

人血小板生成素 cDNA 合成、克隆及序列测定

樊云祯 高彤 叶寅* 田波

(中国科学院微生物研究所 北京 100080)

摘 要 利用 RT-PCR 方法从中国人胎肝总 RNA 中扩增出人血小板生成素(hTPO)的 cDNA。序列分析结果表明,我们所获得的 hTPO cDNA 与文献报道中的基因序列高度同源,其中第 497bp、595bp、767bp 和 795bp 位碱基分别由 T、G、T 和 T 代替了文献中的 G、A、G 和 C,从而导致了 166、199 和 256 位氨基酸由报道中的 Ser、Lys 和 Gly 变为 Phe、Glu 和 Val。

关键词 人血小板生成素, cDNA, RT-PCR, 序列分析

人血小板生成素(hTPO)是人原癌基因 *c-mpl* 编码产物 MPL 的配体,是由 Kelemen 于 1958 年发现并命名的^[1]。研究表明, hTPO 是一种巨核细胞生长、发育、成熟和血小板产生的细胞系特异性调节因子,在体外能促进巨核细胞集落形成,在体内能增加血小板和巨核细胞的数目,临床上对原发性血小板减少症,尤其是骨髓移植、放疗和化疗的癌症患者体内血小板数量的降低有显著的治疗作用^[2~7],已经成为国内外竞相研究开发的新型药物。1994 年,国外几个研究小组几乎同时报道了 hTPO 基因序列,并对其进行了重组、表达、纯化、鉴定^[8~10]。hTPO 基因定位于染色体 3q27,是由 6 个外显子和 5 个内含子组成, cDNA 全长为 1774bp,带有 Poly (A)尾。其中 1062 个碱基的开放阅读框架编码一个 353 个氨基酸的蛋白。其 N-端有一个 21 个氨基酸的信号肽,成熟的 hTPO 蛋白由 332 个氨基酸组成,分为两个部分:与 EPO 高度同源的含 153 个氨基酸的氨基端和一个特异性的羧基端,而且其羧基端有多个潜在的 N-键糖基化位点。目前国内所报道的有关 hTPO 的 cDNA 合成等方面的工作,大都是以从国外购买的人组织 cDNA 文库为材料。

本研究以来源于国内的人胚肝为材料,提取其总 RNA,反转录和 PCR 合成、扩增了 TPO 的 cDNA 基因,分析了其 cDNA 基因序列,并与文献报道的 hTPO 基因进行了比较。

1 材料和方法

1.1 胚肝组织、菌株及质粒

所选用的人胚肝组织是由北京医科大学提供的引产男婴的胎肝。

E. coli DH5 α 、TG1,质粒 pUC19 和 pGEM7Z-f(+)均由本实验室提供。

1.2 工具酶与试剂

所用限制酶、T4 DNA 连接酶为 Promega 公司产品, Taq DNA 聚合酶为华美公司产品, cDNA 合成试剂盒、T4 多核苷酸激酶、T4 DNA 聚合酶及 Klenow 酶为 Boehringer 公司

* 联系人。

本文于 1996 年 11 月 27 日收到。

产品, T7 DNA 序列测定试剂盒为 Pharmacia 公司产品, α - 32 P-dCTP 和 α - 35 S-dATP 为 Dupont 公司产品, 其余化学试剂均为分析纯。

引物 P1 (5'-AGGCTAACGCATATGAGCCCG-3') 和 P2 (5'-GTTAGTCGACTTAC-CCTTCCTGAGACAG-3') 均为本所技术室合成。

1.3 人胚肝总 RNA 的提取及 cDNA 基因的合成

取 lg 冷冻的胎肝组织, 剪碎后冻融, 加 12ml 预冷的变性液 (4mol/L 异硫氰酸胍、0.83% 十二烷基肌氨酸钠、0.2mmol/L 巯基乙醇、42mmol/L 柠檬酸三钠) 研磨, 然后加入 1.2ml 2mol/L 醋酸钠 (pH4.0), 混合均匀, 再加入 12ml 酚: 氯仿: 异戊醇 (25:24:1) 混合均匀, 冰上放置 15min, 然后 10 000g、4℃ 离心 20min 后, 将上清转移到另一个 DEPC 处理过的 50ml 离心管中, 加入等体积的异丙醇, -20℃ 放置 30min 以上, 10 000g、4℃ 离心 15min 沉淀 RNA。用 5ml 变性液重新溶解沉淀, 再加等体积的异丙醇沉淀 RNA, 重复上述步骤, 得到的 RNA 沉淀用预冷的 75% 乙醇洗涤, 真空干燥, 沉淀用无 RNase 的水溶解, -20℃ 保存备用。

1.4 RT-PCR 法合成 TPO cDNA

在 oligo (dT)₁₅ 的引导下合成总 mRNA 的第一链 cDNA, 其反应体系为: 50mmol/L Tris-Cl (pH8.3), 75 mmol/L KCl, 5 mmol/L MgCl₂, 10 mmol/L DTT, 5 μ g 总 RNA、四种 dNTP (各 250 μ mol/L), 10pmol/L oligo(dT)₁₅, 200u M-MLV 反转录酶, 总体积 20 μ l, 37℃ 反应 1h。取 5 μ l cDNA 产物以 PCR 法扩增 TPO cDNA, 其反应混合物为: 10 mmol/L Tris-Cl, 50 mmol/L KCl, 3mmol/L MgCl₂, 10 pmol/L 引物 P1、P2, 四种 dNTPs (各 250 μ mol/L), 3u Taq 聚合酶, 总体积 100 μ l, 循环条件为: 94℃ 变性 45s, 50℃ 退火 1min, 72℃ 延伸 1min。反应结束后, 取 10 μ l 反应混合物在 0.7% 琼脂糖凝胶上电泳检测。

1.5 hTPO cDNA 的克隆

用 0.7% 低熔点琼脂糖凝胶纯化回收 PCR 产物中约 1000bp 的片段, 纯化的片段进行末端磷酸化。取上述处理 DNA 片段, 与经 SmaI 单酶切并经碱性磷酸酶处理的载体 pUC19 连接。连接产物转化大肠杆菌 TG1 感受态细胞, 铺麦康凯培养基培养 18~20h。随机挑取白斑, 用碱法快速提取质粒, 电泳筛选出大于载体的重组克隆, 然后进行酶切图谱分析以鉴定重组质粒。

1.6 序列分析

将重组质粒用合适的限制酶切割后亚克隆, 挑选重组亚克隆并对其进行序列测定^[11]。

2 结果与讨论

2.1 寡核苷酸引物的设计

本研究的目的是为了获得 TPO 全长 cDNA 序列, 在此基础上研究该基因在大肠杆菌和酵母中的表达。因此, 依据文献报道的 hTPO cDNA 序列, 不考虑 5' 端编码前 21 氨基酸信号肽的序列, 从成熟 TPO 的第一个密码子作为基因 5' 端起始设计引物, 其 5' 端引物 P1 为: 5'-AGGCTAACGCATATGAGCCCG-3' (加入一个 NdeI 位点); 3' 端利用其序列本身的终止密码子 TAA, 加入一个 SalI 位点, 设计的 3' 端引物 P2 为: 5'-GTTAGTCGACT-

TACCCTTCCTGAGACAG-3'。

2.2 TPO cDNA 合成和克隆

以胚肝组织总 RNA 的反转录产物为模板的 PCR 反应终止后,取 10 μ l 反应混合物在 0.7% 的琼脂糖凝胶上电泳分析,可见一个大小约 1kb 的特异性片段(图 1),将该片段回收后连接到载体 pUC19 上。因为载体经过去磷酸化处理,其自连的几率很低,同时用 SmaI 打开的载体的连接效率也很低,因此所得的阳性克隆也很少,阳性克隆率在 50% 左右。挑选得到的重组克隆进行限制性酶切分析(图 2),分别用 NdeI、SacI、PstI 和 BamHI 等酶切分析,结果表明与文献中 cDNA 序列中的限制性位点相同,初步确定所克隆的基因为人 TPO 基因,命名该质粒为 pUTPO。

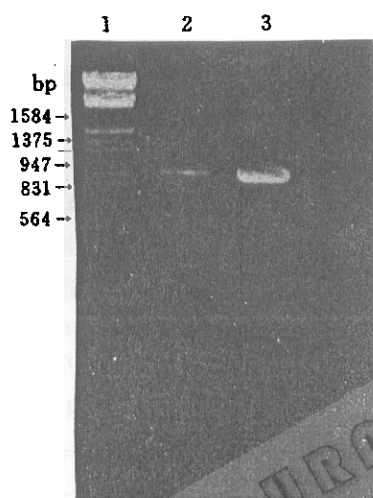


图 1 PCR 产物琼脂糖电泳

Fig.1 Analysis of RT-PCR products

1. λ DNA/HindIII + EcoRI markers
- 2,3. RT-PCR products

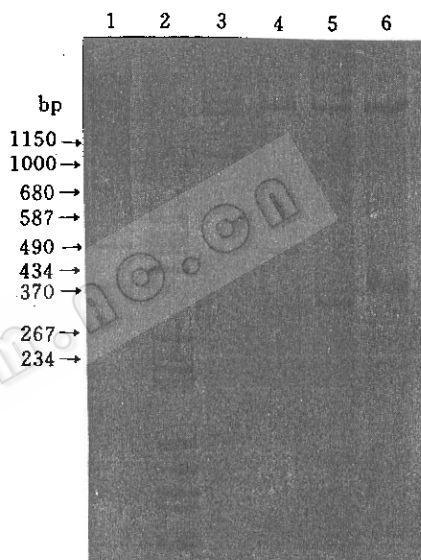


图 2 重组质粒 pUTPO 酶切鉴定

Fig.2 Restriction analysis of recombinant plasmid pUTPO

1. SppI/EcoRI markers
2. pBR322/HaeIII markers
3. pUTPO/NdeI 4. pUTPO/SacI
5. pUTPO/PstI 6. pUTPO/BamHI

2.3 TPO cDNA 的亚克隆和序列分析

为准确测定 TPO cDNA 序列,对重组克隆进行了亚克隆。将重组质粒分别用 XbaI 和 StyI 双酶切、BamHI 单酶切,回收大片段,自连,分别得到 5'端亚克隆 pUTPO1 和 3'端亚克隆 pUTPO3。然后再用 BamHI 和 StyI 双酶切重组质粒,得到的 375bp 片段与用 BamHI 和 XbaI 双酶切的 pGEM7Z-f(+)载体连接,得到克隆有中间片段的第三个亚克隆 pUTPO2。

对上述三个亚克隆序列分析得出了 TPO cDNA 全长(图 3)。由我们测定的序列可以看出,与已报道的 hTPO 序列相比,我们所克隆的基因在第 497bp、595bp、767bp 和 795bp 处与报道的碱基不同,分别由 T、G、T 和 T 代替了 G、A、G 和 C,其余序列与报道中的序列相同,同源率高达 99.5%。但是由于我们选用的是中国人胎肝组织,其中所得到的

AGC CCG GCT OCT OCT GCT TGT GAC CTC CGA GTC CTC AGT AAA CTG CTT CGT GAC TCC CAT GTC CTT	1-66
Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu	1-22
CAC AGC AGA CTG AGC CAG TGC CCA GAG GTT CAC OCT TTG OCT ACA OCT GTC CTG CTG OCT GCT GTG	67-132
His Ser Arg Leu Ser Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala Val	23-44
GAC TTT AGC TTG GGA GAA TGG AAA ACC CAG ATG GAG GAG ACC AAG GCA CAG GAC ATT CTG GGA GCA	133-198
Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys Ala Gln Asp Ile Leu Gly Ala	45-66
GTG ACC CTT CTG CTG GAG GGA GTG ATG GCA GCA CCG GGA CAA CTG GGA CCC ACT TGC CTC TCA TCC	199-264
Val Thr Leu Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser	67-88
CTC CTG GGG CAG CTT TCT GGA CAG GTC CGT CTC CTC CTT GGG GGC CTG CAG AGC CTC CTT GGA ACC	265-330
Leu Leu Gly Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu Leu Gly Thr	89-110
CAG CTT OCT CCA CAG GGC AGG ACC ACA GCT CAC AAG GAT CCC AAT GCC ATC TTC CTG AGC TTC CAA	331-396
Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp Pro Asn Ala Ile Phe Leu Ser Phe Gln	111-132
CAC CTG CTC CGA GGA AAG GTG CGT TTC CTG ATG CTT GTA GGA GGG TCC ACC CTC TGC GTC AGG CGG	397-462
His Leu Leu Arg Gly Lys Val Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Arg	133-154
GCC CCA CCC ACC ACA GCT GTC CCC AGC AGA ACC TTT CTA GTC CTC ACA CTG AAC GAG CTC CCA AAC	463-528
Ala Pro Pro Thr Thr Ala Val Pro Ser Arg Thr Phe Leu Val Leu Thr Leu Asn Glu Leu Pro Asn	155-176
AGG ACT TCT GGA TTG TTG GAG ACA AAC TTC ACT GCC TCA GCC AGA ACT ACT GGC TCT GGG CTT CTG	529-594
Arg Thr Ser Gly Leu Leu Glu Thr Asn Phe Thr Ala Ser Ala Arg Thr Thr Gly Ser Gly Leu Leu	177-198
GAG TGG CAG CAG GGA TTC AGA GGC AAG ATT OCT GGT CTG CTG AAC CAA ACC TCC AGG TCC CTG GAC	595-660
Glu Trp Gln Gln Gly Phe Arg Ala Lys Ile Pro Gly Leu Leu Asn Gln Thr Ser Arg Ser Leu Asp	199-220
CAA ATC CCC GGA TAC CTG AAC AGG ATA CAC GAA CTC TTG AAT GGA ACT CGT GGA CTC TTT OCT GGA	661-726
Gln Ile Pro Gly Tyr Leu Asn Arg Ile His Glu Leu Leu Asn Gly Thr Arg Gly Leu Phe Pro Gly	221-242
CCC TCA CGC AGG ACC CTA GGA GGC CCG GAC ATT TCC TCA GTA ACA TCA GAC ACA GGC TCC CTG CCA	727-792
Pro Ser Arg Arg Thr Leu Gly Ala Pro Asp Ile Ser Ser Val Thr Ser Asp Thr Gly Ser Leu Pro	243-264
CCT AAC CTC CAG CCT GGA TAT TCT OCT TCC CCA ACC CAT OCT OCT ACT GGA CAG TAT ACG CTC TTC	793-858
Pro Asn Leu Gln Pro Gly Tyr Ser Pro Ser Pro Thr His Pro Pro Thr Gly Gln Tyr Thr Leu Phe	265-286
OCT CTT CCA CCC ACC TTG CCC ACC OCT GTG GTC CAG CTC CAC CCC CTG CTT OCT GAC OCT TCT GCT	859-924
Pro Leu Pro Pro Thr Leu Pro Thr Pro Val Val Gln Leu His Pro Leu Leu Pro Asp Pro Ser Ala	287-308
OCA ACG CCC ACC OCT ACC AGC OCT CTT CTA AAC ACA TCC TAC ACC CAC TCC CAG AAT CTG TCT CAG	925-990
Pro Thr Pro Thr Pro Thr Ser Pro Leu Leu Asn Thr Ser Tyr Thr His Ser Gln Asn Leu Ser Gln	309-330
GAA GGG TAA	991-999
Glu Gly ***	331-332

图 3 hTPO cDNA 序列及推测的氨基酸序列

Fig. 3 Nucleotide sequence and deduced amino acid sequence of hTPO cDNA

Asterisks show the nucleotides that differ from the counterparts of hTPO cDNA gene from GeneBank data base[accession nos. L36052].

TPO 基因可能与直接从国外购进的 cDNA 文库中调出的 TPO 基因的序列有所差异。而且,我们测出的序列中与文献序列不同的碱基在 497、595、767、795bp 位点,其中 795bp 位碱基替换是中性突变,其余 3 个不同的碱基导致第 166、199 和 256 位氨基酸分别由报道

中的 Ser、Lys 和 Gly 被 Phe、Glu 和 Val 所替代(均在第 153 个氨基酸之后),这与一般认为的 hTPO N-端前 153 个氨基酸的序列与 EPO 同源性很高且高度保守是一致的。

在本实验中,我们用同一胎肝组织多次提取总 RNA 并经过多次 RT-PCR,得到的约 1kb 的片段克隆到 pUC19 载体上,多次测序的结果表明这几个克隆的序列是一致的。这说明至少对这一个体而言,中国人的 TPO 基因序列与从基因文库中扩增得到的 hTPO 的 cDNA 基因序列是不同的。但是因为材料来源有限,并未对多个胎肝组织的 TPO 基因进行序列测定,所以我们测定的序列是否能代表中国人 TPO 基因序列有待于进一步证明。

目前,我们已将 hTPO cDNA 克隆到大肠杆菌和酵母表达载体上,正在进行表达研究。

参 考 文 献

- [1] Kelemen E, Cserhati I, Tanos B *et al.* Heta Haematol Basel, 1958, 20:350.
- [2] Wendling F, Maraskovsky E, Debili N *et al.* Nature, 1994, 369:571.
- [3] Kaushansky K, Lok S, Holly R D *et al.* Nature, 1994, 369:568.
- [4] Bartly T D, Bogenberger J, Hunt P *et al.* Cell, 1994, 177:1117.
- [5] Broudy V C, Blood, 1995, 86(1):23.
- [6] Kaushansky K, Broudy V C, Lin N, *et al.*, Proc Natl Acad Sci, USA, 1995, 92:3234.
- [7] Farese A M. Blood, 1995, 54.
- [8] Sohma Y, Akahori H, Seki N, *et al.* FEBS Letters, 1994, 353:57.
- [9] de Sauvage F J, Hass P E, Spencer S D *et al.*, Nature, 1994, 369:533~538.
- [10] Foster D C, Sprecher C A, Grant F J, *et al.* Proc Natl Acad Sci USA, 1994, 91:13023.
- [11] Yie Y, Wei Z Y, Tien P. Nucleic Acids Res. 1993, 21(1):361.

cDNA Synthesis, Cloning and Sequencing of Human Thrombopoietin

Fan Yunzhen Gao Tong Ye Yin Tian Bo

(Institute of Microbiology, Academia Sinica, Beijing 100080)

Abstract Human thrombopoietin (hTPO), a lineage-specific cytokine affecting the proliferation and maturation of megakaryocytes, is the major humoral regulator of megakaryo-cytopoiesis and platelet production. A human TPO cDNA was obtained by RT-PCR from chinese human fetal liver. The target gene was cloned and sequenced. The data indicated that there are four bases different from the counterparts previously reported, which leads to the changes of three amino acid residues.

Key words hTPO, cDNA, RT-PCR, DNA sequencing