

一种新的兔内皮素 B 受体基因克隆和全序列分析

杨 林^{1*} 吴秋豫¹ 杨 楠² 龙紫新¹ 王统章¹

(中山大学生物防治国家重点实验室生物医药中心 广州 510275)

(澳大利亚 New Children 's 医院 ,NSW 2124)

关键词 兔内皮素 B 受体 , 原位杂交 , cDNA 克隆和序列分析

中图分类号 TQ92 文献标识码 A 文章编号 1000-3061(2000)03-0403-05

1988 年 ,Yanagisawa 和其同事在培养的猪主动脉内皮细胞上清中分离到能对离体的猪冠状动脉产生收缩作用的多肽即内皮素(Endothelin), 简称为 ET^[1]。ET 不仅具有强大的缩血管作用 , 而且是体内多种细胞的促分裂剂 , 对心血管、肾脏、呼吸、消化、神经、生殖、内分泌各系统均有作用 ; 更为重要的是 , 在多种疾病中 , 它又作为一种致病因子参与疾病的发生 , 具有重要的病理生理意义。ET 只有与靶细胞上特异的 ET 受体结合 , 启动一系列细胞内信息传递系统 , 其中主要是磷酸肌醇系统 , 调节胞内钙浓度 , 才得以发挥其效应。在人类基因组中已发现 3 种内皮素异构肽 , 即 ET-1、ET-2 和 ET-3。ET 主要作用于两种内皮素受体 ETA 和 ETB , 但亲和力并不相同 , ET-1 和 ETA 的结合力是 ET-2 和 ETA 结合力的 2~10 倍 , 是 ET-3 同 ETA 结合力的 100 倍 , 而 3 种 ET 与 ETB 的结合力则相等²⁻⁵。

应用 ET 受体拮抗剂阻断 ET 与其受体的结合 , 有助于治疗 ET 参与发生的各种疾病。而克隆和表达 ET 受体基因 , 将极大的推动 ET 受体拮抗剂的筛选工作。本文以人 ETB 受体基因的保守序列为探针 , 从兔肺 cDNA 文库中筛选和克隆了兔 ETB 受体基因(rETBR 基因) , 并进行了基因全序列分析 , 为下一步开展 ET 受体拮抗剂的筛选工作打下了基础。

1 材料和方法

1.1 材料

兔肺 5'-Stretch cDNA 文库购自 CLONTECH 公司 , 限制性酶、T4 DNA 连接酶、Taq 聚合酶、PGEM-T Vector System、Prime-a Gene Labeling System 购自 Promega 公司 , pBlueScript II SK 由杨楠博士惠赠 , QuickPrep Micro mRNA Purification Kit、First-Strand cDNA Synthesis Kit 和 AutoRead Sequencing

Kit 购自 Pharmacia Biotech 公司 , QIAquick Gel Extraction Kit 和 QIAprep Spin Plasmid Kit 购自 QIAGEN 公司。

1.2 方法

1.2.1 兔肺 mRNA 的制备和 cDNA 第一链的合成 : 参照 Pharmacia Biotech 试剂盒操作手册进行。

1.2.2 探针制备 : 根据已发表的人 hET_BR 基因序列^[2] , 设计并合成一对特异性的 PCR 引物。PCR 反应如下 : 取 2μL cDNA , 加入特异的正反链 PCR 引物各 50 pmol/L , 4 × dNTP 100μmol/L , Taq DNA 聚合酶 2.5u , 补足 d₃H₂O 至 50μL。反应程序为 (1) 94℃ 变性 1 min ; 58℃ 复性 1 min ; 72℃ 延伸 1 min ; 共 30 次循环。 (2) 72℃ 10 min。PCR 产物用 0.8% 琼脂糖凝胶电泳进行分析。PCR 产物经 QIAquick Gel Extraction Kit 纯化后作为制备杂交探针的模板 ; 采用 Promega 公司的 Prime-a-Gene[®] Labeling System 进行标记。

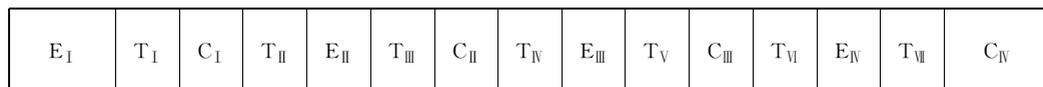
1.2.3 原位杂交筛选兔肺 cDNA 库 : 根据文献报道的方法进行原位杂交^[3]。每个阳性噬斑进行三轮杂交筛选直至挑纯。

1.2.4 克隆和 DNA 序列分析 : 平板裂解法^[6]制备的阳性 λ 噬菌体 DNA 经 EcoRI 酶切后 , 克隆进经 CIP 去磷酸化处理、EcoRI 酶切的窗口中。重组质粒转化进宿主细胞 DH5α , 并用 EcoRI 进行酶切鉴定。

2 结 果

2.1 rETBR 基因部分片段的 PCR 扩增以及探针的制备

从已发表的人、牛、鼠和猪的 ETB 受体基因结构中可发现 ETB 受体基因编码一个 G 蛋白偶联受体 , 具有 4 个胞外域 (E_I~E_{IV}) , 7 个跨膜结构域 (T_I~T_{VII}) 和 4 个胞内环 (C_I~C_{IV})^[2,4-6]。基因结构图如下 :



进一步的分析表明最保守的同源顺序发生在胞内环和疏水性的跨膜结构域部分。故本文以人的 hETBR 基因序列为依据,设计并合成了一对引物,预计将扩增出一段长为 394 bp 的 DNA 片段,它包含 III、IV 和 V 跨膜结构域。引物序列为:

P1 5'ACT GGC CAT TTG GAG CTG AGA 3'
 P2 5'CTG CAT GCC ACT TTT CTT TCT CAA 3'

其中,P1 全长 21 bp,GC 含量为 52.4%;P2 全长 24 bp,GC 含量为 41.7%。

以兔肺 cDNA 第一条链为模板,扩增出了一段长为 394 bp 的 DNA 片段,与预期大小一致。之后用随机引物法制备获得 DNA 探针。

2.2 原位杂交筛选兔肺 cDNA 库

实验确定噬菌体原种 10⁻⁶稀释度为工作稀释度。用随机引物法制备的[α-³²P]dCTP-DNA 探针进行兔肺 cDNA 文库的筛选,共筛选 20 个 90 mm 平板,三轮后获得 11 个阳性克隆。这 11 个克隆分别被命名为 λgt11-rETBR-1~λgt11-rETBR-11。EcoRI 酶切鉴定发现,阳性克隆 λgt11-rETBR-4 无法被切下插入片段,推测 EcoRI 位点发生突变,故利用 λgt11 的 5'/3' insert screening amplicon 扩增出插入子,并克

隆至 pGEM-T 质粒。其余阳性克隆均被切下插入片段,并克隆至 pBluescript II KS 上,构建得到重组质粒 pBluescript-rETBR-1~pBluescript-rETBR-11(除了 rETBR-4)。

2.3 λgt11-rETBR-4 的扩增和克隆

引物序列如下:

5'insert screening amplicon 5'-GAC TCC TGG AGC CCG-3'
 3'insert screening amplicon 3'-CGC GGC CAG CGA TGG-5'

将 PCR 扩增得到的约 680 bp DNA 片段直接克隆至线性化的 pGEM-T 载体质粒上,构建得到重组质粒 pGEM-T-rETBR-4。

2.4 阳性克隆的序列分析

对 11 个阳性克隆的序列测定结果表明,每个克隆均只含有 rETBR 基因的部分序列,其中 pBluescript-rETBR-1、pBluescript-rETBR-2 和 pGEM-T-rETBR-4 的插入片段分别长 1290 bp、1748 bp 和 687 bp,并可由这 3 个片段构建出 rETBR 基因的完整读码框。其余 8 个克隆均只含有 rETBR 基因的较小片段,因此未被采用来组装 rETBR 基因。

2.5 含有完整 rETBR 基因重组质粒的构建

经 DNASIS 软件分析插入子限制酶图谱后,对重组质粒的构建策略设计如下(见图 1):

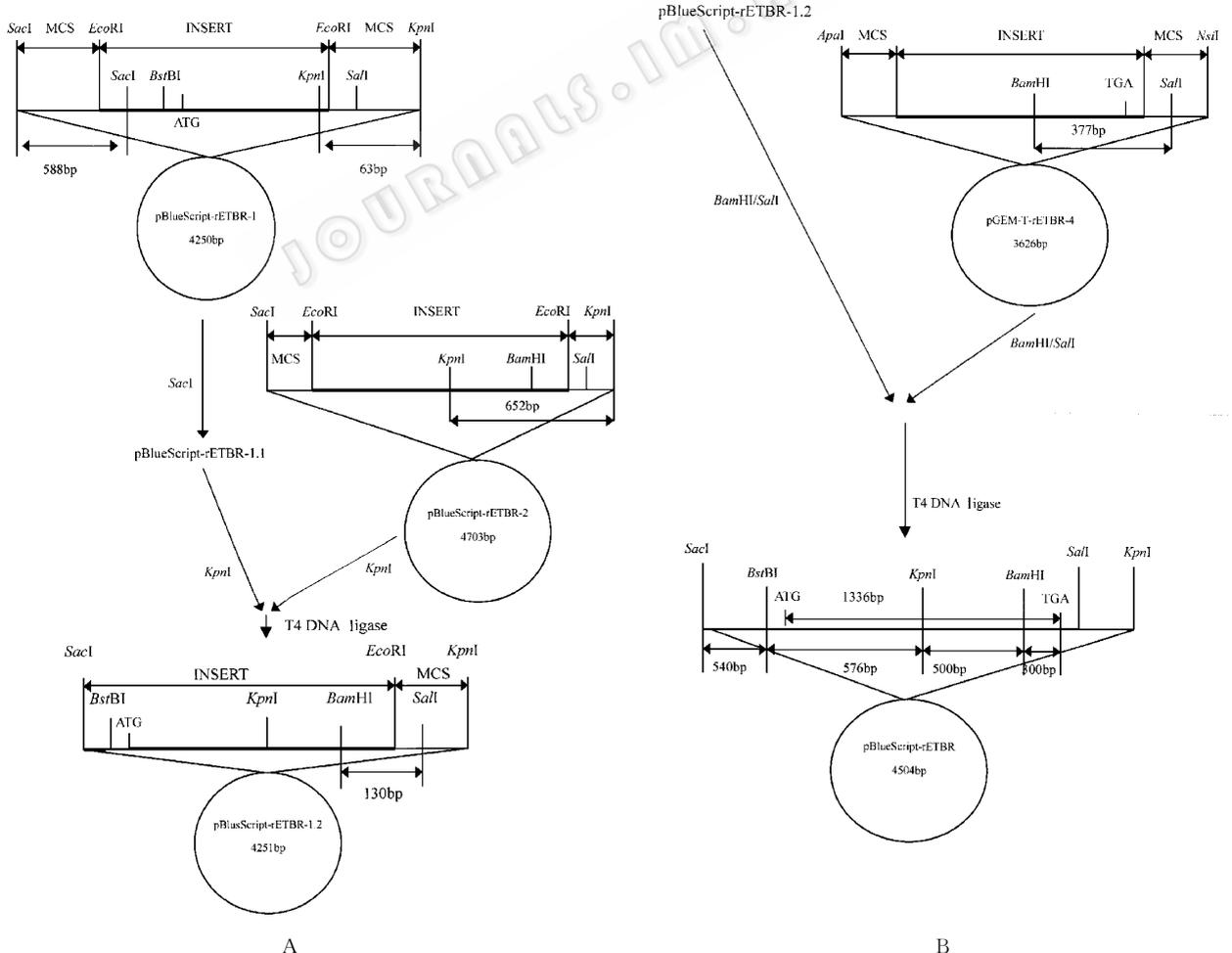


图 1 重组质粒 pBlueScript-rETBR 构建图

在 pBluescript II KS(-) 的多克隆位点中,第一个酶切位点为 *Sac*I,而在 pBluescript-rETBR-1 插入子的 5' 非编码区也有 *Sac*I 位点,故首先用 *Sac*I 单酶切 pBluescript-rETBR-1, 弃掉 588 bp 小片段(包括 insert 的 535 bp 和 MCS 的 53 bp), 回收大片段后进行连接,得到一个重组质粒 pBluescript-rETBR-1.1。

pBluescript-rETBR-1.1 和 pBluescript-rETBR-2 插入子 3' 端后的 MCS 中均有一个 *Kpn*I 位点,故用 *Kpn*I 分别单酶切 pBluescript-rETBR-1.1 和 pBluescript-rETBR-2,前者将得到两个片段—3599 bp 和 63 bp,回收大片段;后者亦将得到两个 4151 bp 和 652 bp 的片段,回收小片段;连接两个回收的

DNA 片段,用 *Sac*I 和 *Sal*I 双酶切鉴定方向,能切出 1338 bp 片段的质粒为正向,遂构建得到新的重组质粒 pBluescript-rETBR-1.2。

用 *Bam*HI 和 *Sal*I 分别双酶切 pBlueScript-rETBR-1.2 和 pGEM-T-rETBR-4,前者得到 2 个 4121 bp 和 130 bp 的片段,回收大片段;后者亦得到 2 个 3249 bp 和 377 bp 的片段,回收小片段,将两个回收的 DNA 片段进行连接,最终构建得到含有 rETBR 基因 1336 bp 完整读码框的重组质粒 pBlue-script-rETBR。

2.6 rETBR cDNA 的序列结果

rETBR 开放读码框全长 1326 bp,序列结果如下显示:

```

          BstBI                                     +1
ATGAGTTTGTTCGAAGGAGTGTGAGCGGGCCACTTGACGCCCCGCTAAAGCAGGCAGCAGCATGCAGCC 70
GCCGCCAAGCCTGTGCGGACTCGCCCTGCTGGCGTGGTTCTTGCCTGCGGGATGGCCGAGGTCTGGGGA 140
GAGGAGAGAGAAATGCCGTCTGCCCGCTACTCCACCGCTTCTGGGAGCCAGCGAGATCCTGACACCTT 210
CCACTAAGACCTCTTGGCCAAGGGATTCACACGCCAGCCTACCACGGTCGTCTGCACCTGCTGAAATACC 280
CAAAGAAGGGAGGACAGCCGGAGCCCGCGACGCCACCCCTCCTCCGTGTACAGACCAACAGAGATTAAG 350
GACACTTTCAAGTACATCAACTGTGGTATCCTGCCTCGTGTTCGTGCTGGGCATCATCGGGAACCTTA 420
CACTGCTGCGAATCATCTACAAGAACAAGTGCATGCGAAATGGCCCAATATCTTGATCGCCAGCCTGGC 490
TCTAGGAGACCTGCTGCACATCATATTGACATCCCCATCAATGTATAACAAGCTGCTGGCAGAAGACTGG 560
CCATTTGGAGCTGAGATGTGTAAGCTGGTACCTTTCATACAGAAGGCCCTCTGTGGGCATCACTGTGCTCA 630
GTCTATGTGCTTTAAGTATTGATAGATATCGAGCTGTTGCTTCTGGAGTCGAATTAAGGAATTGGGGT 700
TCCAAAGTGGACGGCAGTGAAATGTTTTGATTTGGTGGTCTCTGTGATTTGGCTGTCCCTGAAGCT 770
ATAGGTTTTAACCTGGTTACAATTGACTACAAAGGAAGTTACCTGCGAATCTGCTTGCTTAATCCCACTC 840
AAAAACAGCCTTCATGCAGTTTTATAAAACAGCTAAAGATTGGTGGCTATTTAGTTTCTATTTCTGCTT 910
ACCATGGCCATCACTGCCTTTTTTACTCTCTGATGACCTGTGAAATGTTAAGAAAGAAGAGTGAATG 980
CAAATTGCCTTAATGATCACTTAAAACAGAGACGGGAAGTGGCCAAAACGTATTTTGTCTGGTGGCTTG 1050
↑____P2 primer____↑
          BamHI
TTTTTGGCCTCTGTTGGCTGGCCCTTCACCTCAGCAGGATCCTGAAGCTCACACTTTATGATCAGAATGA 1120
CCCCAATAGATGTGAACCTTTTGGCTTTTGTGGTATTGGACTACATTGGAATCAATATGGCCTCCTTG 1190
AATTCCTGCATTAACCAATTGCTCTCTATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1260
TATGTTGCTGGTGCCAGTCATTTGAAGAAAAACAATCCTTGGAGGAAAAGCAGTCTCTGCTTGAAGTCAA 1330
AGCTAACGATCAGGATATGACAACCTCCGTTCCAGTAATAAATACAGCTCATCTTGAAAGAAGGCCGGA 1400
EcoRI
ATTC                                     1404

```

其中,第 1~562 碱基序列为 pBluescript-rETBR-1 插入片段中包含的序列、第 597~1112 碱基序列为 pBluescript-rETBR-2 插入片段中包含的序列、第 1097~1414 碱基序列为 pGEM-T-rETBR-4 的插入片段中包含的序列。

2.7 rETBR 基因核苷酸和氨基酸序列的同源性比较

rETBR 基因全长 1326 bp,与人的 hETBR 基因相比核苷酸序列共有 87.9%的同源性,推导的 rETBR 由 441 个氨基酸残基组成,与人的 hETBR 相比有 90.7%的同源性。

2.8 推断的氨基酸序列分析结果

根据 Kyte 和 Doolittle 方法⁷推测的 rETBR 折叠模式显示其与已建立的 G 蛋白偶联受体模型非常相似,具有 4 个胞外域(E_I~E_{IV})、7 个跨膜结构域(T_I~T_{VII})和 4 个胞内环(C_I~C_{IV})。rETBR 全长 441 个氨基酸,分子量为 49.44 kD。根据 von Heijne's 的方法⁸推测 N-端 18 个氨基酸残基为信号肽序列(Score = 11.11),18~19 之间为信号肽切割

处,因此成熟的 rETBR 全长应为 423 个氨基酸残基,分子量为 47.65 kD。

疏水性分析发现 rETBR 存在 7 个由 22~27 个疏水性氨基酸残基组成的疏水区,推测是跨膜结构域部分(I~VII)这一结构显示出与视紫质(Photoreceptor rhodopsin)和其它 G 蛋白偶联受体显著的序列和局部特征(Topographical)相似性;与其它超家族的膜受体相比,氨基酸的同源性主要集中在跨膜结构域部分。

在 N-端部分第一个跨膜结构域中,第 60 位和第 118 位为预测的 N-糖基化位点(1)60 :wprds N aslpr (2)118 :lgiig N stllr;第 186~202 位为 G 蛋白特征序列:kasvg ITVL-SLCALSIDRYRAV aswsr,进一步分析发现在 rETBR 氨基酸序列中还有一些潜在的磷酸化位点。推导的氨基酸及其特征序列显示如下:

```

      ↓
rETBR - MQPPPSLCGLLALLALVLAACGMAEVWGEEREMPSAPATPELLGASEILTPS -50
      Cho
rETBR - TKTSWPRDSNASLPRSSAPAEIPKEGRTAGAPRRTPPPCQRPTEIKDTFK -100
      Tr Cho Tr
rETBR - YINTVVVSVCLVFVLGLIGNSTLLRLIYKNCMRNGFNLLIASLALGDLLHI -150
      TIII
rETBR - IIDIPINVYKLLAEDWPFGAEMCKLVPFTOKASVGITVLSLCALSIDRYR -200
      (P) TIV
rETBR - AVASWSRIKIGIVPKWTAVEIVLWVSVLLAVPEAIGENLVTTIDYKGSY -250
      (P) (P) Tv
rETBR - LRLICLLNPTQKTAFMQFYKTAKDWWLFSEFYCLPLAITAFFYTLMTCEML -300
      (P) Tvi (P)
rETBR - RKKSGMQIALNDHLKQRREVAKTVFCLVLVEGLCWLALHLSRLKLTLYD -350
      TviI (P)
rETBR - QNDPNRCELLSFLLVLDYIGINMASLNSCINPIALYLVSKRFKNCFKSCL -400
      (P) (P)
rETBR - CCWCQSFEEKQSLEEKQSCLKFKANDHGYDNFRSSNKYSSS -441
↓: stands for the putative signal sequence cleavage site.
Cho: stands for the putative N-glycosylation site.
(P): stands for the putative phosphorylation site.
Tx: stands for the putative transmembrane domain.

```

3 讨 论

本研究获得的 DNA 片段含有一个长为 1336 bp 的开放读码框,与 GenBank 所收集的同类受体编码序列相比较,

该读码框与内皮素 B 受体基因有高达 90% 以上的同源性,而与内皮素 A 受体的同源性则不足 60%,与其它 G 蛋白偶联受体基因的同源性就更低。同源性比较结果参见下表。

	T _I	T _{II}	T _{III}	T _{IV}	T _V	T _{VI}	T _{VII}	E _I	E _{II}	E _{III}	E _{IV}	C _I	C _{II}	C _{III}	C _{IV}
人 ETB	100	96	100	96	100	92	100	71	100	81	92	100	100	100	100
猪 ETB	100	100	100	96	100	90	100	66	92	86	92	100	100	100	100
鼠 ETB	96	96	100	96	96	89	100	60	92	79	75	100	100	100	100
牛 ETA	64	55	77	64	77	73	76	4	50	48	75	0	76	68	34

将所获得的 DNA 片段进行氨基酸序列推导,发现其一级结构为典型的 G 蛋白偶联受体模型。已报道的研究结果

均证实内皮素 B 受体即属于 G 蛋白偶联受体,因此本研究克隆得到的 DNA 片段应为一个新的兔内皮素 B 受体基因。

参 考 文 献

- [1] Yanagisawa M ,Kurihara H ,Kimura S *et al.* *Nature* ,1988 ,**332** :441
- [2] Sakamoto A ,Yanagisawa M ,Sakurai T *et al.* *Biochem Biophys Res Commun* ,1991 ,**178** :656~663
- [3] Arai H ,Hori S ,Aramori I *et al.* *Nature* ,1990 ,**348** :730~731
- [4] Sakurai T ,Yanagisawa M ,Takuwa Y *et al.* *Nature* ,1990 ,**348** :732~735
- [5] Elshourbagy N ,Lee J ,Korman D *et al.* *Mol Pharmacol* ,1991 ,**41** :465~473
- [6] Sambrook J ,Fritsch E F ,Maniatis T. *Molecular Cloning-A Laboratory Manual* Second Edition. NY :Cold Spring Harbor Laboratory ,1989.
- [7] Kyte J ,Doolittle R F. *J Mol Biol* ,1982 ,**157** :105~132
- [8] von Heijne G. *Nucleic Acids Res* ,1986 ,**14** :4683~4690

A Novel Rabbit Endothelin B Receptor Gene Cloning and Sequence Analysis

YANG Lin WU Qiu-Yu LONG Qing-Xin WANG Xun-Zhang

(State Key Laboratory of Biological Control ,Biopharmaceutical Center ,Zhongshan University ,Guangzhou 510275)

YANG Nan

(The New Children 's Hospital ,Parramatta NSW 2124 ,Australia)

Abstract Endothelin(ET)is the most potent mammalian vasoconstrictor identified to date. As a pathogenic factor ,ET is involved in the genesis of many diseases. In this study ,a pair of primers was designed and synthesized according to the human ETB receptor gene(hETBR)sequence. A 394bp of DNA fragment was amplified by polymerase chain reaction(PCR) and labeled with α -³²P-CTP using Random Primer-Labeling method. With this probe ,rabbit lung cDNA library was screened by *in situ* hybridization and 11 positive clones were identified. Sequencing result showed that a complete reading frame of rabbit ETB receptor(rETBR)cDNA could be produced from three positive clones of eleven. By a series of sub-cloning ,a recombinant plasmid including the 1326 bp of rETBR coding sequences ,named pBlu Script-rETBR ,was constructed. The deduced amino acid sequence indicated that the rETBR is 441 residues in length ,with an expected molecular mass of approximately 49.44 kD. N-terminal 18 residues is the potential signal peptide(Score = 11.11)and therefore the molecular mass of mature rETBR is 47.65 kD with 423 amino acid residues. Analysis of the rETBR hydropathy profile indicates the presence of seven hydrophobic regions ,putative transmembrane domains. Potential N-glycosylation sites are the 60th and the 118 th. The structure exhibits a significant sequence and topographical similarity with G protein-coupled receptors.

Key words Rabbit endothelin B receptor , *in situ* hybridization ,cDNA cloning and sequencing