游抑癌基因表达达到抑制细胞凋亡目的。磷脂酰肌醇3激酶作为胰岛素信号转导中的关键酶,在调节糖代谢中起重要作用。2型糖尿病和胰岛素抵抗(IR)患者中,PI3K的含量和活性均降低,并存在对胰岛素刺激的反应缺陷。T、B细胞的活化过程中涉及到多种磷脂酰肌醇的产生及其调节,不同种的磷脂酰肌醇可分别激活下游信号蛋白,使信号逐级传递,通过信号蛋白激

活T、B细胞内的多种酶活化途径,最终导致T、B细胞的增殖、活化,合成并发挥其生物学功能。慢性肝损伤及炎症反应时,常伴有多种可影响不同靶细胞的可溶性物质被激活,如血小板源性生长因子(platelet-derived growth factor, PDGF)及其受体(PDGF receptor, PDGF-R)的表达常可增强。PI3K参与细胞内PDGF信号转导,活化的PDGF-R可使PI3K内的酪氨酸残基磷酸化,从而促进HSC增殖及趋化,PI3K下游效应分子有蛋白激酶C、40S核糖体蛋白S6激酶等。PI3K可催化磷脂酰肌醇-4, 5-二磷酸之D3位磷酸化,其产物磷脂酰肌醇-3, 4, 5-三磷酸可激活PKC基因,而后者可激活Ras,而Ras/MAPK通路与PDGF诱导HSC增殖及移动有关。该研究通过噬菌体展示技术筛选到的部分HBV前-前-S启动子基因结合蛋白为阐明HBV感染的慢性化和肝细胞癌形成机制、并发症的发生机制以及机体免疫功能改变提供了新的研究方向和思路。

7.2 前-前-S蛋白结合蛋白新基因PPSBP9的克隆化研究 基因是通过蛋白质之间的相互作用而实现其功能的,病毒蛋白和肝细胞蛋白之间的作用是病毒致病的关键。为进一步研究前-前-S-基因在HBV致病中的作用,蔺淑梅等^[15]利用酵母双杂交技术对肝细胞cDNA文库中与乙型肝炎病毒前-前-S相互作用的蛋白进行研究。用PCR技术扩增HBV前-前-S蛋白编码基因,连接入酵母表达载体pGBKT-7中构建诱饵质粒,酶切鉴定后,用醋酸锂法转入酵母细胞AH109,并在四缺培养基上培养以排除其自身激活作用。与肝细胞文库酵母细胞配合,对真阳性菌落克隆和分析,共挑选54个阳性克隆测序,并与GenBank数据库进行初步比较,获得了一种能与HBV前-前-S蛋白结合的未知功能蛋白,将编码该蛋白的基因命名为PPSBP9,GenBank注册号为AY553877,其编码序列全长为840个核苷酸,编码产物由279个氨基酸残基组成。

值得关注的是,在对这一未知功能新基因进行进一步生物信息学分析时发现,在用酵母双杂交技术筛选HCV F蛋白与肝细胞文库相互作用的

蛋白时也筛选到这一相同的未知功能的新基因,故将其命名为PPSBP9/FBP2。表明这一未知功能基因能与多种病毒蛋白结合,提示该基因可能在HBV和HCV感染的致病过程中发挥重要作用。新基因的发现及成功克隆化为今后进一步深入地研究新基因的生物功能奠定了基础。

7.3 全S蛋白的反式激活作用 全S基因包括前-前-S基因(135 nt)、前-S1基因(357 nt)、前-S2基因(165 nt)和S基因(678 nt),全长1335 nt,编码445 aa,每一部分都有单独的ATG。白桂芹等^[16]应用抑制消减杂交技术研究了全S蛋白的反式激活作用。构建含全S基因的pcDNA3.1

(一)-全S真核表达载体,然后与Psv-lacZ共转染HepG2细胞,48小时后收获细胞检测 β -半乳糖苷酶的活性进行抑制消减杂交并进行生物信息学分析。结果表明转染pcDNA3.1(一)-全S质粒的HepG2细胞中 β -半乳糖苷酶的活性比对照质粒高6.9倍,扩增的文库含86个阳性克隆。随机挑取35个克隆进行测序,并与GenBank中的已知序列进行比对,共获得了33个编码序列。此外,发现了2个未知功能基因,通过生物信息学分析获得了其中一个基因的全长序列,命名为CSTP1,GenBank注册号为AY553877。获得的基因可能是全-X蛋白反式激活的靶基因,值得注意的是其中有些基因参与细胞周期调控、代谢、免疫、信号转导、细胞凋亡等过程,以及与肝细胞癌的形成有关。

本课题组近年来对前-前-S进行了系统研究,包括编码区界定、启动子序列的确定及转录活性的鉴定、蛋白体外表达、反式激活基因及蛋白-蛋白结合作用等。这些研究证明了前-前-S在基因水平上的真实存在,并初步探讨了前-前-S在HBV致病性方面的可能作用,为未来研究奠定了坚实的基础。进一步研究应首先以表达的前-前-S蛋白为抗原制备多克隆抗体并检测患者体内是否存在前-前-S蛋白表达。如果能在蛋白质水平上证实前-前-S的存在,将对HBV感染的检测、疫苗设计、HBV进入肝细胞机制、表面抗原的表达过程及功能、宿主抗感染机制、肝细胞癌产生机制研究等

均产生重大影响。

参考文献

- [1] World Health Organization. Hepatitis B. World Health Organization Fact Sheet 204 dex. (Revised October 2000). WHO Web site. <http://www.who.int/mediacentre/factsheets/fs204/en/in.html>
- [2] 卫生部. 2006-2010年全国乙型肝炎防治规划. 2006年1月28日颁布.
- [3] Gelibert F, Mandart E, Fitoussi F, et al. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli[J]. *Nature*, 1979, 281: 646-650.
- [4] 中华医学会肝病学会、感染病学会. 慢性乙型肝炎防治指南[J]. *中华肝脏病杂志*, 2005, 13: 881-891.
- [5] Fujiyama A, Miyanohara A, Nozaki C, et al. Cloning and structural analyses of hepatitis B virus DNAs, subtype adr[J]. *Nucleic Acids Res*, 1983, 11: 4601-4610.
- [6] 董菁, 成军. 乙型肝炎病毒基因组中前-前-S区编码基因的界定[J]. *世界华人消化杂志*, 2003, 11: 1091-1096.
- [7] 杨倩, 董菁, 成军. 乙型肝炎病毒基因组前-前-S基因区的分子流行病学研究[J]. *世界华人消化杂志*, 2004, 4: 785-789.
- [8] 段学章, 庄辉, 杜珩, 等. 乙型肝炎病毒基因组前-前-S和前-X区基因分布的研究[J]. *中华微生物学和免疫学杂志*, 2005, 25: 348-350.
- [9] Faure E. Alternative peptide-fusion proteins generated by out-of-frame mutations, just upstream ORFs or elongations in mutants of human hepatitis B virus[J]. *Virus Res*, 2006, 117: 185-201.
- [10] Akuta N, Kumada H. Influence of hepatitis B virus genotypes on the response to antiviral therapies[J]. *J Antimicrob Chemother*, 2005, 55: 139-142.
- [11] 黄耀煌. 肝病分子生物学[M]. 福州: 福建科学技术出版社, 2003.
- [12] 杨倩, 董菁, 成军, 等. 乙型肝炎病毒基因组中前-前-S-编码基因启动子序列的确定及转录活性的鉴定[J]. *解放军医学杂志*, 2003, 28: 761-762, 765.
- [13] 王春花, 成军, 郎振为, 等. 乙型肝炎病毒前-前-S蛋白在大肠杆菌中的表达和纯化[J]. *世界华人消化杂志*, 2005, 13: 1894-1896.
- [14] 黄燕萍, 成军, 张树林, 等. 噬菌体展示技术筛选HBV前-前-S抗原基因启动子结合蛋白基因[J]. *世界华人消化杂志*, 2004, 12: 2801-2804.
- [15] 蔺淑梅, 成军, 张树林, 等. 乙型肝炎病毒前-前-S蛋白结合蛋白新基因PPSBP9的克隆化[J]. *世界华人消化杂志*, 2004, 12: 2733-2736.
- [16] Bai GQ, Liu Y, Cheng J, et al. Transactivating effect of complete S protein of hepatitis B virus and cloning of genes transactivated by complete S protein using suppression subtractive hybridization technique[J]. *World J Gastroenterol*, 2005, 11: 3893-3898.

收稿日期: 2009-05-26