

人载脂蛋白 B 基因 3' 端高可变区不同等位基因 在真核细胞中对基因表达调控的作用

何 平 吕新跃 薛 红^① 陈保生

(中国医学科学院基础医学研究所生物化学及分子生物学研究室, 北京 100005)
(中国协和医科大学)

The Role of Different 3' End of Variable Number of Tandemly Repeat Alleles from Apolipoprotein B Gene in the Expression and Regulation of Gene

HE Ping, LU Xin-Yue, XUE Hong and CHEN Bao-Sheng

(Institute of Basic Medical Sciences, CAMS & PUMC, Beijing 100005, China)

ABSTRACT

Aim The present paper is to study the effect of different 3' end of variable number of tandemly repeat (VNTR) alleles from human apolipoprotein B gene on eukaryotic expression and regulation.

Methods *HVE36* allele distributed widely in normal population and *HVE44* allele distributed mainly in patients with atherosclerosis were cloned into plasmids of luciferase reporter vectors pGL2-promoter and pGL2-control. Then the recombinants were transfected into HeLa and HepG2 cell lines.

Results The level of expression of pGL2-control plasmid inserted with *HVE44* allele was 4 times than those of the recombinant with *HVE36*. The level of expression of pGL2-promoter plasmid inserted by *HVE44* is 1.7 times than those in plasmid inserted by *HVE36* allele in HeLa cell line. In HepG2 cell line, the level of expression of pGL2-control recombinant with *HVE44* is 3.05 times of those in *HVE36* recombinant. But in pGL2-promoter system, the level

of expression of recombinant with *HVE44* alleles is 2.1 time than those in *HVE36* recombinant. In different cell lines, the expression level of recombinant with *HVE44* was much higher than those of *HVE36* recombinant.

Conclusions Different 3' VNTR alleles have regulatory role in gene expression and *HVE44* allele can positively regulate the eukaryotic expression, which may be one of causes of higher apolipoprotein B level in atherosclerosis patients with VNTR-big alleles like *HVE44*. More investigations are carried out for regulatory mechanism of the 3'-end nontranslation region.

KEY WORDS Apolipoprotein B gene; 3' end hypervariable region alleles; Expression regulation

摘要 为探讨人载脂蛋白 B 基因 3' 端高可变区 (hypervariable region, HVR) 不同等位基因对真核细胞基因表达调控的作用。选取在正常人群中分布最广的 *HVE36* 小等位基因和在冠心病病人中分布占优势的大等位基因 *HVE44*, 分别克隆进含荧光素酶基因的 pGL2-promoter 及 pGL2-control 质粒中, 转染 HeLa 细胞和 HepG2 细胞。结果发现, 在不同的细胞系中插入 *HVE44* 等位基因的重组质粒表达均高于插入 *HVE36* 的重组质粒的表达。表明载脂蛋白 B 基因 3' 端高可变区不同等位基因具有调控基因表达的作用, *HVE44* 大等位基因可上调细胞内的基因表达, 这或许是以 *HVE44* 等大等位基因为主体的动脉粥样硬化患者体内载脂蛋白 B 升高的原因之一。

关键词 载脂蛋白 B 基因; 3' 端高可变区等位基因; 表达调控

动脉粥样硬化 (atherosclerosis, As) 是由遗传和环境多种因素共同致病。血浆低密度脂蛋白 (low density lipoprotein, LDL) 及其载脂

蛋白 B (apolipoprotein B) 与 As 和冠状动脉心脏病 (coronary heart disease, CHD) 的形成呈正相关已被公认。作为导致 As 的重要候选基因, 载脂蛋白 B 基因具有高度多态性, 其中载脂蛋白 B 基因 3' 端高可变区 (hypervariable region, HVR) 或称可变数目串联重复序列 (variable number of tandemly repeated short DNA sequence, VNTR) 就是载脂蛋白 B 基因 3' 端下游一段富含 AT、长度约在 400~1 200 个碱基对的高可变重复序列。根据重复序列长度的不同^[1], 现已发现 22 种不同的等位基因。可将其分为 VNTR—大等位基因 (拷贝数大于或等于 38) 和 VNTR—小等位基因 (拷贝数小于或等于 36)。载脂蛋白 B 基因 3' VNTR 多态性对于冠心病易感性具有独立判别意义, 我组最近的研究发现 3' VNTR-大等位基因 *HVE42*、*HVE44*、*HVE46* 在 CHD 病人中占优势, 而正常人群中 *HVE36* 小等位基因分布最广^[2,3]。VNTR 的功能研究甚少, 尚未见到不同 VNTR 等位基因对基因表达、调控作用的研究报道。因此我们选择了 *HVE36* 和 *HVE44* 两个等位基因, 分别克隆到含荧光素酶的报告基因中, 并转染 Hela 细胞和 HepG2 细胞, 观察其对两种真核细胞基因表达调控的影响。

1 材料和方法

1.1 实验材料

1.1.1 载体 荧光素酶报告基因载体 (luciferase reporter vectors): pGL2-control 型质粒 pGL2-promoter 型质粒购于 promega 公司。

1.1.2 工具酶及 DNA 标准分子量 BamH I 内切酶、T4 DNA 连接酶、RNase A、溶菌酶均为 Promega 公司产品。Sal I 和 Bgl II 内切酶为协和友谊开发公司产品。klenow 大片段酶为 Pharmacia 公司产品。荧光素酶报告基因检测试剂盒 (Luciferase Assay System) 为 Promega 公司产品。

1.1.3 荧光发光测定仪 (Monolight 2010 Instrument Analytical Luminescence Laboratory) 和 PCR 扩增仪均为美国产品。

1.2 实验方法

1.2.1 扩增提取 3' VNTR 的 *HVE36* 和 *HVE44* 等

位基因 以已知携带这两种等位基因的纯合子个体的基因组 DNA 为模板, 用多聚酶链反应 (polymerase chain reaction, PCR) 方法扩增这些片段, 反应总体积为 100 μ l。PCR 引物为:

5' primer: 5'-ATGGAAACGGAGAAATTATG-3',

3' primer: 5'-CCTTCTCACTTGGCAAATAC-3'。

PCR 条件为 94℃ 预变性 5 min \rightarrow 94℃ 变性 30 s \rightarrow 58℃ 复性及延伸 60 s \rightarrow 58℃ 10 min, 共 30 个循环。每个样品共扩增 500 μ l。行 2% 低熔点琼脂糖电泳, 观察并回收、纯化扩增的两种等位基因, 其两端均为平端。扩增的 *HVE36* 和 *HVE44* 等位基因分别为 680 和 800 碱基对。

1.2.2 克隆载体的处理 分别取 5 μ g 环形质粒 pGL2-promoter (含 SV40 promoter) 和 pGL2-control 质粒 (含 SV40 promoter 及 enhancer)、1 \times Buffer E 和 10 单位的 BamH I 酶, 体积 50 μ l。37℃ 保温 2 h, 完全酶解后, 加入 1 μ l Klenow 大片段酶补齐平端, 使之成为平端。

1.2.3 连接、转化及重组克隆筛选 分别将 *HVE36* 和 *HVE44* 克隆至 pGL2-promoter 和 pGL2-control 质粒, 克隆位点选择质粒中报告基因下游 Poly A 附加信号之后的 BamH I 位点, 使得片段处于表达基因的 3' 端。这与它们在体内载脂蛋白 B 基因中所处的位置相当。取消补齐的载体和扩增回收的两种等位基因 (摩尔比为 1:3) 加 T4 连接酶后 14℃ 连接 12~14 h。然后将连接的质粒转化入 JM109 感受态菌中, 通过酶切筛选出正向插入的克隆。大量提取质粒并纯化, 检测所提取的质粒 DNA 的纯度和浓度要大于 1.8。

由此, 我们得到了 4 个重组正向克隆。它们是 *HVE36* 插入 pGL2-promoter 质粒 (简称为 P36), *HVE36* 插入 pGL2-control 质粒 (简称为 C36), *HVE44* 插入 pGL2-promoter 质粒 (简称为 P44), *HVE44* 插入 pGL2-control 质粒 (简称为 C44)。

1.2.4 真核细胞培养 Hela 细胞和 HepG2 细胞在含 10% 小牛血清的 DMEM 培养液中作常规培养。

1.2.5 转染细胞 分别取 2.6×10^5 个培养细胞接种至 35 mm 的培养皿培养 12 h, 在转染前 3 h 换液。各取 10 μ g 重组的质粒以磷酸钙共沉淀法转染至 Hela 细胞和 HepG2 细胞。

1.2.6 荧光素酶的检测 转染 48 h 后裂解细胞, 收集细胞碎片, 检测荧光素酶的活性 (按 Luciferase Assay System 说明进行)^[4]。

1.2.7 数据的统计学处理 在平衡蛋白浓度以后, 对所得到的数据作以下处理, 每次实验中同一个样品

作三个平行板,除去一个数值偏差大的,其余两个取平均值,也就是一次实验一个样品得到一个数值。共重复实验4次,每个样品有4个数值。为避免每组数值偏差较大,将pGL2-control质粒的测定数值作为1,将C36(HVE36插入pGL2-control质粒的简称)和C44(HVE44插入pGL2-control质粒的简称)的测定数值与之作倍数比。将倍数作统计学处理,这样就消除实验中每组数据不平均的误差。将所测数据进行单因素方差分析。

2 结果

2.1 重组的荧光素酶报告基因在Hela细胞系中的表达

2.1.1 pGL2-control体系的表达 两种重组3'VNTR等位基因的pGL2-control质粒在

Hela细胞系中的表达结果见表1(Table 1)。可见C44重组体的表达量是对照pGL2-control质粒表达量的2.025倍。C36重组体的表达量仅为对照的一半,C44的表达量则是C36的4倍($P<0.01$),具有高度显著性差异。

2.1.2 pGL2-promoter体系的表达 载脂蛋白B基因3'VNTR等位基因HVE36和HVE44分别克隆入pGL2-promoter质粒3'端BamH I位点。得到的两个重组体P36(将HVE36等位基因插入pGL2-promoter质粒组成重组体的简称)和P44(将HVE44等位基因插入pGL2-promoter质粒组成的重组体的简称),表达结果见表2(Table 2)。

Table 1. Comparison of expression times of different pGL2-control recombinants with HVE36 and HVE44 alleles in Hela cell line.

Sample	count				$\bar{x} \pm s$
pGL2-control	1	1	1	1	1
pGL2-control recombinant-HVE36	0.39	0.35	0.60	0.69	0.508 ± 0.142^a
pGL2-control recombinant-HVE44	1.52	1.92	1.96	2.70	2.025 ± 0.426^b

a: $P<0.01$, compared with pGL2-control; b: $P<0.01$, compared with pGL2-control recombinant-HVE36.

Table 2. Comparison of expression times of different pGL2-promoter recombinants with HVE36 and HVE44 alleles in HeLa cell line.

Sample	count				$\bar{x} \pm s$
pGL2-promoter	1	1	1	1	1
pGL2-promoter recombinant-HVE36	0.43	0.46	0.74	0.51	0.535 ± 0.122^a
pGL2-promoter recombinant-HVE44	0.91	0.96	0.86	0.89	0.905 ± 0.036^b

a: $P<0.01$, compared with pGL2-promoter; b: $P<0.05$, compared with pGL2-promoter recombinant-HVE36.

2.2 重组的荧光素酶报告基因在HepG2细胞系中的表达

2.2.1 pGL2-control体系的表达 用重组体C36和C44以及对照质粒pGL2-control转染HepG2细胞,发现C36表达量比对照质粒低,而C44表达量是对照质粒表达量的2.19倍。各重组体的表达量具有显著性差异(表3, Table 3)。

2.2.2 pGL2-promoter体系的表达

pGL2-promoter体系在HepG2细胞中的表达结果见表4(Table 4)。表明P44重组体的表达量是对照质粒pGL2-promoter的1.68倍。P36重组的表达量是对照质粒0.80倍。而P44重组体的表达量是P36的2.1倍。两者具有高度显著性差异($P<0.01$)。

Table 3. Comparison of expression times of different pGL2-control recombinants with *HVE36* and *HVE44* alleles in HepG2 cell line.

Sample	count					$\bar{x} \pm s$
pGL2-control	1	1	1	1	1	
pGL2-control recombinant- <i>HVE36</i>	0.72	0.40	0.83	0.91		0.715 ± 0.194^a
pGL2-control recombinant- <i>HVE44</i>	3.21	1.41	1.29	2.86		2.193 ± 0.853^{bc}

a: $P < 0.05$, compared with pGL2-control; b: $P < 0.01$, compared with pGL2-control; c: $P < 0.01$, compared with pGL2-control recombinant-*HVE36*.

Table 4. Comparison of expression times of different pGL2-promoter recombinants with *HVE36* and *HVE44* alleles in HepG2 cell line.

Sample	count					$\bar{x} \pm s$
pGL2-promoter	1	1	1	1	1	
pGL2-promoter recombinant- <i>HVE36</i>	0.78	1.07	0.61	0.72		0.795 ± 0.170^a
pGL2-promoter recombinant- <i>HVE44</i>	1.99	2.22	1.19	1.30		1.675 ± 0.439^b

a: $P < 0.05$, compared with pGL2-promoter; b: $P < 0.01$, compared with pGL2-promoter recombinant-*HVE36*.

3 讨论

3.1 研究表明不同长度的载脂蛋白 B 基因 3'端高可变区等位基因对真核细胞基因具有不同的调控作用。在所做的两个细胞系的四组实验中,*HVE36* 等位基因两个重组体(C36 和 P36)的表达量均少于对照质粒。而 *HVE44* 等位基因的两个重组质粒在不同细胞系中的表达与 *HVE36* 重组体有明显差别。在四组实验中,除 Hela 细胞系的 P44 以外,*HVE44* 的重组体(C44 和 P44)的表达量均高于相应的对照质粒,且更高于 *HVE36* 重组体。因而,在 Hela 和 HepG2 细胞系中 *HVE44* 等位基因能够提高其表达量,可能起正调控作用。

3.2 3'端高可变区不同等位基因对基因表达作用的机制探讨

不同等位基因内部结构有差异。*HVE36* 全长约 680 bp, *HVE44* 全长约 800 bp。*HVE44* 的(XY)重复序列比 *HVE36* 多 3 个拷贝数,而(Xii Yii)重复序列拷贝数比 *HVE36* 多 1 个。

就载脂蛋白 B 基因 3'VNTR 的二级结构来看,整个重复序列排列成串联的十字架结构。这种十字架的二级结构很有可能在结合某些蛋白因子以后发生构象变化。*HVE36* 与 *HVE44*

在(XY)重复区和(XiiYii)重复区的长度不同,有可能影响整个十字架结构结合蛋白的稳定性,进而对表达活性出现不同的影响。另外,载脂蛋白 B 基因 3'端的染色质结构上存在两个 DNase I 高敏位点(DNase I hypersensitive)^[5],第一个 DNase I 高敏位点于 3'端 AT 高可变重复序列内,同时具有丰富的拓扑异构酶 II (Topoisomerase II) 识别位点,并可使基因的 3'端结合核基质。第二个 DNase I 高敏位点在第一个高敏位点的上游。DNase I 高敏位点这些部位往往具有转录活性。

转录复合物通过基因时,复合物前方产生正向超螺旋(或在后方产生负向超螺旋),拓扑异构酶 II 将在此处起到解旋作用。DNase I 高敏位点使此过程变得容易进行。3'端的 DNase I 高敏位点的其它作用可能是保持载脂蛋白 B 基因区对 DNase I 高敏感性和载脂蛋白 B 基因在肝脏和小肠细胞中的转录状态。因而,允许负责转录调控的特异细胞内信号对载脂蛋白 B 基因转录作进一步的调节。就我们所研究的载脂蛋白 B 基因的 5'和 3'末端均可以结合核基质,这些核基质当处于 5'和 3'的 A/T 丰富区时,可激活拓扑异构酶 II,并完成与识别位点的结合过程;3'端结合核基质的位点则存在于

3'VNTR 区中^[6]。由此可以推测,处于载脂蛋白 B 基因 3'端的高重复序列,对载脂蛋白 B 基因表达调控有一定的影响。

两个等位基因重复序列内部的拓扑异构酶 II 作用位点和骨架蛋白结合区(scaffold attachment region, SAR)的分布也不相同^[7]。拓扑异构酶 II 同复制有关。而 SAR 是 DNA 样环连接染色体支架上的区域,这种支架可能就是核基质(核骨架)。HVE44 的(XY)重复序列的拷贝数比 HVE36 多 3 个,(Xii Yii)重复序列的拷贝数比 HVE36 多 1 个。一个(XY)重复序列有两个拓扑异构酶 II 位点和一个 SAR 位点。一个(Xii Yii)重复序列有一个拓扑异构酶 II 位点。所以,HVE44 等位基因的拓扑异构酶 II 位点和 SAR 位点比 HVE36 多。由于拓扑异构酶 II 位点和 SAR 位点数目的不同,导致结合的核基质数目也不同。虽然,HVE36 和 HVE44 同样具有拓扑异构酶 II 位点和 SAR 位点以及十字架二级结构,但在表达上却相差很大。其原因也许就是因为位点数目的不同,结合的核基质不同。因而,对于基因的复制和转录也有相应的差异。

3.3 载脂蛋白 B 基因 3'端高可变区与动脉粥样硬化的关系

载脂蛋白 B100 在人肝脏中合成,体外培养的人肝肿瘤衍生细胞系 HeGp2 细胞株亦可表达分泌。我们的实验结果表明:HVE44 等位基因对 HeGp2 等细胞的基因表达有上调作用。可以推测带有 HVE44 的载脂蛋白 B 基因的转录活性高于带有 HVE36 的载脂蛋白 B 基因。在相同代谢水平下,带有 HVE44 的个体其血清中载脂蛋白 B 浓度可能会高于带有 HVE36 的个体。载脂蛋白 B 水平的升高将会导致

LDL 增高。使不同个体对 As 的敏感性呈现不同的反应。这可以部分解释以 VNTR—大等位基因为主体的 As 患者血浆载脂蛋白 B 水平升高的原因。但就人们所知,动脉粥样硬化是一种多基因病,受环境和遗传双重因素的作用。动脉粥样硬化病的形成是一长期缓慢的过程,受多种微效基因和环境因素“相互作用”的影响。载脂蛋白 B 基因 3'VNTR—大等位基因也可能就是这些微效基因的一种。

参考文献

- 1 Friedl W, Ludwig EH, Paulweder F, et al. Hypervariability in a minisatellite 3' end of the apolipoprotein B gene in patient with coronary heart disease compared normal controls. *J Lipid Res*, 1990, **31**: 659~665.
- 2 Ye Ping, Chen Baisheng, Wang Shiwen. Association of polymorphisms of the apolipoprotein B gene with coronary heart disease in Han Chinese. *Atherosclerosis*, 1995, **117**: 43~50.
- 3 陈保生,郭中民,何平,等. 中国汉族和蒙古族人群载脂蛋白 B 基因 3'端高可变区等位基因的分布频率及序列分析. *中国动脉硬化杂志*, 1996, **4** (2): 93~99.
- 4 Suit E. Luciferase Assay Guide Book: Protocols and information for measuring firefly luciferase expressed in cells. *San Diego USA*, 1992, 8~10.
- 5 Beatriz Levy-Wilson. DNase I hypersensitive sites at the 3' end of the human apolipoprotein B gene. *Biochem Biophys Res Comm*, 1990, **171** (1): 162~168.
- 6 Behringer RR, Hammer RE, Brinster RL, et al. Two 3' sequences direct adult erythroid-sogenic expression of human-biolin gene sineans genicmice. *Proc Natl Acad Sci USA*, 1987, **84**: 7 056~63.
- 7 Buresi C, Desmarais E. Structural analysis of the minisatellite present at the 3' end of the human apolipoprotein B gene. New definition of the alleles and evolutionary implications. *Hum Mol Genet*, 1995, **12**: 568~577.

(1996-12-10 收到, 1997-02-28 修回)