

[19]中华人民共和国专利局

[51]Int.Cl⁶



[12] 发明专利申请公开说明书

C12N 15/19
 C07K 14/52 C07K 16/24
 C12N 1/21 C12N 5/10
 C12N 1/19 A01K 67/027
 A61K 38/19 C12Q 1/68
 C12N 15/11

[21] 申请号 94194997.4

[43]公开日 1997年4月23日

[11] 公开号 CN 1148408A

[22]申请日 94.8.5

[30]优先权

[32]94.2.14 [33]US[31]08 / 196,025

[32]94.2.25 [33]US[31]08 / 203,197

[32]94.3.21 [33]US[31]08 / 215,203

[32]94.6.1 [33]US[31]08 / 252,491

[86]国际申请 PCT / US94 / 08806 94.8.5

[87]国际公布 WO95 / 21920 英 95.8.17

[85]进入国家阶段日期 96.8.13

[71]申请人 津莫吉尼蒂克斯公司

地址 美国华盛顿

共同申请人 华盛顿大学

[72]发明人 R·D·霍利 S·罗克

D·C·弗斯特 F·S·哈根

K·考山斯奇 J·F·贵皮尔

C·罗夫顿-得 P·J·欧特

S·K·伯克哈德

[74]专利代理机构 中国国际贸易促进委员会专利商标
事务所

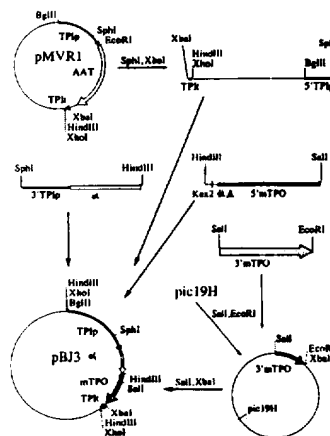
代理人 李 瑛

权利要求书 7 页 说明书 97 页 附图页数 2 页

[54]发明名称 血细胞生成蛋白及其制造材料和方法

[57]摘要

本发明提供了血细胞生成蛋白及其多肽片段，包括小鼠和人的蛋白质和多肽。本发明也提供了编码这些蛋白质和多肽的 DNA 分子，以及在其生产中有用的载体和细胞。本发明还提供了与所述蛋白质上的抗原决定簇结合的抗体。这些蛋白质和多肽对体内和体外治疗有用，也可以作为细胞培养和研究细胞增生和发育的试剂。



权 利 要 求 书

1.一种选自由以下物质组成的组的分离的蛋白质。

a) 包含 SEQ ID No:2 的从第 45 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列的蛋白质。

b) 包含 SEQ ID No:2 的从第 45 个氨基酸残基到第 206 个氨基酸残基的氨基酸序列的蛋白质。

c) 包含 SEQ ID No:19 的从第 22 个氨基酸残基到第 173 个氨基酸残基的氨基酸序列的蛋白质。

d) 包含 SEQ ID No:19 的从第 22 个氨基酸残基到第 175 个氨基酸残基的氨基酸序列的蛋白质。

e) (a), (b), (c), (d) 的等位变异体; 和

f) (a), (b), (c), (d) 或 (e) 的物种同系物, 其中所述的蛋白质刺激骨髓样或淋巴样前体的增生或分化。

2.一种按照权利要求 1 的分离的蛋白质, 其中所述的蛋白质包含 SEQ ID No:2 的从第 45 个氨基酸残基到第 379 个氨基酸残基的氨基酸序列。

3.一种按照权利要求 1 的分离的蛋白质, 其中所述的蛋白质包含 SEQ ID No:19 的从第 22 个氨基酸残基到第 353 个氨基酸残基的氨基酸序列。

4.一种按照权利要求 1 的分离的蛋白质, 其中所述的蛋白质为一种小鼠蛋白质。

5.一种按照权利要求 1 的分离的蛋白质, 其中所述的蛋白质为一种人体蛋白质。

6.一种按照权利要求 1 的分离的蛋白质, 其中所述的蛋白质包括:

SEQ ID No:2 中所示的第 45 个氨基酸残基到第 379 个氨基酸残基的氨基酸序列;

SEQ ID No:2 中所示的第 24 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列;

SEQ ID No:2 中所示的第 24 个氨基酸残基到第 206 个氨基酸残基的

氨基酸序列；

SEQ ID No:2 中所示的第 24 个氨基酸残基到第 379 个氨基酸残基的氨基酸序列；

SEQ ID No:2 中所示的第 1 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列；

SEQ ID No:2 中所示的第 1 个氨基酸残基到第 206 个氨基酸残基的氨基酸序列；或

SEQ ID No:2 中所示的第 1 个氨基酸残基到第 379 个氨基酸残基的氨基酸序列。

7.一种按照权利要求 1 的分离的蛋白质，其中所述蛋白质包括：

SEQ ID No:19 中所示的第 1 个氨基酸残基到第 173 个残基的氨基酸序列；

SEQ ID No:19 中所示的第 1 个氨基酸残基到第 175 个残基的氨基酸序列；

SEQ ID No:19 中所示的第 1 个氨基酸残基到第 353 个残基的氨基酸序列；或

SEQ ID No:19 中所示的第 22 个氨基酸残基到第 353 个残基的氨基酸序列。

8.一种基本上由选自以下序列组成的组的氨基酸序列组成的分离的蛋白本质：

SEQ ID No:2 中从第 45 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列；

SEQ ID No:2 中从第 45 个氨基酸残基到第 206 个氨基酸残基的氨基酸序列；

SEQ ID No:2 中从第 45 个氨基酸残基到第 379 个氨基酸残基的氨基酸序列；

SEQ ID No:19 中从第 22 个氨基酸残基到第 175 个氨基酸残基的氨基酸序列；和

SEQ ID No:19 中从第 22 个氨基酸残基到第 353 个氨基酸残基的氨基酸序列。

9.一种可刺激骨髓样或淋巴样前体增生或分化的分离的蛋白质，其中所述的蛋白质包含与 SEQ ID No:2 中的第 45 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列或 SEQ ID No:19 中的第 22 个氨基酸残基到第 173 个氨基酸残基的第 173 个氨基酸残基的氨基酸序列在氨基酸水平上至少具有 80 % 的同源性的区段。

10.一种编码按照权利要求 1 的蛋白质的分离的多核苷酸分子。

11.一种按照权利要求 10 的分离的多核苷酸分子，其中该分子为包含一条编码链的 DNA 分子，所述编码链包含 SEQ ID No:1 中的第 237 个核苷酸到第 692 个核苷酸的核苷酸序列。

12.一种按照权利要求 10 的分离的多核苷酸分子，其中该分子为包含一条编码链的 DNA 分子，所述编码链包含 SEQ ID No:18 中的第 64 个核苷酸到第 519 个核苷酸的核苷酸序列。

13.一种按照权利要求 10 的分离的多核苷酸分子，其中该分子编码 SEQ ID No:2 中的第 45 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列。

14.一种按照权利要求 10 的分离的多核苷酸分子，其中该分子编码 SEQ ID No:19 中的第 22 个氨基酸残基到第 173 个氨基酸残基的氨基酸序列。

15.一种选自由以下物质组成的组的分离的多核苷酸分子：

(a) 编码血细胞生成蛋白并包含 SEQ ID No:1 中所示的第 237 个核苷酸到第 692 个核苷酸的核苷酸序列的 DNA 分子；

(b) 编码血细胞生成蛋白并包含 SEQ ID No:18 中所示的第 64 个核苷酸到第 519 个核苷酸的核苷酸序列的 DNA 分子；

(c) (a) 或 (b) 的等位变异体；

(d) 编码血细胞生成蛋白的 DNA 分子，该蛋白与 (a)，(b) 或 (c) 编码的蛋白质在氨基酸序列方面至少有 90 % 的同源性；和

(e) 与 (a)，(b)，(c) 或 (d) 互补的分子。

16.一种按照权利要求 15 的分离的多核苷酸分子，其中该分子编码的血细胞生成蛋白与 (a)，(b) 或 (c) 编码的蛋白在氨基酸序列方面至少具有 90 % 的同源性。

17.一种按照权利要求 15 的分离的多核苷酸分子，其中该分子包含 SEQ ID No:1 的第 237 个核苷酸到 722 个核苷酸或 SEQ ID No:18 的第 64 个核苷酸到第 525 个核苷酸。

18.一种选自由以下物质组成的组的分离的 DNA 分子：

(a) 质粒 pZGmpl-1081 (ATCC 69566) 的 EcoRI - XhoI 插入片段；

(b) (a) 的等位变异体；和

(c) 编码与 (a) 或 (b) 编码的蛋白质在氨基酸序列方面至少具有 80 %同源性的蛋白质的 DNA 分子，其中该分离的 DNA 分子编码具有血细胞生成活性的蛋白质。

19.一种按照权利要求 18 的分离的 DNA 分子，其中该分子编码包含 SEQ ID No:2 中的第 45 个氨基酸残基至第 196 个氨基酸残基的氨基酸序列的多肽。

20.一种包含以下可操作连接的元件的表达载体：

一个转录启动子；

一个选自由以下物质组成的组的 DNA 区段：

(a) 编码血细胞生成蛋白并包含 SEQ ID No:1 中所示的第 237 个到第 692 个核苷酸的核苷酸序列的 DNA；

(b) 编码血细胞生成蛋白并包含 SEQ ID No:18 中所示的第 64 个核苷酸到第 519 个核苷酸的核苷酸序列的 DNA；

(c) (a) 或 (b) 的等位变异体；和

(d) 编码血细胞生成蛋白的 DNA 分子，该蛋白与 (a)、(b) 或 (c) 编码的蛋白在氨基酸序列方面至少具有 80 %的同源性和一个转录终止子。

21.一种按照权利要求 20 的表达载体，其中所述的 DNA 区段编码的血细胞生成蛋白与 (a)、(b) 或 (c) 编码的蛋白在氨基酸序列方面至少具有 90 %的同源性。

22.一种按照权利要求 20 的表达载体，其中所述的 DNA 区段包含 SEQ ID No:1 中的第 237 个核苷酸到第 722 个核苷酸或 SEQ ID No:18 中的第 64 个核苷酸到第 525 个核苷酸。

23.一种按照权利要求 20 的表达载体, 该载体还包括可经操作与所述 DNA 区段连接的分泌信号序列。

24.一种培养细胞, 该细胞中已引入了按照权利要求 20 的表达载体, 其中所说的细胞表达的由 DNA 区段编码的血细胞生成蛋白。

25.一种按照权利 24 的培养细胞, 其中所说的细胞为一种真菌细胞。

26.一种按照权利要求 25 的培养细胞, 其中所说的细胞为一种酵母细胞。

27.一种按照权利要求 24 的培养细胞, 其中所说的细胞为一种哺乳动物细胞。

28.一种按照权利要求 24 的培养细胞, 其中所说的细胞为一种细菌细胞。

29.一种非人哺乳动物, 该哺乳动物种系中已引入选自以下物质组成的组的异说 DNA 区段:

(a) 编码血细胞生成蛋白并包含 SEQ ID No:1 中所示的第 64 个核苷酸到第 519 个核苷酸的核苷酸序列的 DNA 区段;

(b) 编码血细胞生成蛋白并包含 SEQ ID No:18 中所示的第 64 个核苷酸到第 519 个核苷酸的核苷酸序列的 DNA 区段;

(c) (a) 或 (b) 的等位变异体; 和

(d) 编码与 (a), (b) 或 (c) 编码的蛋白在氨基酸序列方面至少具有 80%同源性的血细胞生成蛋白的 DNA 区段;

其中上述哺乳动物产生由所述 DNA 区段编码的血细胞生成蛋白。

30.一种按照权利要求 29 的非人哺乳动物, 该哺乳动物选自由猪、山羊、绵羊、牛和小鼠组成的组。

31.一种按照权利要求 29 的非人哺乳动物, 其中所述 DNA 区段包含 SEQ ID No:1 的第 237 个核苷酸到第 722 个核苷酸或 SEQ ID No:18 的第 64 个核苷酸到第 525 个核苷酸。

32.一种产生血细胞生成蛋白的方法, 该方法包括: 培养已引入按照权利要求 20 的表达载体的细胞, 其中所说的细胞表达由所说 DNA 区段编码的血细胞生成蛋白。

回收血细胞生成蛋白。

33.一种按权利要求 32 的方法，其中所说的血细胞生成蛋白由所述细胞分泌并从培养所述细胞的培养基中回收。

34.一种药物组合物，该组合物包含按照权利要求 1 的蛋白质和药理学上可接受的载体。

35.一种抗体，该抗体结合按照权利要求 1 的蛋白质的抗原决定簇。

36.一种在哺乳动物中刺激血小板产生的方法，该方法包括给所述哺乳动物施用治疗有效量的选自以下物质组成的组的血细胞生成蛋白：

a) 包含 SEQ ID No:2 的从第 45 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列的蛋白质；

b) 包含 SEQ ID No:19 的从第 22 个氨基酸残基到第 173 个氨基酸残基的氨基酸序列的蛋白质；

c) (a) 或 (b) 的等位变异体；和

d) (a)，(b) 或 (c) 的物种同系物。

其中所述的蛋白质与药理学上可接受的载体一起刺激骨髓样或淋巴样前体增生或分化。

37.一种探针，该探针包含至少 14 个核苷酸的寡聚核苷酸，其中所述的寡聚核苷酸的序列与以下序列的同一长度部分至少有 80 % 的同源性；

(a) SEQ ID No:1；

(b) SEQ ID No:18；

(c) SEQ ID No:28；或

(d) 与 SEQ ID No:1，SEQ ID No:18 或 SEQ ID No:28 互补的序列。

38.一种在 DNA 分子的混合物中检测编码血小板生成素的 DNA 分子的方法，该方法包括用探针探测 DNA 分子的混合物，所述探针包含至少 14 个核苷酸的寡聚核苷酸，其中所述的寡聚核苷酸的序列与以下序列的同一长度部分至少有 80 % 的同源性：

(a) SEQ ID No:1；

(b) SEQ ID No:18；

(c) SEQ ID No:28; 或

(d) 与 SEQ ID No:1, SEQ No:18 或 SEQ ID No:28 互补的序列;

和

检测与所述探针杂交的 DNA 分子。

39.一种刺激细胞增生的方法,该方法包括向培养的骨髓细胞中以足以刺激细胞增生的量加入按照权利要求 1 的分离的蛋白质。

40.一种按照权利要求 39 的方法,其中所述的细胞为巨核细胞或巨核细胞前体。

41.一种提纯血小板生成素的方法,该方法包括:

把含血小板生成素的溶液暴露于附加有固体支持物的抗体中,按照权利要求 1,其中的所述的抗体与按照权利要求 1 的蛋白质的抗原决定簇结合;

洗涤所述的抗体,去除未结合的污染物;

从所述的抗体中洗脱结合的血小板生成素; 和

回收所述洗脱的血小板生成素。

说 明 书

血细胞生成蛋白及其制造材料和方法

相关申请的交互参见

本申请是案号为 08/215, 203 的申请 (1994 年 3 月 21 日申请) 的部分后续申请, 后者是案号为 08/203197 的申请 (1994 年 2 月 25 日申请) 的部分后续申请, 08/203,197 号申请又是案号为 08/196,025 的申请 (1994 年 2 月 14 日申请) 的部分后续申请, 这些申请是未决申请, 本文一并参考。

发明背景

血细胞生成 (Hematopoiesis) 是一个使血细胞从骨髓中多能性干细胞发育并分化的过程。该过程涉及通过靶细胞上膜结合受体起作用的多肽生长因子 (细胞因子) 的复合物相互作用。细胞因子的作用导致细胞增生和分化, 对一个特别的细胞因子的反应通常为谱系特异性的和阶段特异性的。一个单细胞型 (如血小板) 从干细胞的发育可能需要多种细胞因子以合适的次序的协同作用。

已知的细胞因子包括白细胞介素 (如 IL - 1, IL - 2, IL - 3, IL - 6, IL - 8, 等等) 和集落刺激因子 (如 G - CSF、M - CSF, GM - CSF, 促红细胞生成素 (EPO), 等等)。一般来说, 白细胞介素充当免疫和炎症反应介体的作用。集落刺激因子刺激骨髓来源细胞的增生, 激活成熟白细胞, 并且另外形成对炎症、侵袭和免疫激发的寄生反应的一个组成部分。

各种细胞因子已开发为治疗剂。例如, 促红细胞生成素, 它可刺激红细胞的发育, 用来治疗肾衰竭引起的贫血。几种集落刺激因子与癌症的化学疗法一起已被用于加速病人免疫系统的康复。白细胞介素-1 β , α -干扰素和 β -干扰素用于疗某些癌症。在患血小板减少症的动物体液中, 已鉴别出一种可刺激巨核细胞生成和血小板生成的活性, 在文献中被称之为“血小板生成素 (thrombopoietin)” (McDonald 最近的综

述, Exp. Hematol. 16:201 - 205, 1988 和 McDonald, Am. J. Ped. Hematol. Oncol. 14:8 - 21, 1992)。虽然有三十多年的研究, 但尚未确定对产生这种活性起决定作用的某个特定因子或多种因子的特征, 部分原因是由于缺乏好的材料来源, 缺乏好的试验方法和缺乏像产生位点方面的知识。

与血小板机能失调相关的温和出血失调症, 相对较为常见 (Bachmann, Seminars in Hematology 17:292 - 305, 1980), 它是一些血小板功能的光天性失调, 包括 Bernard - Soulier 综合症 (缺乏血小板 GpIb) 格兰茨曼血小板机能不全病 (缺乏 GpIIb 和 GpIIIa) 和先天性无纤维蛋白原血症 (血浆和血小板中纤维蛋白原的减少或缺乏) 以及灰色血小板综合症 (缺少 α - 颗粒)。此外, 还有一些与血小板分泌, 贮藏库 (storage pool) 缺损, 血小板花生四烯酸途径畸变, 血小板环加氧酶和凝血恶烷合成酶的缺损及血小板激活的缺损有关的失调症 (reviewed by Rao and Holmsen, Seminars in Hematology 23:102 - 118, 1986)。目前, 尚未阐明大部分这些缺损的分子基础。

血小板蛋白的分离和特征描述对许多血小板机能失调的缺陷基础的阐明提供了非常宝贵的工具。对详细的分子分析的一个主要限制步骤在于很难从血小板或其前体巨核细胞中获得 mRNA, 以便分析和 cDNA 文库构建。血小板缺乏细胞核和转录作用。仍然与血小板相联的微量的 mRNA 很难分离并常常易于降解。血小板 cDNA 文库的构建至今为止需要大量的血小板, 典型地需要来自 25 至 250 个单位全血的血小板 (Izumi et al., Proc. Natl. Acad. Sci. USA 87:7477 - 7481, 1990; Wicki et al., Thrombosis and Haemostasis 61:448-453, 1989; and Wenger et al., Blood 73:1498 - 1503, 1989) 或来自具有实质性血小板增多引起的升高的血小板计数的病人的 Pheresis 的血小板 (Roth. et al, Biochem. Biophys. Res. Comm. 160:705-710, 1989)。分离出血小板特异性 cDNA 的地方, mRNA 很可能是最稳定的或总 mRNA 种类丰富的, 并且很可能只代表一个部分总的血小板编码库。

获得一个血小板 cDNA 文库的另一途径是由巨核细胞 (即血小板的直接细胞前体) 分离的 mRNA 的分离和文库的构建。巨核细胞为多倍体细胞, 预期包含编码全套血小板和巨核细胞蛋白质的 mRNA。然而, 现

已证明很难分离出足够数量和纯度的巨核细胞。

分子生物学的最近进展大大地增加了我们对血细胞生成的理解，但同时已表明它是一个十分复杂的过程。尽管对许多细胞因子已确定特征，并且有些已证实有临床用途，但本领域尚需要一些其它药剂，刺激髓样和淋巴样前体的增生和分化以及成熟血细胞的产生。尤其需要刺激巨核细胞谱系的细胞（包括血小板）发育和增生的药剂。本领域还需要可用于治疗血细胞减少症（包括血小板减少，即循环的血小板数量低得反常的状况（约不到 1×10^5 血小板/mm³），及其他血小板失调症）的药剂。本发明可满足这些需要并提供其他相关的优点。

发明概要

本发明的一个目的是提供具有血细胞生成（hematopoietic）活性的分离的蛋白质。

本发明的另一个目的是提供生产具有血细胞生成活性的蛋白质的方法及该方法中使用的DNA分子，载体和细胞。

本发明的另一个目的是提供一些可结合血细胞生成蛋白（hematopoietic protein）上的抗原决定簇的抗体。

本发明的另一个目的是提供在包括人在内的哺乳动物中刺激巨核细胞，血小板和嗜中性白细胞产生的方法。

本发明的另一个目的是提供用于骨髓细胞发育、分化和增生研究和骨髓细胞发育，分化和增生过程中以畸变为特征的疾病检测的各种工具。

一方面，本发明提供了一种分离的蛋白质，该蛋白质选自：（a）包含 SEQ ID No:2 的第 45 个氨基酸残基到第 196 个氨基酸残基的氨基酸序列的蛋白质；（b）包含 SEQ ID No:2 的第 45 个氨基酸残基到第 206 个氨基酸残基的氨基酸序列的蛋白质；（c）包含 SEQ ID No:19 的第 22 个氨基酸残基到第 173 个氨基酸残基的氨基酸序列的蛋白质；（d）包含 SEQ ID No:19 的第 22 个氨基酸残基到第 175 个氨基酸残基的氨基酸序列的蛋白质；（e）（a）、（b）、（c）、（d）的等位变异体；和（f）（a）、（b）、（c）、（d）、（e）的种类同系物，其中所述蛋白质刺激髓样和淋巴前体（precursor）的增生和分化。在某些实

施方案中，所述蛋白质包含 SEQ ID No:2 的第 45 个氨基酸残基至第 379 个氨基酸残基的氨基酸序列或包含 SEQ ID No:2 的第 22 个氨基酸残基至第 353 个氨基酸残基的氨基酸序列。

在相关的一方面，本发明提供了一种编码上述蛋白质的分离的多核苷酸分子。在一个实施方案中，所述多核苷酸分子为一个 DNA 分子，该 DNA 分子包含一个编码链，所述编码链包含 SEQ ID No:1 的第 237 个核苷酸至第 692 个核苷酸的核苷酸序列或 SEQ ID No:18 的第 64 个核苷酸至第 519 个核苷酸的核苷酸序列。在其他实施方案中，该分子包含 SEQ ID No:1 的 237 - 1241，174-1241，105 - 1241，105 - 722，174 - 722 或 237 - 722 位核苷酸或 SEQ ID No:18 的相应区。本发明进一步提供了这些分子和编码血细胞生成蛋白的 DNA 分子的等位变异体，所述分子编码的蛋白质在氨基酸序列上与由 SEQ ID No:1 或 SEQ ID No:18 的以上叙述的任一部分编码的蛋白质至少有 80 % 的同源性。本发明也提供了与这些序列互补的分子。

另一方面，本发明提供了一种分离的 DNA 分子，该分子选自：(a) 质粒 pZGmpL - 1081 (ATCC 69566) 的 EcoRI - XhoI 插入片段，(b) (a) 的等位变异体，和 (c) 编码氨基酸序列与 (a) 或 (b) 编码的蛋白质至少有 80 % 的同源性的蛋白质的 DNA 分子，其中，分离的 DNA 分子编码具有血细胞生成活性的蛋白质。

另一方面，本发明提供一种表达载体，该载体包含下列可操作连接的元件 (element)：一个转录启动子；一个 DNA 区段 (segment)，该区段选自：(a) 编码血细胞生成蛋白并包含 SEQ ID No:1 中所示的第 237 至第 692 个核苷酸的核酸序列的 DNA 区段，(b) 编码血细胞生成蛋白质并包含 SEQ ID No:18 中所示的第 64 至第 519 个核苷酸的核酸序列的 DNA 区段，(c) (a) 和 (b) 的等位变异体，以及 (d) 编码氨基酸序列与 (a)、(b) 或 (c) 编码的蛋白质至少有 80 % 的同源性的血细胞生成蛋白质的 DNA 区段；和一个转录终止子。

另一方面，本发明提供了一种已导入上述表达载体的培养细胞，其中，细胞表达被所述 DNA 区段编码的血细胞生成蛋白。在某些实施方案中，所述细胞为真菌细胞，哺乳动物细胞或细菌细胞。

另一方面，本发明提供了一种已向某种系导入编码上述血细胞生成蛋白质的异源 DNA 分子的非人体哺乳动物，其中，哺乳动物产生由所述 DNA 区段编码的血细胞生成蛋白质。

另一方面，本发明提供了在哺乳动物中刺激血小板产生的方法。这些方法包括给哺乳动物施用治疗有效量的血细胞生成蛋白，该血细胞生成蛋白，选自：（a）包含 SEQ ID No:2 的第 45 个氨基酸残基至第 196 个氨基酸残基的氨基酸序列的蛋白质；（b）包含 SEQ ID No:19 的第 22 个氨基酸残基至 173 个氨基酸残基的氨基酸序列的蛋白质；（c）（a）和（b）的等位变异体；和（d）（a）、（b）或（c）的物种同系物（Species homologs），其中，所述的蛋白质与药理上可接受的载体一起刺激髓样或淋巴样前体的增生或分化。

参考下面详细的描述及附图，本发明的这些方面及其他方面将显得十分清楚明了。

附图简要说明

图 1 为载体 pDX 的部分限制性酶切图谱。使用的记号 SV40，Ori，表示 SV40 的复制起点；SV40E，为 SV40 增强子；MLP，为腺病毒主要晚期启动子；L1-3，为腺病毒三联体前导区；SS，为拼接信号；pA，为多聚腺苷酸位点。

图 2 表示质粒 pBJ3 的构建过程。使用的记号 TPIp，为 TPI1 启动子；TPIt，为 TPI1 终止子；AAT，为 α -1 抗胰蛋白酶 cDNA： α ，为 α -1 因子前导区；mTPO，为鼠 TPO 编码序列。

发明的详细说明

在详细描述本发明之前，对本文使用的某些术语下定义将是有帮助的。

等位变异体：另外一种形式的经突变而产生的基因，或突变基因编码的改变的多肽。基因突变可以是沉默的（编码的多肽没有改变）或者可以编码具有改变的氨基酸序列的多肽。

cDNA：互补 DNA，由信使 RNA 模板通过逆（反）转录而产生，或者是这种分子的克隆或扩增拷贝。互补 DNA 可以是单链的或者双链的。

表达载体：一种 DNA 分子，为线状或环状，包含编码多肽的区段。

该区段与一些可提供其转录的附加区段可操作连接。这些附加区段包括启动子和终止子序列，也可包括一个或多个复制起点，一个或多个选择标记，一个增强子，一个多聚腺苷酸化信号，等等。表达载体一般地来源于质粒或病毒 DNA，或含有这两种因子。术语“可操作连接”是指编排所述区段，以便它们为其预期的目的协同作用，如在启动子区开始转录和从编码区段到终止子进行转录。

基因：编码多肽链的一个染色体 DNA 区段。一个基因包括编码氨基酸的一个或多个区域（在某些情况下，这些区域可被一些非编码的“间隔序列”（“内含子”）间隔开）以及提供编码序列转录的侧翼非编码区。

与之互补的分子：与一个参照序列相比，具有互补的碱基序列和颠倒的方向的多核苷酸分子。例如，序列 5'ATGCACGGG 3' 与序列 5'CCCGTGCAT 3' 互补。

启动子：基因的一部分，RNA 聚合酶与其结合，mRNA 的合成在其上开始。

如上述所指明的，本发明提供了用于生产具有血细胞生成活性的蛋白质的材料和方法。如本文中所使用的，术语“血细胞生成”表示按标准试验方法测定的刺激髓样或淋巴样前体增生和/或分化的能力。参见例如 Metcalf, *Proc. Natl. Acad. Sci. USA* 77:5327-5330, 1980; Metcalf et al., *J. Cell. Physiol.* 116:198-206, 1983; 和 Metcalf et al., *Exp. Hematol.* 15:288-295, 1987. 典型地，在一个试验样品和一个对照样品存在时，对骨髓细胞进行培养。然后根据视觉检查和/或染色对培养物的细胞增生和分化进行计数。一个特别优选的测定法是 Mosman (*J. Immunol. Meth.* 65:55-63, 1983; 在此一并参考) 的 MTT 比色测定法，在下面的例子中将被更详细地说明。

本发明部分基于发现了一种通过 MPL 受体刺激细胞生长的活性。在此发现之前，这种受体 (Souyri et al., *Cell* 63:1137-1147, 1990)，为一种“孤儿”受体，其自然配体尚不清楚。通过随后实施例中详细描述克隆和突变方法，发明者培育了一种细胞系，它依赖于其生存和生长的 MPL 受体一连结途径的刺激作用，且它是能够受体自泌刺激的。研究

发现这些不依赖于白细胞介素-3 (IL-3) 的细胞的条件培养基能支持表达 MPL 受体并且还依赖于 IL-3 的细胞生长。抗体中和实验表明, 这种活性并非由于 IL-3 或 IL-4 的作用, 并且这种活性可被一种可溶形式的 MPL 受体中和。然后从不依赖于 IL-3 的细胞系中制备一个 cDNA 文库。用 DNA 转染小仓鼠肾 (BHK) 细胞, 从所述转染体得到的培养介质用于试验刺激 MPL 依赖细胞增生的能力。分离到一个阳性克隆, 并产生重组的 MPL 配体。研究发现, 重组蛋白刺激广谱髓样和淋巴样前体的增生, 并且尤其是刺激骨髓中祖细胞来源的巨核细胞和嗜中性白细胞的产生。此外, 还发现重组蛋白刺激试验动物血小板的产生。鉴于这性活性 (作用), 该蛋白被指定为血小板生成素 (TPO)。

本发明提供了编码血小板生成素的分离的多聚核苷酸分子。在这点上, 有用的多聚核苷酸分子包括 mRNA, 基因组 DNA, cDNA, 合成的 DNA 和不同来源的片段连接产生的 DNA 分子。对重组 TPO 的生产来说, 在多数表达系统中最好使用缺乏内含子的 DNA 分子。“分离”的意思是指从其自然的遗传背景中移出这些分子。因此, 本发明提供了不含其他通常与其相关的基因的 DNA 分子。这些分子尤其不含外来的或不需要的编码序列, 适用于遗传工程蛋白质生产系统。

编码代表性的小鼠和人的 TPO 蛋白的 cDNA 克隆的序列分别在 SEQ ID No:1 和 SEQ ID No:2 中表示出来, 相应的氨基酸序列分别在 SEQ ID No:2 和 SEQ ID No:19 中表示出来。本领域的技术人员将识别出 SEQ ID Nos:1, 2, 18 和 19 中所示序列以及 SEQ ID Nos:28 和 29 中所示的基因组序列, 这些序列与鼠或人的基因的单个等位基因相对应, 并预期有等位变异存在。SEQ ID No:1, SEQ ID No:18 和 SEQ ID No:28 中所示的 DNA 序列的等位变异体 (包括含沉默突变的那些以及那些导致氨基酸序列变化的突变的那些) 都在本发明的范围之内, SEQ ID No:2 和 SEQ ID No:19 的等位变异体的蛋白质也一样。也很显然的是, 本领域技术人员可以利用可变密码子创造有利于核酸序列操作的一些位点。

本文中公开的小鼠和人的序列对从其他动物种类 (“物种同系物”) 制备编码 TPO 蛋白的分离的多聚核苷酸分子是有用的工具。优选的这些物种同系物包括哺乳动物同系物 (如: 牛的、犬的、猪的、绵羊的、马

的、尤其是灵长类的蛋白质)。在本领域中熟知一些用第一物种的序列信息克隆第二物种相应的多核苷酸序列的方法。参见例如, Ausubel et al., eds., Current Protocols in Molecular Biology, John Wiley and Sons, Inc., NY, 1987。本发明的编码 TPO 的 DNA 分子在序列上与 SEQ ID No:1 和 SEQ ID No:18 及其等位变异体一般至少有 60%, 优选地至少有 80%, 和可能有 90 - 95% 或更高的同源性。血小板生长素分子的特征是它们能够与同种来源的 MPL 受体特异性地结合和刺激活体中血小板的产生。在正常的试验动物中, 在开始每天给动物施药的十天内, TPO 可 100% 或更多地增加血小板水平。

对 mRNA 分布的分析表明, 编码 TPO 的 mRNA 存在于人和鼠的几种组织中, 在肺, 肝, 心脏, 骨骼肌和肾脏中, mRNA 的丰度更高。因此, 为了从其他物种中分离同系物, 最好是从已发现可产高水平 mRNA 的一种组织中制备 cDNA 文库。本技术领域已熟知一些制备 cDNA 文库的方法。参见例如, Sambrook et al., eds., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, 1989 及其中引用的参考文献。为了检测编码 TPO 的一些分子, 用本文公开的小鼠或人的 cDNA 或其片段或用一种或多种更小的基于已详细描述序列的探针探测所述文库。一些特别实用的探针包含一个寡聚核苷酸, 其长度至少为 14 个或更多的核苷酸, 并且可多达 25 个或更多的核苷酸, 它们与 SEQ ID No:1, SEQ ID No:18, SEQ ID No:28 中的相同长度的部分或其互补序列至少有 80% 的同源性。最好在一个较低的杂交严谨度下, 即约 $2 \times$ SSC 和约 50 °C 的杂交温度下用标记的探针探测文库。用标准的检测方法检测与探针杂交的分子。通过序列分析和活性检测确认阳性克隆, 如检测与同源的 MPL 受体结合的能力 (即从同种来源的 MPL 受体作为 cDNA) 或刺激同源骨髓细胞的血细胞生成能力。本领域技术人员清楚的是, 可以利用其他的一些克隆方法。

通过从产生 MPL 配体和显示自泌生长刺激作用的细胞系克隆也可分离一些编码 TPO 的多核苷酸分子 (包括等位变异体和本文公开的物种同系物)。简单地说, 对因子依赖性细胞系进行转染, 以表达 MPL 受体 (Vigon et al., Proc. Natl. Acad. Sci. USA 89:5640 - 5644, 1992;

Skoda et al., EMBO J. 12: 2645-2653,1993;和 SEQ ID No:17) , 然后进行诱变处理, 选择一些非因子依赖性细胞。然后这些细胞可用作 TPO mRNA 的来源。一些稳定的因子依赖性细胞系包括 IL - 3 依赖性 BaF3 细胞系 (Palacios and Steinmetz, Cell 41:727 - 734 , 1985 ; Mathey - Prevot et al., Mol. Cell. Biol. 6:4133-4135,1986) ,FDC-P1 (Hapel et al., Blood 64: 786-790, 1984) ,和 MO7e (Kiss et al., Leukemia 7:235 - 240, 1993) 。按照已发表的一些方法 (例如 Greenberger et al., Leukemia Res. 8: 363-375, 1984; Dexter et al., in Baum et al., Eds., Experimental Hematology Today, 8th Ann. Mtg. Int. Soc. Exp. Hematol. 1979, 145-156, 1980) , 可建立生长因子依赖性细胞系。在一个典型的方法中, 从兴趣组织中移出细胞 (如骨髓, 脾脏, 胎儿肝脏) , 并在一个常规的, 补充有血清的培养基中培养, 如补充有 10 % 胎儿牛血清 (FBS) , 15 % 马血清, 和 10^{-6} M 氢化可的松的 RPMI 1640 的培养基中培养间隔一至二周后, 收获非粘着细胞, 并给这些培养物饲喂新鲜培养基。冲洗收获的非粘着细胞, 并在添加有一种生长因子源 (例如, RPMI 1640 + 10 % FBS + 5 - 20 % WEHI - 3 的条件培养基作为 IL - 3 的来源) 的培养基中培养, 每隔一至二周给这些细胞饲喂新鲜培养基, 随着培养物的生长, 这些细胞也扩展开来。几周到几个月以后, 通过把这些细胞涂布于半固体培养基 (如含有甲基纤维素的培养基) 上或通过限制性稀释的方法分离一些单独的克隆。通过在不存在生长因子的条件下培养这些单独的克隆为确认这些克隆的因子依赖性。用逆转录病毒感染或化学诱变可高频率地获得生长因子 - 依赖性细胞。转染因子依赖性细胞, 以表达 MPL 受体, 然后作诱变处理, 如通过化学处理, 暴露于紫外线下, 暴露于 X - 射线下或逆转录病毒插入突变。突变的细胞然后在细胞生存依赖于自泌生长因子产生, 即在母细胞需要的外来生长因子不存在的条件下培养。通过筛选, 如通过检验表达和不表达 MPL 受体的细胞的培养条件或检验在可溶性 MPL 受体或与已知细胞因子结合的抗体存在时的培养条件的活性, 确认 TPO 的产生。

本发明也提供了一些分离的蛋白质, 它们实质上与 SEQ ID No:2 或 SEQ ID No:19 的蛋白质和其物种同系物具有同源性。“分离的”的意思

是指一种蛋白质在一种非原来的自然环境中（如除去了血和动物组织）的条件下发现的。分离的蛋白质一种优选的形式是实质上不含其他蛋白质，尤其是不含其来源动物的其他蛋白质。最好是提供高度纯化形式的一些蛋白质，如纯度大于 95%，纯度最好大于 99%。术语“实质上同源的”此处用来表示这些蛋白质与 SEQ ID No:2 或 SEQ ID No:19 或其物种同系物的氨基酸序列具有 50%，更好为 60%，最好至少具有 80% 的同源性。这些蛋白质与 SEQ ID No:2 或 SEQ ID No:19 或其同系物更好至少具有 90%，最好具有 95% 或更高的同源性。序列同源性的百分率是通过常规的方法测定的。参见，如，Altschul et al., Bull. Math. Bio. 48:603 - 616, 1986 和 Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915 - 10919, 1992。简单地说，用缺口开放障碍（gap opening penalty）10，缺口延伸障碍（gap extension penalty）1，和表 1（用标准单字母码表示氨基酸）所示的 Henikoff 和 Henikoff（同上）“评分矩阵对比两个氨基酸序列，使得序列对比评分最适当。然后按下式计算百分同源性：

相同匹配的总数

较长序列加为对比两个序列引入到较长序列中的缺口数的长度

表 1

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
C	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													
G	0	-2	0	-1	-3	-2	-2	6												
H	-2	0	1	-1	-3	0	0	-2	8											

I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

实质上同源的蛋白质的特征是具有一个或更多氨基酸替换，缺失或增加。这些变化最好只是性质上较微小的，即只是一些保守的氨基酸替换，对蛋白质的折叠或活性并无重大影响（见表2）；小的缺失，典型地缺失掉一到大约三十个氨基酸；小的氨基-或羧基-末端的延伸（如一个氨基末端甲硫氨酸残基，一个达约20-25个残基的小接头肽），或有利于纯化的一个小的（氨基酸）延伸（如一个多一组氨基酸束，一个抗原决定簇或一个结合结构域）。一般性的情况参见，Ford et al., Protein Expression and Purification 2:95-107,1991,该文献本文一并参考。

表 2

保守氨基酸替换

碱性的	精氨酸
	赖氨酸
	组氨酸
酸性的	谷氨酸
	天冬氨酸
极性的	谷氨酰胺
	天冬酰胺
疏水性	亮氨酸
	异亮氨酸

芳香族的

小的

缬氨酸

苯丙氨酸

色氨酸

酪氨酸

甘氨酸

丙氨酸

丝氨酸

苏氨酸

甲硫氨酸

按照本领域中已知的一些方法，如定点突变或丙氨酸一扫描突变（Cunningham and Wells, Science 244, 1081-1085, 1989），可鉴定 TPO 中的一些必需氨基酸。在后面一种技术中，对分子的每个残基引入单个丙氨酸突变，并检测产生的突变体分子的生物活性（如，受体结合，体外和体内增生活性），以鉴定对分子活性起关键作用的氨基酸残基。也可以通过诸如核磁共振，晶体学或光亲和性标记等技术测定的晶体结构的分析，测定配体—受体相互作用的一些位点。参见，例如，de Vos et al., Science 255:306 - 312, 1992；Smith et al., FEBS Lett. 309 59-64, 1992。

一般地，细胞因子预测有一个 4 - α 螺旋结构，第一个和第四个螺旋在配体—受体相互作用中具有极为重要的作用，在其家族成员间具有更加高度的保守性。查看 SEQ ID No:19 中所示的人体 TPO 氨基酸序列，细胞因子序列的排列表明这些螺旋分别由氨基酸残基 29 和 53，80 和 89，108 和 130 及 144 和 168 形成分界面（分界面为 \pm 个残基）。通过与人的序列对比，可测定鼠及其他的非人的 TPO_s 的螺旋分界面。TPO 其他重要的结构方面包括 SEQ ID No:2 中的第 51，73，129 和 195 个位置上的半胱氨酸残基（相当于 SEQ ID No:19 中的 28，50，106 和 172 位点）。

此外，本发明的蛋白质（或其多肽片段）可连接到其他的生物活性分子上（尤其是其他的细胞因子）。提供一些多功能性的分子。例如，血小板生成素的 C - 端结构域可与其他的细胞因子连接，以提高其生物

效能或生产效率。血小板生成素似乎由两个结构域组成。大约 150 个氨基酸的第一个结构域（氨基末端）在大小和结构方面与红细胞生成素和几种其他的血细胞生成激动素相似。在第一个结构域之后是约 180 个氨基酸的第二个结构域，其结构与数据库中已知的任何一种蛋白质没有明显的相似之处。此第二个结构域富含与 N 键结合的糖基化位点和富含丝氨酸，脯氨酸及苏氨酸残基，这些氨基酸残基是与 O 键结合的糖基化位点的标志。这种明显较高含量的碳水化合物表明该结构域在使第一个疏水性的结构域变为相对更为可溶性方面起着某种作用。实验证据表明，与第二个结构域相关的碳水化合物参与了正当的细胞间装配和其生物合成期间蛋白质的分泌。第二个结构域在稳定第一个结构域抗蛋白质水解的降解作用和延长该分子在活体中的半衰期方面也可能起着某种作用，并且可能增强生物信号传递或蛋白质的特异活性。

本发明因而提供了一系列新型的杂种分子，其中 TPO 的第二个结构域与第二个细胞因子相连。TPO 的 C-端结构域与该第二种细胞因子的 C-端相连是优选的。连接最好是在 DNA 水平上通过拼接完成，以便让这些嵌合分子在重组生产系统中表达。然后对所形成的分子进行检测，如提高的可溶性，增加的稳定性，延长的清除半衰期的，或提高的表达和分泌水平，和药效学等方面的检测。这些嵌合细胞因子的具体例子包括其中 TPO 的第二结构域与 EPO，G-CSF，GM-CSF，IL-6，IL-3 或 IL-11 的 C 端相结合那些。如上面所指出的，这可通过 DNA 融合方便地完成。然后，将融合的 cDNA 按照常规的方法亚克隆到一个合适的表达载体中，转化或转染到寄生细胞或有机体中。用常规的色谱纯化技术（如色谱技术）纯化产生的融合蛋白，比较它们与那些天然的、非融合的母亲细胞因子的一些性质。这些杂合分子在其组成蛋白或多肽之间可以还包括一些附加的氨基酸残基（如多肽接头）。

除以上所述的血细胞生成蛋白外，本发明还包括这些蛋白质的一些片段和编码这些片段的分离的多核苷酸分子。是与 MPL 受体结合的长度至少为 10 个氨基酸的片段，和编码这些多肽的长度至少为 30 个核苷酸的多核苷酸分子。通过已知的筛选方法，如通过消化完整的蛋白质或合成较小的重叠多肽或多核苷酸（和表达后者），选择性地与以上描述的

结构分析技术相结合的方法鉴定这种类型的多肽。然后测定这些产生的多肽特异地结合 MPL 受体和通过 MPL 受体刺激细胞增生的能力。通过常规的方法测定结合能力,如 Klotz 公开的方法 (Science 217:1247, 1982 (“ Scatchard analysis ”))。简单地讲,在增加浓度的未标记的 TPO 存在下,把放射性标记的试验多肽与带有 MPL 受体的细胞一起培养。用通过邻苯二甲酸酯油状物的离心方法把与细胞结合的,标记的多肽与未标记的多肽分离开来。根据纵座称上结合未标记物对横座标上结合标记物的比率作图,测定试验多肽的结合亲和力。通过与细胞因子(而不是 TPO)的竞争测定结合特异性。通过由 MPL 受体(或其结合配体的细胞外结构域)沉淀试验化合物也可测定受体结合。简单地讲,将受体或它的一部分固定于一种不溶性的支持物上。例如,就一个重组的试验化合物而言,通过代谢标记寄生细胞或通过常规的体外标记方法(例如放射性碘化作用),对试验化合物进行标记。标记的化合物再与固定化受体相结合,去掉未结合的物质,检测结合的标记化合物。本领域中已知检测种种标记物的方法。用带有 MPL 受体的细胞用 MTT 比色试验法常规测定细胞增生的刺激作用。在各种浓度下,典型地在一个从 1nm 至 1nm 的范围内测定多肽的活性。

(本发明)也提供了多达 50 个或更多个残基、优选地 100 个或更多个残基,更优选地约 140 个或更多个残基以及多达整个成熟蛋白质大小的多肽。例如,对 SEQ ID No:2 中所示的从第 51 个至第 195 个残基(这两个氨基酸残基也包括在内)的氨基酸序列和 SEQ ID No:19 中所示的从第 28 个至第 172 个残基(这两个氨基酸残基也包括在内)的氨基酸序列的分析和模型建立表明所述分子的这些部分是类似于细胞因子一样的可自我装配的结构域。含有这种核心的类似于细胞因子的结构域加上最初翻译产物的一个或更多个区段或结构域的这些分子也是重要的。因此,重要的其他多肽包括表 3 中所示的那些多肽分子。

表 3

小鼠 TPO (SEQ ID NO:2) :

Cys (残基 51) -- Val (残基 196)

Cys (51) -- Pro (206)

Cys (51) - - Thr (379)
Ser (45) - - Cys (195)
Ser (45) - - Val (196)
Ser (45) - - Pro (206)
Ser (45) - - Thr (379)
Met (24) - - Cys (195)
Met (24) - - Val (196)
Met (24) - - Pro (206)
Met (24) - - Thr (379)
Met (1) - - Cys (195)
Met (1) - - Val (196)
Met (1) - - Pro (206)
Met (1) - - Thr (379)

人 TPO (SEQ ID NO:19)

Cys (28) - - Val (173)
Cys (28) - - Arg (175)
Cys (28) - - Gly (353)
Ser (22) - - Cys (172)
Ser (22) - - Val (173)
Ser (22) - - Arg (175)
Ser (22) - - Gly (353)
Met (1) - - Cys (172)
Met (1) - - Val (173)
Met (1) - - Arg (175)
Met (1) - - Gly (353)

本领域技术人员将认识到这些分子的中间形式 (如具有 SEQ ID No:2 中的第 196 和 206 残基之间的 C - 末端的或具有 SEQ ID No:19 中的第 22 和 28 个残基之间的 N - 末端的那些分子) 也是重要的, 以上所述的具有一个或更多氨基酸替换, 缺失, 插入, 或 N - 端或 C - 端延伸的那些多肽同样如此。因此, 本发明提供的血细

胞生成多肽长度至少为 10 个氨基酸残基，较好的是至少为 50 个残基，更好的是至少为 100 个残基，最好的是至少为约 140 个残基，其中所述多肽与 SEQ ID No:2 或 SEQ ID No:19 中的类似大小的多肽实质上是同源的。

本发明的蛋白质可在遗传工程的寄生细胞中通过常规的技术产生。稳定的寄生细胞是那些可以被外源 DNA 转化或转染及能人工培养的细胞类型，包括细菌，真菌细胞和培养的高等真核细胞。操作克隆的 DNA 分子和向各种寄生细胞导入外源 DNA 的技术由 Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, 和 Ausubel et al., 同上公开，这些文献本文一并参考。

一般来说，编码本发明的蛋白质的 DNA 序列在操作上可与表达载体内的转录启动子和终止子相连接。载体一般包含一个或多个选择标记和一个或多个复制起点，尽管本领域的技术人员将认识到在某些系统内，选择标记可以提供于各个单独的载体上，并且通过整合到寄生细胞基因组中可以提供外源 DNA 的复制。启动子、终止子、选择标记、载体和其他因子的选择是日常设计的事情，都在本领域一般技术人员的水平之内。许多这种因子在文献中已有描述，并可通过商业供应商获得。

为了把本发明的蛋白质导入到一些寄生细胞的分泌途径中，在表达载体中提供了一种分泌信号序列（也称前导序列，前置序列或前序列）。分泌信号序列以正确的阅读框架与编码本发明的蛋白质的 DNA 序列相连。分泌信号序列一般地位于编码兴趣蛋白质的 DNA 序列的 5' 端，虽然某些信号序列可以位于兴趣 DNA 序列的其他地方（参见，例如，Welch et al., 美国专利 5,037,743; Holland et al., 美国专利 5,143,830）。分泌信号序列可能是与本发明的蛋白质通常相有关的序列，或可从编码其他分泌蛋白的基因而来。

酵母细胞（尤其是糖酵母属的细胞）是本发明内使用的一种优选寄生。用外源 DNA 转化酵母细胞及从中产生重组蛋白的一些方

法已有描述，例如， Kawasaki, U.S. Patent No. 4, 599,311; Kawasaki et al., 美国专利 4,931,373; Brake, 美国专利 4,870,008; Welch et al., 美国专利 5,037, 743;and Murray et al., 美国专利 4,845, 075，这些文献本文一并参考。根据选择性标记（一般为药物抗性或在缺乏某种特定营养因子（如亮氨酸）的条件下生长的能力）测定的表现型选择转化细胞。在酵母中，一个优选使用的载体系统是 Kawasaki 等（美国专利 4,931,373）详细描述的 POT₁ 载体系统，它允许转化细胞在含葡萄糖的培养基中选择性生长。酵母中使用的一种优选分泌信号序列是 *S.Cerevisiae* MF_{α1} 基因序列（ Brake, *ibid*; Kurjan et al., 美国专利 4,546,082）。适合于酵母中使用的一些启动子和终止子包括那些来源于酵解酶的基因的那些（参见，例如， Kawasaki, 美国专利 4,599,311; Kingsman et al., 美国专利 4,615, 974; and Bitter, 美国专利 4,977,092, 本文一并参考）和酒精脱氢酶基因的那些。也参见美国专利 4,990,446; 5,063, 154; 5,139, 936 and 4,661, 454，这些文献本文一并参考。供其他酵母的转化系统包括本领域中已知的多形汉逊氏酵母、粟酒裂殖糖酵母、乳克鲁维氏酵母、脆壁克鲁氏酵母、玉蜀黍黑粉菌、巴斯德毕赤氏酵母、季也蒙氏毕赤氏酵母和麦芽糖假丝酵母。参看，例如， Gleeson et al., J. Gen. Microbiol. 132:3459-3465, 1986 和 Cregg, 美国专利 4,882,279。

其他的真菌细胞也适合作为寄生细胞。例如，可按照 Mcknight 等美国专利 4,935,349（本文一并参考）公开的方法使用。曲霉属细胞。转化 Acremonium Chrysogenum 的方法已由 Sumino 等，美国专利 5,162,228 公开。

培养的哺乳动物细胞也是本发明内的优选寄生。外源 DNA 导入哺乳动物寄生细胞的方法包括磷酸钙一介导的转染（ Wigler et al., Cell 14:725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7:603, 1981; Graham and Van der Eb, Virology 52:456, 1973), 电击穿（ Neumann et al., EMBO J. 1:841 - 845, 1982 ）和 DEAE - 葡萄糖介导的转染（ Ausubel et al., eds.,

Current Protocols in Molecular Biology, John Wiley and Sons, Inc., NY, 1987), 这些方法本文一并参考, 培养的哺乳动物细胞中的重组蛋白质的产生已被公开, 例如, 被 Levinson et al., 美国专利 4,713,339; Hagen et al., 美国专利 4,784,950; Palmiter et al., 美国专利 4,656,134 (本文在此一并参考) 公开。优选培养的哺乳动物细胞包括 COS-1 (ATCC No. CRL1650), COS-7 (ATCC No. CRL 1651), BHK (ATCC No. CRL 1632), BHK 570 (ATCC N. CRL 10314), 293 (ATCC No. CRL1573; Graham et al., J. Gen. Virol. 36: 59-72, 1977) 和中国仓鼠卵巢 (例如 CHO-K1; ATCC No. CCL61) 细胞系。另外合适的细胞系在本领域中已公知, 并可从一些公共菌种保藏中心 (如 American Type Culture Collection, Rockville, Maryland) 得到。一般来说, 较强的转录启动子是优选的, 如 SV-40 或腺病毒的启动子。参见, 例如, 美国专利 4,956,288。其他合适的启动子包括那些金属硫蛋白基因来源的启动子 (美国专利 4,579,821 和 4,601,978, 在此一并参考) 和腺病毒主要晚期启动子。

药物抗性一般用来选择已插入外源 DNA 的人工培养的哺乳动物细胞。这些细胞一般被称为“转染子”。在选择性因子存在下能被培养且能把兴趣基因传递给其后代的细胞被称之为“稳定的转染子”。一个优选的选择标志是编码抗新霉素抗生素的基因。在一种新霉素类药物, (如 G-418 或其类似药物) 存在的情况下进行选择。选择系统也可被用来增加兴趣基因的表达水平 (被称之为“扩增”的过程)。扩增的实施是通过在低水平的选择因子存在下培养转染子, 然后增加选择因子的剂量以选择那些产生高水平的导入基因产物的细胞。一种优选的扩增选择标记是二氢叶酸还原酶, 它可赋予氨甲蝶呤抗性。也可用其他的药物抗性基因 (例如, 潮霉素抗性, 多药物抗性, 嘌呤霉素乙酰转移酶)。

其他的高等真核细胞也可被用作寄主细胞, 这些细胞包括昆虫细胞, 植物细胞和鸟类细胞。昆虫细胞的转化和其中外源蛋白的产生已被公开 (Guarino et al., 美国专利 5,162,222; Bang et al., U.S.

美国专利 4,775,624;和 WIPO 出版物 WO 94/06463, 本文一并参考)。使用发根病土壤杆菌作为植物细胞基因表达的一种载体已由 Sinkar 等评述 (Sinkar et al., J.Biosci.(Bangalore) 11:47-58,1987)。

在实施本发明中,尽管芽孢杆菌属和其他一些属的细菌也有用处,但优选使用的原核寄主细胞为细菌大肠杆菌的一些菌株。在本领域中已熟知转化这些寄主和表达在其中已克隆的外源 DNA 序列的一些方法(参看,例,Sambrook et al.,同上)。当在象大肠杆菌细菌中表达蛋白质时,蛋白质一般作为不溶性颗粒保持在细胞质中,或通过一种细菌分泌序列指导蛋白质分泌至外周胞质空间中。在前一种情形下,裂解细胞,回收这些颗粒并使用例如异硫氰酸胍进行变性。通过稀释变性剂,然后使变性的蛋白质重新折叠。在后一种情形下,通过破碎细胞(如通过超声波或渗透压震扰)以释放外周胞质的内含物和回收蛋白质的方法,可从外周胞质中回收可溶性和功能形式的蛋白质。

按照常规方法,在含有营养素和选择的寄主细胞生长所必需的其他成分的培养基中培养转化或转染的寄主细胞。本领域中已公知一些合适的培养基,包括限定培养基和复合培养基,一般包括碳源,氮源,必需氨基酸,维生素和矿物质。若需要的话,培养基也含有一些象生长因子或血清一类的成分。例如,含有外源添加的 DNA 的细胞的生长培养基一般通过药物选择或必需营养物缺乏来选择,该营养物缺乏可通过表达载体上携带的或共转染到寄主细胞的选择标记补充。

在本发明中,可使用转基因动物技术生产 TPO。最好产生雌性哺乳动物寄主乳腺内的一些蛋白质。在乳腺中的表达和兴趣蛋白随后向奶中的分泌可克服在从其他来源分离蛋白质过程中遇到的许多困难。奶易于收集,可大量获得,并且可很好地作生化特性分析。而且,主要的奶蛋白高浓度地存在于奶中(约 1 至 15g/L)

从商业观点来看,很显然最好是使用具有高产量奶的寄主物种。尽管可使用一些象小鼠和大鼠一类的小动物(和最好是在

proof-of-concept 阶段)，但在本发明内，最优选使用的牲畜动物包括但不限于猪、山羊、绵羊和牛。绵羊为特别优选的动物，这是由于这种动物已有转基因的先前历史，奶的产量，和很容易得到收集羊奶的设备等因素。为比较影响寄生物种选择的一些因素。可参见 WIPO 出版物 WO 88/00239。一般较理想的是选择一种为生产奶而饲养的寄主动物，如 East Friesland 山羊，或在较晚的时期通过繁育转基因品系采用供奶家畜。无论如何，应该使用已知的健康状况良好的动物。

为了在乳腺中表达，可使用来源于奶蛋白基因的转录启动子。奶蛋白基因包括那些编码酪蛋白（参见美国专利 5,304,489，本文一并参考）、 β -乳球蛋白、 α -乳白蛋白和乳清酸性蛋白的基因。就绵羊的 β -乳球蛋白基因来说，一般使用至少近 406bp 的 5'侧翼序列的区域（尽管优选使用 5'侧翼序列的多达约 5Kbp 的较大部分），如包含 β -乳球蛋白基因的 5'侧翼启动子和非编码部分的一个 4.25Kbp DNA 区段。参看 Whitelew et al., Biochem J. **286:31-39**,1992。源于其他动物种类启动子 DNA 的类似片段也是合适的。

如同欲实现表达的基因的基因组区域一样， β -乳球蛋白基因的其它区域也可结合于构建载体中。本领域一般认为，缺少内含子的构建体当与那些含有这些 DNA 序列的构建体相比较时，其表达水平很低（参见 Brinster et al., Proc. Natl. Acad. Sci. USA **85:836-840**, 1988; Palmiter et al., Proc. Natl. Acad. Sci. USA **88:478-482**,1991; Whitelaw, et al., Transgenic Res. **1:3-13**,1991; WO 89/01343; WO 91/02318）。在这一点上，可能的情况下一般优选使用含编码兴趣蛋白或多肽的基因的所有或某些天然内含子的基因组序列，因此，优选的是进一步包括如 β -乳球蛋白基因的至少某些内含子。一种这样的区域是一个 DNA 片段，该 DNA 片段从绵羊的 β -乳球蛋白基因的 3'非编码区提供内含子的拼接和 RNA 聚腺酸化。当用一个基因天然的 3'非编码序列代替时，该绵羊的 β -乳球蛋白片段可以增加和稳定所述兴趣蛋白或多肽的表达水平。在其他的实施方案中，TPO 序列起始 ATG 周围的区域被

一个特定奶蛋白基因的相应序列替换。这种替换提供了一种推定的组织特异性的起始环境，以增强表达。虽然可以替换较小的区域，但较便利的是用那些如 BLG 基因的一些序列代替完整的 TPO 前 (pre - pro) 和 5' 非编码序列。

为了在转基因动物中表达 TPO，将编码 TPO 的 DNA 区段操作连接到其表达产生表达单位所必需的附加 DNA 区段上。这些附加片段包括上述提到的启动子以及提供转录终止和 mRNA 聚腺苷酸化的序列。表达单位还包括编码分泌信号序列的 DNA 片段，该分泌信号序列可操作地与编码 TPO 的片段相连结。分泌信号序列可以是天然的 TPO 分泌信号序列或者是另一蛋白（如奶蛋白）的分泌信号序列。参见，例如，Von Heinje, Nuc. Acids.Res. 14:4683-4690, 1986; and Meade et al., U.S. Patent No. 4,873, 316 (本文一并参考)。

供转基因动物中使用的表达单位的构建可很方便地通过向质粒或含附加 DNA 区段的噬菌体载体中插入 TPO 序列得以完成（尽管本质上通过任何连接的序列可以构建表达单位）。尤为方便的是提供一种含编码奶蛋白的 DNA 区段的载体和用 TPO 多肽的编码序列代替此奶蛋白的编码序列，由此创造包括此奶蛋白基因的表达控制序列的基因融合。无论如何，在质粒或其他的载体中克隆表达单位使 TPO 序列的扩增很便利。在细菌（如大肠杆菌）寄主细胞中可方便地实施扩增，因此，载体一般地包括在细菌寄主细胞中发挥功能的一个复制起点和一个选择标记。

然后将表达单位导入选择的寄主动物种的受精卵（包括早期胚胎）中。异源 DNA 的导入由几种路径之一完成，包括显微注射（例，美国专利 4,873,191），逆转录感染（Jaenisch, Science 240:1468-1474, 1988）或用胚胎干（ES）细胞的点指导整合（reviewed by Bradley et al., Bio/Technology 10:534-539, 1992）。然后将这些受精卵种植于假孕性动物的输卵管或子宫中，并让其发育至分娩期。在其种系中携带导入 DNA 的后代可按正常的孟德尔规律把 DNA 传递给子代，并让转基因兽群发育。

产生转基因动物的一般步骤在本领域中是已知的。参见,例如, Hogan et al., Manipulating the mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory, 1986; Simons et al., Bio/technology 6: 179-183, 1988; Wall et al., Biol. Reprod. 32:645-651, 1985; Buhler et al., Bio/Technology 8:140-143, 1990; Ebert et al., Bio/Technology 9:835-838, 1991; Krimpenfort et al., Bio/Technology 9:844-847, 1991; Wall et al., J. Cell. Biochem. 49:113-120, 1992; 美国专利 4,873,191 and 4,873,316; WIPO 出版物 WO 88/00239, WO 90/05188, WO 92/11757; and GB 87/00458, 这些文献本文一并参考。向哺乳动物或其种系中导入外源 DNA 序列的技术最早是在小鼠上建立的。参见,例如, Gordon and Ruddle, Science 214:1244-1246, 1981; Palmiter and Brinster, Cell 41:343-345, 1985; Brinster et al., Proc. Natl. Acad. Sci. USA 82:4438-4442, 1985; 和 Hogan et al. (同上)。后来这些技术被用于一些较大的动物,包括家畜物种(参看,例 WIPO 出版物 WO, Bio/Technology 6:179-183,1988)。总之,在至今使用的最有效的产生转基因小鼠或家畜的途径中,按照本领域中标准的技术,已有几百种兴趣线性 DNA 分子被注射到受精卵的前核酸中。DNA 也可以注射到合子中。

也可利用转基因植物生产。表达可在整个植物或特定的器官(如块茎)中进行。参见, Hiatt, Nature 344:469-479, 1990; Edelbaum et al., J. Interferon Res. 12:449-453, 1992; Sijmons et al., Bio/Technology 8:217-221, 1990; 和欧洲专利局出版物 EP 255, 378。

按照本发明的方法制备的 TPO 一般用本领域公知的方法进行提纯,如根据蛋白质的大小、电荷和可溶性和其他性质的亲和纯化和分离。当在培养的哺乳动物细胞中生产时,优选在无血清培养其中培养,以限制污染蛋白质的量。收获培养基并对其分级分离,优选对分级分离的方法包括伴刀豆蛋白 A 和其他凝聚素上的亲和层析,从而利用存在于蛋白质上的碳水化合物。利用固定化的 MPL

受体蛋白或其配体结合部分或利用亲和标记（如多聚组氨酸，物质 P 或可获得其抗体或其他特异性结合因子的其他多肽或蛋白质），也可纯化蛋白质。在兴趣蛋白质和亲和标记之间可提供一个特异的切割位点。

本发明的蛋白质可用于治疗学上，无论用于什么地方，它们对增加骨髓细胞的增生都是理想的，如在治疗白细胞减少症（如由再生障碍性贫血、髓样发育不全综合症、化学疗法或先天性白细胞减少引起的）中；骨髓移植病人中；周围血干细胞移植病人中；和在治疗引起骨髓衰竭疾病（如骨髓发育不良综合症）。这些蛋白质对增加血小板的产生也是有用的，如在治疗血小板减少症中。血小板减少症与多种多样的疾病和临床症状有关，这些疾病或临床症状可单独或协同作用产生这种状况（即血小板减少）。例如，较低的血小板计数可能由例如血小板产生缺陷（由于例如，先天失调，如 Fanconi 综合症、血小板减少综合症、Wiskott Aldrich、May Hegglin 异常、Bernard-Soulier 综合症、Menneapolis 综合症、Epstein 综合症、Montreal 血小板综合症和 Eckstein 综合症）、血小板异常分布、因大量输血引起的稀释损失、血小板的异常破坏或脾机能亢进病人的脾脏中血小板的异常分离（sequestration）（由于例如，肝硬变或充血性的心脏衰竭）引起。例如，癌症治疗中使用的化疗药物可抑制骨髓中血小板先祖细胞的发育，由此产生的血小板减少限制了化学治疗且必须为病人输血。此外，某些恶性肿瘤可损害血小板的产生和血小板分布。用于杀死恶性细胞的辅射疗法也可杀死血小板先祖细胞。血小板减少症也可由各种血小板自身免疫失调（药物、新生的异源免疫、血小板输血异源免疫和病毒（包括 HIV）感染引起的）引起。本发明的蛋白质可减少或排除对输血的需要。并由此减少血小板异源免疫的发生。血小板的异常破坏可因下列原因造成：（1）血管移植或创伤组织的增加的血小板消耗；或（2）与例如药物诱导的血小板减少、自发性血小板减少的紫癜病（ITP）、自身免疫疾病、血液失调症（例如与骨髓有关的白血病、淋巴细胞癌和转移性癌症）。本发明的蛋白质的其他

适应症包括再生障碍性贫血和药物引起的骨髓抑制（例如，因化疗或用 AZT 治疗 HIV 感染而造成的）。

血小板减少症表现为增加出血，如鼻一口区域或肠胃道的出血或伤口、溃疡或注射位点的血的渗出。

为了药学用途，除本发明的蛋白质按常规的方法配制成供肠胃外的（尤其是静脉或皮下）释放的制剂。静脉施药典型地以一至几个小时的期限通过药团（bolus）注射或灌输。总之，药物配方将包括与药理学上可接受的载体（如盐溶液，缓冲盐溶液，溶于水的 5% 的葡聚糖或类似物）相结合的血细胞生成蛋白。配方可另外包括一种或几种赋形剂、防腐剂、增溶剂、缓冲剂、白蛋白等等，以防止小药瓶表面上蛋白质的损失。此外，本发明的血细胞生成蛋白也可与其他的细胞因子组合，尤其是早期作用的细胞因子，如干细胞因子、IL - 3、IL - 6、IL - 11 或 GM - CSF。当利用这样一种组合疗法时，细胞因子可以在一种单独的配方中组合，或在不同的配方中施用。配方的配制方法在本领域中是公知的，并且在例如 Remington's Pharmaceutical Sciences, Gennaro, ed., Mack Publishing Co., Easton PA, 1990（本文一并参考）中公开。治疗剂量一般在每天每 kg 病人体重 0.1 - 100 μ g 的范围内，优选剂量为每天每公斤 0.5 - 20 μ g，准确的剂量由临床医师按接受的标准，考虑到治疗病情的性质和严重程度，病人的特点等等确定。剂量的确定在本领域一般技术人员的水平范围之内。所述蛋白质一般在化疗或骨髓移植后达 28 天的时间期限内或直到血小板数量大于 20,000/ m^3 、优选大于 50,000/ m^3 时施用。更普遍地是所述蛋白质在一周或更短（常常在一到三天期限内）施用。总之，TPO 治疗有效量为足以在临床上显著增加淋巴样或髓样先祖细胞的增生和（或）分化的量，其表现为成熟细胞（如血小板或嗜中性白细胞）的循环水平的增加。由此，将血小板失调的治疗继续至血小板数量至少达到 20,000/ m^3 ，最好达到 50,000/ m^3 。本发明的蛋白质也可与其他的细胞因子（如 IL - 3、- 6 和 - 11；干细胞因子；红细胞生成素；G - CSF 和 GM - CSF）结合施药。在结合疗法的方

式内，其他细胞因子的日剂量一般为：EPO， $\leq 150\text{u/kg}$ ；GM-CSF， $5 - 15\mu\text{s/kg}$ ；IL-3， $1 - 5\mu\text{g/kg}$ ；和G-CSF， $1 - 25\mu\text{g/kg}$ 。例如，具有低水平EPO的贫血患者需要用EPO的结合疗法。

本发明的蛋白质也是体外研究造血细胞的分化和发育（如阐明细胞分化的机制和测定成熟细胞的谱系）的宝贵工具；也可以找到在细胞培养中作为增生因子的用途。

本发明的蛋白质也可体外使用。如自体骨髓培养中。简单地说，在化疗之前从患者体中取出骨髓，用TPO处理，可与可不与一种或更多的其他细胞因子结合处理。处理的骨髓在化疗之后再放回患者体中，以加快骨髓的恢复。此外，本发明的蛋白质也可用于骨髓或外周血先祖细胞（PBPC）的体外扩增研究。在化疗处理之前，用于细胞因子（SCF）或G-CSF刺激骨髓，让早期先祖细胞释放至外周循环中。从外周血液中收集和浓缩这些先祖细胞，然后用TPO培养处理，可与可不与一种或更多的其他细胞因子（包括但不限于SCF、G-CSF、IL-3、GM-CSF、IL-6或IL-11）结合处理，分化和增生至高密度的巨核细胞培养物，在高剂量的化疗之后再把它放回患者体内。

本发明也提供了结合蛋白质上抗原决定簇的一些抗体。这些抗体可通过本领域中已知的各种方法产生。非人体的单克隆抗体的产生方法是熟知的，例如，通过用重组或合成的TPO或其选择的多肽片段免疫动物（如小鼠、大鼠、兔子、山羊、绵羊或豚鼠）来完成。优选用一种高度纯化的蛋白质或多肽片段免疫动物。为增强免疫反应，用佐剂（如福氏佐剂）与蛋白质或多肽结合免疫动物也是优选的。虽然一次注射抗原足以诱导动物体内抗体的产生，但一般优选开始大量注射一次后，紧接着在几周到几个月期间通过一次或多次加强注射免疫动物。参见，例如 Hurrell, ed., Monoclonal Hybridoma Antibodies: Techniques and Applications, CRC Press Inc., Boca Raton, FL, 1982（在此一并参考）。从动物体中取血，使血液凝结成块后，用常规的技术（如盐析、离子交换层析、亲和

层析或高效液相色谱) 从血清中分离抗体。

单克隆抗体的使用一般优于多克隆抗体。单克隆抗体具有易于生产, 特异性和可再生性的优点。本领域中产生单克隆抗体的方法已熟知并已被公开 (例如 Kohler and Milstein, Nature 256:495, 1975. and Eur.J.Immunol. 6:511-519, 1976)。也参见 Hurrell, 同上, 和 Hart, 美国专利 5,094,941 (本文一并参考)。简言之, 使从免疫动物体内获得的产生单克隆抗体的细胞长期不断地维持增殖状态并筛选, 或为了与 TPO 结合的抗体的产生首先进行筛选。通过用骨髓瘤细胞融合使阳性细胞存活。非人的抗体可以按已知的技术“人抗体化”。参见, 例如美国专利 4,816,397; 欧洲专利局出版物 173,494 和 239,400; 和 WIPO 出版物 WO 87/02671 和 WO 90/00616, 这些文献本文在此一并参考。简单地说, 人的恒定区基因与合适的人的或非人的可变区基因相连。例如, 代表母体 (非人的) 单克隆抗体的抗原结合位点 (CDR₁ 或互补性一决定区) 的氨基酸序列与 DNA 水平上可被嫁接到人的可变区框架序列中。本领域中该技术的一些方法已熟知并已被公开 (例如, Jones et al., Nature 326:522-525, 1986; Richmann et al. Nature 322:323-327, 1988 and Queen et al. Proc. Natl. Acad. Sci. USA 86:10029-10033, 1989)。然后用连接的基因转染寄主细胞, 再按常规的方法培养这些寄主细胞。或者, 用克隆的人恒定区基因和同源重组产生的嵌合抗体基因转染单克隆抗体产生细胞。因此, 可能组装成具有人的单克隆抗体重要部分结构的单克隆抗体, 由此提供更适合于用多种施药方法施用给患者的抗体。

通过重组的, 一般由可变区的轻链序列典型地通过接头多肽与可变区的重链序列相连的多肽的表达产生单链抗体。产生单链抗体的方法在本领域中已公知, 并由 Davies 等公开 (Davies et al. Biotechnology 9:165-169, 1991)。

与 TPO 抗原决定簇结合的抗体是有用的, 例如, 在以血小板, 巨核细胞或其他血细胞或先祖细胞的减少为特征的疾病诊断中, 这些疾病与先祖细胞的增生或分化的缺乏症有关。一般用常规的免疫

试验方法（如酶联免疫吸附试验（ELISA）或辅射免疫试验）通过测定血液或血浆实施这种诊断。这些类型的试验方法在本领域中已公知。参见，例如，Hart et al., Biochem. 29:166-172, 1990; Ma et al., British Journal of Hematology 80:431-436, 1992; and Andre et al., Clin. Chem. 38/5:758-763, 1992。TPO活性的诊断试验对鉴定最有可能受益于TPO治疗的患者群体是有用处的。TPO的抗体在TPO的纯化方面也有用处，如通过把抗体结合于固体支持物上（如包装于柱子中的颗粒基质），和用含蛋白质的溶液过程。然后用合适的缓冲液洗脱结合的蛋白。总之，蛋白质在较低离子强度和近中性pH的生理条件下结合于柱子上，然后用洗涤柱子，以洗脱未结合的污染物。结合蛋白的洗脱通过改变离子强度或pH值（如用3M KSCN（分批或梯度）或低pH的柠檬酸缓冲液）进行。一般应避免低于约2.5的pH值。

本发明也提供了一些产生大量的巨核细胞和血小板的方法，这些巨核细胞和血小板可用于制备cDNA文库。由于血小板被导引到一些受伤位点，它们被认为是伤口愈合的媒介，在某些情况下，为致病过程的媒介。因此，对血小板和巨核细胞分子生物学的详细理解将洞察体内平衡和血小板功能的临床上有关的失调。本发明的蛋白质提供了一种改进的产生巨核细胞或血小板cDNA文库的方法。

将重组血小板生成素施用于动物或应用于培养的脾或骨髓细胞时可引起来自前体细胞的巨核细胞的增生。在施用TPO后，巨核细胞和其前体的扩增及巨核细胞成熟使高纯度的和可供mRNA分离及cDNA文库构建的充足数量的巨核细胞的分离成为可能。通过调整TPO的剂量和施药方式，可从原初脾或骨髓细胞选择性地扩大早期或完全成熟的巨核细胞以及那些主动脱去血小板的巨核细胞。相应地，按照巨核细胞生成早期、中期或晚期构建代表性的cDNA文库。

产生的cDNA文库有许多用途。例如，这些文库可用于在各种血小板功能失调中起作用的低丰度蛋白质的鉴定和克隆。通过毫不

费力地对患者的巨核细胞进行扩增和对其 mRNA 进行分离分析，大大地帮助了对这些疾病的分子分析。这些文库也是新型生长因子和具有潜在治疗效用的蛋白质的克隆来源。已克隆的有用的血小板蛋白包括从血小板来源的生长因子 (Ross et al., Cell 26:155-169,1986) ;转化的生长因子 (Miletich et al., Blood 54:1015-1023,1979; Roberts and Sporn, Growth Factors 8:1-9,1993) ;血小板来源的内皮细胞生长因子 (Miletich et al., Blood 54:1015-1023,1979) 和 PF - 4 (Doi et al., Mol. Cell. Biol. 7:898-904, 1987; Poncz et al., Blood 69:219-223,1987) 。新型生长因子可通过表达 cDNA 文库的功能筛选或用已知的生长因子探针在较低的严紧度下杂交筛选进行鉴定。新型生长因子的分离通过利用与已知生长因子保守区域相对应的简并引物经聚合酶链式反应完成。此外，对文库系统和完整的 DNA 序列分析提供了巨核细胞 cDNA 序列的数据库。这种数据库可用于通过各种基于计算机的检索规则系统挖掘一些有用序列。

根据上面描述制备的巨核细胞也可用于制备蛋白质文库。该蛋白质文库与其 cDNA 文库互补。由蛋白质文库获得的氨基酸序列信息使迅速分离编码兴趣蛋白 cDNA 成为可能。使用蛋白质序列数据设计 DNA 分离的引物可排除在常规的文库制备方法中因相对的 mRNA 丰度引起的一些问题。蛋白质和 cDNA 文库的结合也推进了特定兴趣序列的目标克隆。

通过从巨核细胞中按已知的方法提取蛋白质，然后通过双向凝胶电泳分离蛋白质，制备蛋白质文库。分离的蛋白质然后经过原位胰蛋白酶消化，接着通过微孔 HPLC 分离。然后分离的片段通过质谱仪分析。通过与蛋白质序列数据库对比，对形成的质量构成进行检索，推导蛋白质的同源性。未鉴定的肽通过 Edman 降解进行序列分析。

cDNA 和蛋白质文库是一些新的蛋白质及编码这些蛋白质的宝贵来源。血小板被认为是伤口愈合和在某些情况下发病的重要媒介。已对许多重要的血小板蛋白进行了鉴定和特征描述，所述蛋白

包括血小板来源的生长因子，转化生长因子- β ，血小板来源的内皮生长因子和血小板因子4。对其他血小板蛋白的鉴定和特征描述将非常有助于阐明伤口愈合和发病的基本过程，并预期产生一些重要的治疗因子和策略。

正如下面将更加详细描述，人的TPO基因已被定位于染色体3q26上。这一信息与人TPO基因序列（SEQ ID No:28）结合起来，通过遗传筛选，可直接诊断TPO基因中的遗传失调或其表达调节。这些失调包括导致表达水平上的增加或减少的启动子序列的改变，编码或非编码区的染色体易位和TPO位点上的新的调节序列的并置。可应用的诊断方法在本领域中已公知。例如，从基因组序列中可设计至少为5个核苷酸，优选为15-30个或更多核苷酸长度的引物或杂交探针，用于检测染色体畸变或测定mRNA水平。各种合适的检测和测定方法在本领域中已公知，这些方法包括“Southern”印迹，聚合酶链反应（Mullis, 美国专利4,683,202），和连接酶链反应（Barany, PCR Methods and Applications 1:5-16, Cold Spring Harbor Laboratory Press, 1991）。例如，用一种或多种限制性酶消化患者DNA并转移到硝酸纤维素膜上，产生Southern印迹。再用探针探查印迹，检测因限制性位点识别序列突变产生的片段大小的总的变化。在另一种方法中，对异常基因序列的分析及正常和异常序列的比较可设计用于鉴定异常基因（如断裂的或易位的）的引物。通过聚合酶链反应扩增患者DNA，检测表示正常基因特征的或特定基因重排的扩增产物。

本发明通过下面非限制性的实施例作进一步的说明。

实施例 I. 人 MPL 受体 cDNA_s 的分离

利用设计的与已发表的编码受体的氨基和羧基末端序列（Vigon et al., Proc. Natl. Acad. Sci. USA 89: 5640 - 5644）相对应的引物，通过逆转录酶聚合酶链反应（RT - PCR）从人的红细胞白血病细胞（HEL）（Martin and Papayannopoulou, Science 216:1233-1235, 1982）中分离人MPL - P和MPL - K受体同种型编码cDNA。利用引物ZC5499（SEQ ID No:3）从 polyd

(T) - 选择的 poly (A)⁺ RNA 合成模板 HEL 细胞 cDNA 。浓度为 1 μ g/ml 的 13 μ l 的 HEL 细胞 poly (A)⁺ RNA 与 3 μ l 的 20pm/ μ l 第一条链引物 ZC5499 (SEQ ID No:3) 混合, 混合物在 65 $^{\circ}$ C 下加热 4 分钟, 置于冰上冷却。

通过加入 8 μ l 的第一条链缓冲液 (250mM Tris-HCl, pH8.3, 375mMKCl, 15mM MgCl₂) (5 \times SUPERSRIPTTM 缓冲液; GIBCOBRL, Gaithersburg, MD), 4 μ l 的 100mM 二硫苏糖醇和 3 μ l 的一种双脱氧核苷酸三磷酸溶液 (各含 10mM 的 dATP, dGTP, dTTP 和 5 - 甲基 - dCTP) (Pharmacia LKB Biotechnology Inc., Piscataway, NJ) 开始第一条链的合成。反应混合物在 45 $^{\circ}$ C 下温育 4 分钟后, 加入 10 μ l 的 200U/ μ l 的 RNaseH - 逆转录酶 (SUPERSRIPTTM 逆转录酶, GIBCO BRL) 至 RNA - 引物混合物中。反应液再于 45 $^{\circ}$ C 下温育 1 小时, 接着在 50 $^{\circ}$ C 下温育 15 分钟。加入 60 μ l 的 TE (10mM Tris:HCl, pH8.0, 1mM EDTA) 至反应液中, 再在 400 孔径大小的凝胶过滤柱子 (CHROMA SPIN + TE - 400TM; Clontech Laboratories Inc. Palo Alto, CA) 上通过层析去除多余的引物。

利用与编码受体蛋白的氨基和羧基末端区域 (Vigon et al., 同上) 相对应的引物, 用第一条链 HEL 细胞的 cDNA 作为模板, 扩增人的 MPL - P 受体 cDNA 。每个引物中各包含一个不同的限制性酶切割位点, 有助于扩增产物的定向克隆 (ZC5746, SEQ ID No:4, 含有一个 EcoRI 位点; ZC5762, SEQ ID No:5, 含有一个 XhoI 位点) 。准备一种 100 μ l 的反应混合液, 其中含 10mg 的模板 cDNA, 50pM 的每种引物; 200 μ M 的各种脱氧核苷酸三磷酸 (Pharmacia LKB Biotechnology, Inc.) ; 1 μ l 的 10 \times PCR 缓冲液 (Promega Corp. Madison, WI) ; 和 10 个单位的 Taq 聚合酶 (Roche Molecular Systems, Inc., Branchburg, NJ) 。聚合酶链反应进行 35 个循环 (95 $^{\circ}$ C 1 分钟, 60 $^{\circ}$ C 1 分钟和 72 $^{\circ}$ C 2 分钟为一个循环, 在每个连续循环再加 1 秒钟) , 然后于 72 $^{\circ}$ C 下温育 10 分钟。

通过聚合酶链反应扩增从 HEL 细胞 cDNA 中分离人的 MPL - K 受体 cDNA，其分离方式与以上描述的 MPL - P 受体 cDNA 相同，只是用引物 ZC5742 (SEQ ID No:6) 代替引物 ZC5762。PCR 引物 ZC5742 为人的 MPL - K cDNA 的 3' 端所特有，并结合有一个便于克隆的 XhoI 位点。

反应产物用苯酚/氯仿 (1:1) 抽提两次，然后用氯仿再抽提一次和乙醇沉淀。产物用 EcoRI 和 XhoI 消化之后，在 0.8% 的低熔点琼脂糖凝胶 (SEA PLAQUE GTGTM 低熔点琼脂糖； FMC Corp., Rockland, ME) 上分级分离。通过用 β - 琼脂糖酶 (New England Biolabs, Inc., Beverly, MA) 消化凝胶基质，从切下的凝胶带中回收相当于人 MPL - P 受体 cDNA 的一个 1.9kb 的扩增产物和相当于人 MPL - k 受体 cDNA 的一个 1.7kb 的扩增产物，再用乙醇沉淀。将上述分离的 cDNA_s 再亚克隆到 pBluescript[®]SK⁺ (Stratagene Cloning Systems, La Jolla, CA) 载体上，通过序列分析予以证实。

实施例 II. 小鼠 MPL 受体 cDNA 的分离

从 C57BL/K_sJ-db/db 小鼠体内取出脾脏，立即置于液氮中。用异硫氰酸胍 (Chirgwin et al., Bioistry 18:52-94,1979) 处理脾脏组织，接着通过一个 CsCl 梯度离心步骤，从脾脏组织中制备总 RNA。用 Oligo - d (T) 纤维素层析 (Aviv and Leder, Proc. Natl. Acad. Sci. U.S.A 69:1408-1412,1972) 分离脾脏的 poly(A)⁺ RNA。

浓度为 1.7 μ g/ml 的 7.5 μ l 的 poly d (T) - 选择的 poly(A)⁺ 小鼠脾 RNA 与含有一个 Not 限制性位点的 3 μ l 的 20pM/ μ l 第一链引物 ZC6091 (SEQ ID No:7) 混合。混合物于 65 $^{\circ}$ C 下加热 4 分钟，然后置于冰上冷却。通过向 RNA - 引物混合物中加入 8 μ l 的 250mM Tris-HCl, pH8.3, 375mM KCl, 15mM MgCl₂ (5 \times SUPERSRIPTTM 缓冲液； GIBCO BRL)，4 μ l 的 100mM 的二硫苏糖醇和 3 μ l 的各含有 dATP、dGTP:dTTP 和 5 - 甲基 - dCTP (Pharmacia LKB Biotechnology Inc.) 的脱氧核苷酸三磷酸溶液开始第一链 cDNA 的合成。反应混合物于 45 $^{\circ}$ C 下温育 4 分钟，再

加 10 μ l 的 200U/ μ l Rnase H⁻ 逆转录酶 (GIBCO BRL) 。在一个平行反应中, 通过向一个 10 μ l 反应混合物等分样中加入 10 μ Ci 的 ³²p - α dCTP 以标记反应, 分析第一链合成的效率。反应液于 45 $^{\circ}$ C 下温育 1 小时, 再于 50 $^{\circ}$ C 下温育 15 分钟。在一个 400 孔径大小的凝胶过滤柱 (CHROMA SPIN + TE - 400TM; Clontech Laboratories Inc.) 中通过层析去除标记反应中未接合的 ³²p - α dCTP。在 8 μ g 的糖原载体, 2.5M 的醋酸铵和 2.5 倍体积的乙醇存在时通过二次沉淀去除掉未标记的第一链反应中未结合的核苷酸。将未标记的 cDNA 重新悬浮于 50 μ l 的水中, 用于第二链的合成。通过琼脂糖凝胶电泳测定标记的第一链 cDNA 的长度。

在启动的第一链指导第二链合成产生 DNA 发夹结构的条件下, 据第一链进行第二链的合成。在室温下和由 50 μ l 未标记的第一链 cDNA, 16.5 μ l 的水, 20 μ l 的 5 \times 聚合酶 I 缓冲液 (100mM Tris HCl, pH7.4, 500mM KCl, 25mM MgCl₂, 50mM (NH₄)₂SO₄), 1 μ l 的 100mM 的二硫苏糖醇, 2 μ l 的含有各种脱氧核苷酸三磷酸的溶液, 3 μ l 的 5mM β - NAD, 15 μ l 的 3U/ μ l 的大肠杆菌 DNA 连接酶 (New England Biolabs Inc., Beverly, MA) 和 5 μ l 的 10U/ μ l 的大肠杆菌 DNA 聚合酶 I (Amersham Corp., Arlington Heights, IL) 组成的混合物中组装反应混合物。反应液于室温于温育 5 分钟, 再加入 1.5 μ l 的 2U/ μ l Rnase H (GIBCO BRL), 通过向 10 μ l 第二链合成混合物等分样中加入 10 μ Ci 的 ³²p - α dCTP 标记平行反应, 以检测第二链合成的效率。反应液于 15 $^{\circ}$ C 下温育 2 小时, 紧接着在室温下温育 15 分钟。经一个 400 孔径大小的凝胶过滤柱 (Clontech Laboratories, Inc) 通过层析去除标记反应中未结合的 ³²p - α dCTP, 之后再通过琼脂糖凝胶电泳分析。通过用苯酚/氯仿抽提二次和用氯仿抽提一次后, 在 2.5M 醋酸铵存在时用乙醇沉淀来终止未标记的反应。

用绿豆核酸酶切割具发夹结构的单链 DNA。反应混合物含 100 μ l 的第二链 cDNA, 20 μ l 的 10x 的绿豆核酸酶缓冲液 (Stratagene Cloning Systems, La Jolla, CA), 16 μ l 的 100mM

的二硫苏糖醇，51.5 μ l 的水和 12.5 μ l 的以绿豆核酸酶稀释缓冲液以 1:10 稀释的绿豆核酸酶（Promega Corp; 终浓度为 10.5U/ μ l）。反应于 37 $^{\circ}$ C 下温育 15 分钟，通过加入 20 μ l 的 1M Tris-HCl，pH8.0，再通过如上所述的相继用苯酚/氯仿和氯仿抽提的方法终止反应。抽提之后，用乙醇沉淀 DNA 并重新悬浮于水中。

重新悬浮的 cDNA 用 T₄ DNA 聚合酶补平。重新悬浮于 190 μ l 水中的 cDNA 与 50 μ l 5x 的 T₄ DNA 聚合酶缓冲液（250mM Tris-HCl, pH8.0, 250mM KCl, 25mM MgCl₂），3 μ l 0.1 的二硫苏糖醇，3 μ l 的含 10M 的各种脱氧核苷酸三磷酸的溶液和 4 μ l 的 1U/ μ l 的 T₄ DNA 聚合酶（Boehringer Mannheim Corp., Indianapolis, IN）混合。在 10 $^{\circ}$ C 下温育 1 小时后，通过加入 10 μ l 的 0.5M EDTA，接着用与上述描述相同的几次苯酚/氯仿和氯仿抽提终止该反应。通过一个 400 孔径大小的凝胶过滤柱（Clontech Laboratories Inc., Palo Alto, CA）对 DNA 作层析处理，以除去微量蛋白质和不到 400bp 长度的较低的 cDNA。在 12 μ g 糖原和 2.5M 醋酸铵存在时用乙醇沉淀 DNA，再重新悬浮于 10 μ l 的水中。根据渗入的 ³²P- α -dCTP，估计由 12.5 μ g 的起始 mRNA 模板合成的 cDNA 产量为 ~2 μ g。

在 cDNA 的 5' 末端连接 EcoRI 接头，使之能够克隆到入噬菌载体中。10 μ l 等份的 cDNA（约 2 μ g）和 10 μ l 的 65pM/ μ l 的 EcoRI 接头（Pharmacia LKB Biotechnology Inc.）与 2.5 μ l 10x 连接酶缓冲液（Promega Corp.），1 μ l 的 10mM ATP 和 2 μ l 的 15U/ μ l T₄ DNA 连接酶（Promega Corp.）混合，该反应物于 0 $^{\circ}$ C - 18 $^{\circ}$ C 的温度范围内过夜（0 $^{\circ}$ C - 18 $^{\circ}$ C）温育，该反应物再于 12 $^{\circ}$ C 下温育过夜。通过加入 75 μ l 的水和 10 μ l 的 3M 醛酸钠，再于 65 $^{\circ}$ C 下温育 30 分钟终止该反应。温育之后，cDNA 用上述描述的苯酚/氯仿和氯仿抽提，用 2.5M 醛酸钠和 1.2 倍体积的异戊醇沉淀 cDNA。离心之后，cDNA 沉淀物用 70% 乙醇洗涤，空气干燥，重新悬浮于 89 μ l 的水中。

为了便于 cDNA 向入噬菌体的定向克隆，用 NotI 消化 cDNA，

产生 5' 具有 EcoRI 和 3' 具有 NotI 的粘性末端的 cDNA。cDNA 3' 端的 NotI 限制性位点已事先通过引物 ZG6091 (SEQ ID No:7) 引入。限制性酶的消化是在含有以上描述的 89 μ l 的 cDNA, 10 μ l 的 6mM Tris:HCl, 6mM MgCl₂, 150mM NaCl, 1mM DTT (10x 缓冲液 D; Promega Corp., Madison, WI) 和 1 μ l 的 12U/ μ l 的 NotI (Promega Corp.) 的反应物中进行的。消化于 37 $^{\circ}$ C 下进行 1 小时。通过连续的苯酚/氯仿和氯仿抽提终止该反应。cDNA 经乙醇沉淀, 70% 乙醇洗涤, 空气干燥, 重新悬浮于 20 μ l 的 1x 含凝胶的缓冲液中 (10nM Tris:HCl, pH8.0, 1mM DETA, 5% 甘油和 0.125% 的溴酚兰)。

重新悬浮的 cDNA 加热至 65 $^{\circ}$ C 持续 5 分钟, 冰上冷却, 再于 0.8% 低熔点琼脂糖凝胶 (SEA, PLAQUE GTGTM 低熔点琼脂糖; FMC Corp.) 上电泳。从胶上切下来结合的接头和低于 1.6Kb 长的 cDNA。再把电极逆转, 电泳直至 cDNA 浓缩至泳道起点。切下含浓缩 cDNA 的凝胶部分, 置于一个微量离心管中, 测定胶条的大致体积。加入约为凝胶体积的三倍的水等分样, 再把琼脂糖加热至 65 $^{\circ}$ C 持续 15 分钟。当样品平衡至 42 $^{\circ}$ C 时, 加 10 μ l 的 1U/ μ l 的琼脂酶 I (New England Biolabs, Inc.), 混合物温育 90 分钟以消化琼脂糖。温育之后, 加入 40 μ l 的 3M 醋酸钠至样品中, 混合物再于冰上温育 15 分钟。以 14,000g 于室温下离心样品 15 分钟, 去除未消化的琼脂糖。上清液中的 cDNA 经乙醇沉淀, 70% 的乙醇洗涤, 空气干燥后, 重新悬浮于 37 μ l 的水中, 用于激酶反应, 使连接的 EcoRI 接头磷酸化。

向上面描述的 37 μ l cDNA 溶液中加入 10 μ l 10x 连接酶缓冲液 (Stratagene Cloning Systems), 再把反应混合物加热至 65 $^{\circ}$ C, 持续 5 分钟。于冰上冷却该混合物, 加入 5 μ l 的 10mM ATP 和 3 μ l 的 10U/ μ l 的 T₄ 多聚核苷酸激酶 (Stratagene Cloning Systems)。反应于 37 $^{\circ}$ C 下温育 45 分钟, 通过加热至 65 $^{\circ}$ C 持续 10 分钟, 再用一连续的苯酚/氯仿和氯仿抽提终止该反应。磷酸化的 cDNA 在 2.5M 醋酸存在时经乙醇沉淀, 用 70% 乙醇洗涤, 空气干燥和用

12.5 μ l 的水重新悬浮。经估测，磷酸化的 cDNA 浓度约为 40fmol/ μ l。

产生的 cDNA 克隆至购买的预先用 EcoRI 和 NotI 消化和在磷酸化的入噬菌体载体入 ExCellTM (Pharmacia LKB Biotechnology Inc.) 上。cDNA 与载体的连接在包含 2 μ l 的 20fmol/ μ l 制备的入 ExCellTM 噬菌体臂，4 μ l 的水，1 μ l 10x 连接酶缓冲液 (Promega Corp.)，2 μ l 的 40fmol/ μ l cDNA 和 1 μ l 的 15U/ μ l T₄DNA 连接酶反应物中进行，于 4 $^{\circ}$ C 下连接 48 小时。按照销售商提供的说明书，用 GIGAPACK[®] II Gold 包装提取物 (Stratagene Cloning Systems)，将约 50 % 的连接混合物包装至噬菌体中。产生的 cDNA 文库含有超过 1.5×10^7 独立的重组体，具有不到 1.5 % 的非插入性的噬菌体的背景水平。

一个 ³²p - 标记的人 MPL - K 受体 cDNA 探针用来从小鼠脾 cDNA 噬菌体文库中分离小鼠 MPL 受体 cDNA。cDNA 文库涂布于大肠杆菌的 SURE[®] 细胞中 (Stratagene Cloning Systems)，密度为每 150mm 直径的平板 40,000 至 50,000 PFU。三十三个平板的噬菌斑转移至尼龙膜上 (Hybond NTM; Amersham Corp., Arlington Heights, IL) 并按照生产商的说明进行处理。处理的滤膜在真空炉中于 80 $^{\circ}$ C 下烘烤 2 小时，再于洗涤缓冲液 (0.25 \times SSC，0.25% SDS, 1mM EDTA) 中洗涤几次，于杂交溶液 (5 \times SSC，5 \times Denhardt's 溶液，0.1 % SDS，1mM EDTA 和 100 μ g/ml 热变性鲑鱼精 DNA) 中在杂交箱 (model HB - 2; Techne Inc., Princeton, NJ) 中预杂交过夜。预杂交之后，弃去杂交溶液，用含有利用商业上提供的标记试剂盒 (MEGAPRIMETM 试剂盒; Amersham Corp., Arlington Heights, IL 制备的约 2×10^6 cpm/ml 的 ³²p - 标记的人 MPL - K cDNA 的新鲜杂交溶液替换，探针加到杂交溶液之前于 98 $^{\circ}$ C 下变性 5 分钟，于 65 $^{\circ}$ C 下杂交过夜。滤膜用洗涤缓冲液 (0.25 \times SSC，0.25% SDS, 1mM EDTA) 于 55 $^{\circ}$ C 下洗涤，在 - 70 $^{\circ}$ C 下，置于 XAR - 5 胶片 (Kodak Inc., Rochester, NY) 上用增感屏放射自显影 4 天。用放射自显影

图作为模板，从对应于初级信号的平板区域中收获琼脂塞，浸泡于 SM (0.1M NaCl; 50mM Tris:HCl, pH7.5, 0.02%明胶) 中洗提噬菌体，以提纯噬斑。七个经噬菌斑提纯的噬菌体被分离出来，它们携带的插入片段可与人 MPL - K 受体探针杂交。按照厂商提供的说明书。用体内的重组系统收获入 ExCell™ 噬菌体中含有的噬菌粒，通过 DNA 序列分析证实 cDNA 插入的同源性。

分离的克隆编码的蛋白质表现与人 MPL - P 受体和最近报道的小鼠 MPL 受体 (Skoda et al., EMBOJ. 12:2645-2653,1993) 的一种较高级别的序列同源性。七个克隆分为两类，彼此互为不同，三个克隆缺失掉编码近 N - 端的 60 个氨基酸残基的序列，编码没有缺失的蛋白质的 cDNA 称为小鼠类型 I 受体 cDNA。类型 II 受体 cDNA 缺乏编码 SEQ ID No:17 中的第 131 至 190 的氨基酸残基的序列。此外，类型 I 和类型 II 受体不同于报道的小鼠 MPL 受体序列 (Skoda et al., *ibid*)，在于这两类受体中存在一个编码插入于第 222 个氨基酸残基后的氨基酸残基 Val-Arg-Thr-Ser-Pro-Ala-Gly-Glu (SEQ ID No:9) 的序列，和在第 241 个位置上 (位置指类型 I 小鼠受体) 甘氨酸残基被丝氨酸取代。

为了在哺乳动物细胞中表达，把类型 I 和类型 II 小鼠 MPL 受体 cDNA_s 亚克隆至质粒载体 pHZ-1 上。质粒 pHZ-1 为一种表达载体，它可用来在哺乳动物细胞中或蛙卵母细胞翻译系统中由体外转录的 mRNA 表达蛋白质。PHZ - 1 表达单元由小鼠金属硫蛋白 - 1 启动子，噬菌体 T₇ 启动子 (其两侧由含有供插入克隆序列的单一限制性位点的多克隆区)，人生长激素终止子和噬菌体 T₇ 终止子。此外， pHZ-1 还含有一个大肠杆菌的复制起点；一个细菌的 β - 内酰胺酶基因；一个包含 SV40 启动子和起点的哺乳动物可选择标记表达单位、一个新霉素抗性基因和 SV40 转录终止子。为了便于定向克隆至 PHZ-1，使用合适的引物，用聚合酶链反应分别在翻译起始密码子的上游和翻译终止密码子的下游创造一个 EcoRI 位点和一个 XhoI 位点。聚合酶链反应在一种混合物中进行，该混合物含有 10μl 10x 的 ULTMA™ DNA 聚合酶缓冲液 (Roche Molecular

Systems, Inc., Branchburg, NJ), 6 μ l 的 25mM MgCl₂, 0.2 μ l 的含各 10mM 的 dATP、dGTP、dGTP 和 dCTP (Pharmacia LKB Biotechnology Inc.) 的脱氧核糖核酸三磷酸溶液, 2.5 μ l 的 20pmol/ml 引物 ZC6603 (SEQ ID No:8), 2.5 μ l 的 20pmol/ μ l 引物 ZC5762 (SEQ ID No:5), 32.8 μ l 的水, 1 μ l 的含类型 I 或类型 II 小鼠 MPL 受体质粒的较早的对数期的细菌培养物, 1 μ l 的 6U/ μ l DNA 聚合酶 (ULTMATM 聚合酶; Roche Molecular Systems, Inc., Branchburg, NJ.)。在反应中按厂商的说明使用 AmpliwaxTM (Roche Molecular Systems, Inc.) 聚合酶链反应进行 25 个循环 (95 $^{\circ}$ C 1 分钟, 55 $^{\circ}$ C 1 分钟和 72 $^{\circ}$ C 3 分钟), 紧接着于 72 $^{\circ}$ C 下温育 10 分钟。扩增产物用苯酚/氯仿和氯仿连续抽提, 然后在 6 μ g 的糖原载体和 2.5M 醋酸铵存在时用乙醇沉淀。沉淀带重新悬浮于 87 μ l 的水中, 再向水中加入 10 μ l 的 10x 缓冲液 H (Boehringer Mannheim Corp.), 2 μ l 的 10U/ μ l EcoRI (Boehringer Mannheim) 和 1 μ l 的 40U/ μ l XhoI (Boehringer Mannheim Corp.)。消化于 37 $^{\circ}$ C 下进行 1 小时。通过加热至 65 $^{\circ}$ C, 持续 15 分钟终止反应, 并通过 400 孔径大小的凝胶过滤柱子 (CHROMA SPIN + TE - 400TM; Clontech Laboratories Inc.) 进行层析。

按上面描述分离的受体插入片段连接至 EcoRI 和 XhoI 消化和脱磷酸化的 pHZ-1 受体上。连接反应含有 1 μ l 的 50ng/ μ l 的制备的 pHZ-1 载体, 5 μ l 的 5mg/ μ l cDNA 插入片段, 2 μ l 的 10x 连接酶缓冲液 (Promega Corp.), 11.75 μ l 的水和 0.25 μ l 的 4U/ μ l T₄ DNA 连接酶 (Stratagene Cloning Systems)。于 10 $^{\circ}$ C 下过夜进行连接。用连接的 DNA_s 按厂商的说明转染大肠杆菌 (MAX EFFICIENCY DH10BTM 感受态细胞; GIBCO BRL)。通过 DNA 序列分析证实 pHZ-1 中的类型 I 和类型 II 小鼠和人 MPL-P 受体插入片段的正确性。产生的质粒 pSLmpl-8 和 pSLmpl-q 分别携带小鼠类型 II 和类型 I 受体 cDNA_s。质粒 pSLmpl-44 携带人 MPL - P cDNA 插入片段。

实施例 III, 表达 MPL 受体的 BaF₃ 细胞系的构建

将依赖于白细胞介素 - 3 的来源于鼠骨髓的前淋巴样细胞系 BaF₃ (Palacios and Steinmetz, Cell 41:727-734,1985; Mathey-Prevot et al., Mol. Cell.Biol. 6:4133-4135,1986) 保存于完全培养基 CRPMI 1640 培养基 (JRH Bioscience Inc., Lexexa KS) 中, 并补充 10 % 的热钝化小牛胎儿血清, 4 % 的来源于培养的 WEHI - 3 细胞 (Becton Dickinson Labware, Bedford, MA) 的条件培养基, 2mM L-谷氨酰胺, 2 - 巯基乙醇 (1:280,000 终浓度)和 PSN 抗体 (GIBCO BRL)。在电击穿至 BaF₃ 细胞中之前, 使用氯化铯抽提的质粒 pSLmpl - 8, pSLmpl - 9 和 pSLmpl-44 在 NdeI 位点线性化。供电击穿的 BaF₃ 细胞在 RPMI 1640 培养基中洗涤一次, 并重新悬浮于 RPMI 1640 培养基中, 细胞密度为 10⁷ 细胞/毫升。1ml 重新悬浮的 BaF₃ 细胞与各 30μg 的线性化质粒 DNA_s 混合并转移至分隔的可自由使用的电击穿室 (GIBCO BRL) 中。室温下温育 15 分钟后, 细胞受到电击穿装置 (CELL - PORATORTM; GIBCO BRL)提供的连续电击(800μFad/300v; 1180μFad/300v)。经过 5 分钟的恢复之后, 经电击穿的细胞转移到 10ml 的完全培养基上, 并置于培养箱中 15 - 24 小时 (37 °C, 5 % CO₂)。细胞经离心, 重新悬浮于 10ml 的含 1600μg/ml G418 的完全培养基中, 以有限的稀释度涂布于 96 - 孔的组织培养皿中, 分离抗 G418 的克隆。通过 Northern 印迹分析 BaF₃ mRNA 中 MPL 受体转录物的存在推导抗 G418 的 BaF₃ 克隆中 MPL 受体的表达。发现一个指定为 BaF₃/MPLR1.1 的细胞系可表达高水平的类型 I 小鼠 MPL 受体 mRNA, 该细胞系用于条件培养基中转染 BHK570 细胞 MPL 配体活性的后续检测。表达类型 II 的受体 mRNA 的一个细胞系被指定为 BaF₃/MPLR₂。

实施例 IV. 可溶性小鼠 MPL 受体的产生

通过把 pSLmpl-9 (一个含有以上描述的编码全长小鼠类型 I 的 MPL 受体 cDNA 的哺乳动物表达质粒) 中的 DNA 区段与 pSLmpl-26 (一个构建的在细菌中产生可溶性小鼠类型 I 的 MPL 受体的表达质粒) 相连接, 产生编码可溶性小鼠类型 I MPL 受体

的哺乳动物表达质粒。

以全长受体质粒 pSLmpl-9 为模板, 利用引物 ZC6704 (SEQ ID No:10) 和引物 ZC6703 (SEQ ID No:11) 通过 PCR 分离编码小鼠类型 I 的 MPL 可溶性受体 cDNA 区段。为了便于定向克隆, 在引物 ZC6704 和 ZC6703 各自的 5' 端分别引入 EcoRI 和 XhoI 位点。引物 ZC6703 也编码一个蛋白质激酶的框架内的共有靶序列, 使得用 ^{32}P - ATP 可以在体外标记提纯的可溶性受体 (Li et al., Proc Natl. Acad. Sci. USA. 86:558-562,1989)。在一种混合物中进行 PCR, 该混合物含 10 μl 的 10x ULTMATM DNA 聚合酶缓冲液 (Roche Molecular Systems, Inc.), 6 μl 的 25mM MgCl₂, 0.2 μl 的各含 10mM 的 dATP, dGTP, dTTP 和 dCTP (Pharmacia LKB Biotechnology Inc.) 的脱氧核苷酸三磷酸溶液, 11 μl 的 4.55pmol/ μl 引物 ZC6704 (SEQ ID No:10), 21 μl 的 2.43pmol/ μl 引物 ZC6703 (SEQ ID No:11), 50.3 μl 的水, 1 μl 的 50ng/ μl 的 Hind IV 和 XbaI 消化的 pSLmpl-9 和 1 μl 的 6U/ μl ULTMATM DNA 聚合酶 (Roche Molecular Systems, Inc.)。按厂商的说明在反应中使用 AmpliWaxTM (Roche Molecular Systems, Inc.)。聚合酶链反应进行三个循环 (95 $^{\circ}\text{C}$ 1 分钟, 50 $^{\circ}\text{C}$ 1 分钟, 72 $^{\circ}\text{C}$ 2 分钟) 后, 再在增加杂交严谨度时进行 11 个循环 (95 $^{\circ}\text{C}$ 1 分钟, 55 $^{\circ}\text{C}$ 30 秒, 72 $^{\circ}\text{C}$ 2 分钟), 然后于 72 $^{\circ}\text{C}$ 下温育 10 分钟。扩增的产物经一系列苯酚/氯仿和氯仿抽提后, 通过一个 400 孔径大小的凝胶过滤柱 (Clontech Laboratories, Inc.) 层析。PCR 产物在 20 μg 糖原载体和 2.5M 醋酸铵存在时用乙醇沉淀。沉淀带重新悬浮于 32 μl 的水中。向 16 μl 的重新悬浮的 PCR 产物中加入 2 μl 10x 的缓冲液 H (Boehringer Mannheim Corp), 1 μl 的 10U/ μl EcoRI (Boehringer Mannheim Corp.) 和 1 μl 的 40U/ μl XhoI (Boehringer Mannheim Corp.) 消化于 37 $^{\circ}\text{C}$ 下进行 1 小时。通过加热至 65 $^{\circ}\text{C}$, 持续 15 分钟终止消化, 再于 0.7% 的低一熔点琼脂糖凝胶上纯化消化的 PCR 产物。通过用 β - 琼脂酶 I (New England Biolabs) 消化胶基质, 从低熔点凝胶中回收片段。

产生的 PCR 产物编码小鼠类型 I MPL 受体的 N - 端胞外结构域 (SEQ ID No:17) 的第 27 至 480 个残基)。在缺乏推定的受体跨膜结构域 (SEQ ID No:17 第 483 至 504 个残基) 时, 表达的蛋白质预期在一种合适的信号肽存在下分泌。除了用 pSLmpl-8 作为模板外, 用上面描述的 PCR 条件, 获得一个小鼠类型 II 可溶性 MPL 的受体编码 cDNA。两种受体片段经 DNA 序列测定进行验证。

可溶性小鼠类型 I 和类型 II MPL 受体编码 DNA 片段克隆于 EcoRI 和 XhoI 消化的载体 pOmpA2 - 5 中, 分别产生 pSLmpl-26 和 pSLmpl-27。质粒 pOmpA2 - 5 由 pOmpA₂ 改变而来 (Gbrayab et al., EMBO.J 3: 2437-2442,1984), pOmpA₂ 为一种设计的把重组蛋白引入胞外周质中的细菌表达载体。通过用一个合成的 42bp 序列代替 pOmpA₂ 的 EcoRI 和 BamHI 位点间的一个 13bp 的序列, 完成 pOmpA2-5 的构建。通过两个 42 个核苷酸互补的容聚核苷酸 (ZC6707, SEQ ID No:12; ZC6706, SEQ ID No:13) 退火产生所述序列, 当碱基配对形成 EcoRI 和 BamHI 粘性末端时, 该序列有利于定向克隆至 EcoRI 和 BamHI 消化的 pOmpA₂ 上。在该插入序列内, 是一个有关细菌前导序列和以上描述的小鼠 MPL 可溶性受体编码 cDNA 读码框架内的 XhoI 位点以及位于 Xho 3' 端的读码框架内的 6 个组氨酸密码子束 (该密码子束可通过金属螯合作用亲合层析 (Houchuli et al., Bio/Fechnol. 6:1321-1325,1988) 提纯重组蛋白)。在编码组氨酸束的序列之后是一个框架内的终止密码子。 pOmpA2-5, pSLmpl-26 和 pSLmpl-27 的确实性由 DNA 序列测定来证实。

pLDmpl-53 为一种产生可溶性小鼠类型 I MPL 受体的哺乳动物表达质粒, 其构建是通过把 pSLmpl-9 和 pSLmpl-26 的 DNA 区段连接至表达载体 pHZ-200 (pHZ-1 中, 用一个二氢叶酸不原酶序列代替新霉素粘性基因) 中而完成的。在细菌表达质粒 pSLmpl1-26 的构建中, 1164 bp 的 EcoRI/BamNI cDNA 片段代替缺失的哺乳动物信号序列。 pSLmpl-26 的 416bp - BamHI 片段为

可溶性 MPL 受体的羧基末端，激酶标记结构域，多聚组氨酸束和翻译终止子提供了编码序列。这两个片段经凝胶纯化，再克隆至 pBluescript[®] KSt(Stratagene Cloning Systems)的 EcoRI/BamHI 的位点上，产生质粒 pBS8.76LD-5。就 pBS8.76LD-5 中的 1164bp 的 pSLmpl-9 来源的 EcoRI/BamHI 片段而论 416bp 的 pSLmpl-26 来源的 BamHI 片段的正确方向用引物 ZC6603 (SEQ ID No:8) 和引物 ZC6703 (SEQ ID No:11) 通过 PCR 来测定。PBS8.76LD-5 中的多接头序列中的 XbaI 位点可以使重组的受体 cDNA 切割成一个 1.5Kb EcoRI/XbaI 片段；在用 EcoRI 和 XbaI 消化载体之后，该 1.5Kb 的片段可克隆至 pHZ-200 载体中。大量制备产生的哺乳动物表达质粒 pLDmpl-53，以转染至 BHK 细胞中。

用磷酸钙沉淀法，把 20 μ g 提纯的 pLDmpl-53 质粒转染至 BHK570 细胞中。5 小时之后，细胞用 15% 的甘油刺激 3 分钟，以利于 DNA 的吸收。加入新鲜的培养基过夜。第二天把细胞分成各种稀释度，然后加入含 1 μ M 的氨基嘌呤的选择培养基。大约 2 周之后，可以看见不连续的（即分散的）抗氨基嘌呤的菌落。抗性菌落或者被收集起来，或者作为不同的克隆保存起来。对来源于已收集菌落的耗尽营养的培养基立即检测可溶性 MPL 受体蛋白的存在。

通过蛋白质羧基末端存在的多聚组氨酸束与含有固定的 Ni²⁺ (HIS - BIND[™]; Novagen, Madison, WI) 的金属螯合树脂的相互作用分离可溶性 MPL 受体蛋白。将 pLDmpl-53 库来源的无血清的营养耗尽的培养基通过此树脂，用 1M 的咪唑洗提结合的蛋白质。SDS - PAGE 分析显示在大约 67KD_a 处有一条单一带。该蛋白再进行 N - 末端的氨基酸序列分析，并证实为小鼠 MPL 受体。

可溶性小鼠 MPL 受体由 BHK 转染子库中提纯而来，该转染子库已用可溶性小鼠类型 I MPL 受体表达质粒 pLDmpl-53 转染过。纯化的可溶性受体基本上按照厂商说明固定于 CNBr - 活化的 SEPHAROSE[™]4B (Pharmacia LKB Biotechnology, Inc.) 基质

上，并用于 24 - 11 - 5 细胞的培养培养基中的 MPL 活性的亲和提纯。亲和基质包装于 XK16 柱子（Pharmacia LKB Biotechnology Inc.）中。细胞 24 - 11 - 5 的培养培养基在 10KD 剪下的空心纤维膜（A/G Tology Corp., Needham, MA）上浓缩，并以 1ml/分钟流速上样至 MPL 受体亲和柱的底部。用含 0.5M NaCl 和 0.0% 叠氮化钠的磷酸盐缓冲液（PBS）洗涤柱子。用 3M 的异硫氰酸钾（Sigma Chemical Company, St. Louis, MO）以 0.5ml/分钟流速从柱子中洗提 MPL 活性。通过透析 PBS，去除异硫氰酸钾。通过 MTT 增生试验（在实施例 VII 中描述的）鉴定活性部分。

实施例 V. MPL 受体的配体表达细胞系的分离和特征

BaF₃/MPLR1.1 细胞为依赖于 IL - 3 的细胞，该细胞表达稳定转染的类型 I 小鼠 MPL 受体。

设计突变和选择方案，经突变 BaF₃/MPLR1.1 细胞和在缺乏外源 IL - 3 时选择自泌生长来分离表达 MPL 受体的配体细胞系。大约 1.2×10^6 BaF₃/MPLR1.1 的细胞经沉淀收集，并用 GM（RPMI 1640 培养基，辅以 2 - 巯基乙醇（1:240,000 的终浓度），2ml L - 谷氨酰胺，110 μ g/ml 丙酮酸钠，50 μ g/ml G418 和 10% 的热钝化小牛胎血清）洗涤。细胞重新悬浮于 2ml 的含有 0.15%（V/V）的诱变剂甲基磺酸乙酯（EMS）的 GM 中，并于 37 $^{\circ}$ C 下温育 2 小时。经温育之后，细胞用 PBS 洗涤一次，用 GM 洗涤一次，然后以大约 40,000 细胞/ml 的密度涂布于补加有 5% 作为 IL - 3 来源的 WEHI - 3 条件培养基（Becon, Dickinson Labware, Bedford, MA）的 GM 中的 10cm 平板上。在选择不依赖于 IL - 3 的生长之前，细胞经七天恢复期（于 5% 的 CO₂ 下，37 $^{\circ}$ C 温育）。温育期之后，培养物中密布有可生长的细胞。这些细胞用 GM 洗涤，再于缺乏 WEHI - 3 的条件培养基的 GM 中培养。经过 11 天的选择之后，可观察到少量可生长的细胞。不依赖于 IL - 3 的培养物中的可生长的细胞密度估计为 250 细胞/毫升。1ml 的不依赖于 IL - 3 的培养物涂布于 24 - 孔培养皿中的 19 个孔中，以作进

一步的特性分析。

对上述不依赖于 IL - 3 生长的 BaF₃/MPLR1.1 细胞来源的条件培养基测定 BaF₃/MPLR 细胞的增生活性。发现来源于所有 19 个不依赖于 IL - 3 生长的细胞库的培养基在 MTT 增生检测中具有活性（实施例 VII 中描述）。在 2μg/ml 大鼠抗小鼠 IL - 3，抗小鼠 IL - 4 或两种中性抗体（Pharmingen, San Diego, CA）都存在时对阳性培养基检测增生活性，以鉴定不依赖于 IL - 3 生长的表达那些细胞因子的突变体。（在一个先前的实验中，发现 BaF₃ 细胞与 IL - 4 也有反应。）发现只有从平板井 11 的细胞（指定为“24 - 11”细胞）来源的条件培养基具有不被 IL - 3 或 IL - 4 抗体中和的活性。

以上描述的突变和选择方案被应用于 5 种其他的 BaF₃/MPLR1 克隆（BaF₃/MPLR1 克隆井 4，9，12，15 和 18，分别指定为 BaF₃/MPLR1 · 4，· 9，· 12，· 15 和 · 18）。发现 17 个分离物具有刺激 BaF₃/MPLR1 细胞增生的条件培养基。发现所有培养基的活性可以被抗 - IL - 3 或 IL - 4 抗体单独或联合中和。这些克隆没有作进一步的特征描述。

对 24 - 11 库来源的条件培养基的增生活性作了详细的特征描述。24 - 11 细胞库可再分为 19 个亚库，并对其条件培养基再检测活性。所有 19 个亚库（如 24 - 11 - 1 到 24 - 11 - 19）在缺乏外源的 IL - 3 时均可刺激依赖于 IL - 3 生长的细胞增生。其活性不能被 IL - 3 或 IL - 4 的中和抗体或两种抗体的联合抑制。

实施两个实验测定 24 - 11 活性的特异性。对其条件培养基检测对不表达 MPL 受体的对照 BaF₃ 细胞的增生活性。在缺乏外源 IL - 3 时，在任何一种 19 个 24 - 11 细胞亚库来源的条件培养基中，没有观察到对照 BaF₃ 细胞的增生。在第二个实验中，检测增生活性被提纯的可溶性 MPL 受体的抑制作用。BaF₃/MPLR1 细胞培养于补加有 50% 的 24 - 11 的条件培养基的 GM 中。向每个样品中加入类型 I 小鼠可溶性受体，至终浓度为 0.0，0.625, 1.25, 2.5 或 5.0μg/ml。4 天之后通过 MTT 细胞增生试验对结果实行计分。

在 0.625 至 1.25 $\mu\text{g}/\text{ml}$ 可溶性 MPL 受体浓度下, 24 - 11 条件培养基的增生活性被完全封闭。具完全抑制活性的可溶性受体浓度对 BaF₃/MPLR₁ 细胞用 IL - 3 或 IL - 4 刺激没有影响。结果表明可溶性受体竞争 24 - 11 培养基的刺激活性, 并与 24 - 11 细胞表达 MPL 受体配体的假设一致。

通过有限的稀释涂布平板, 分离了来源于 24 - 11 细胞的克隆。一个指定为 24 - 11 - 5 # 3 的克隆在其与 24 - 11 库相关的条件培养基中显示出高水平的增生活性。发现其增生活性相当于 WEHI - 3 细胞的以 1:2000 稀释的条件培养基的活性 (Becton Dickinson Labware) 。

实施例 VI. 24-11-5#3 cDNA 文库的构建

用异硫氰酸胍处理, 再经 CsCl 离心, 从大约 2.7×10^8 24 - 11 - 5 # 3 细胞中提取总 RNA (Chirgwin et al., *ibid*)。用一个 OLIGOTEX - dT - mRNA 分离试剂盒 (Qiagen Inc., Chatsworth, CA) 按照厂商的说明分离 poly (A)⁺ RNA。

24 - 11 - 5 # 3 细胞第一链在 4 个单独平行的反应中合成。每个反应含有浓度为 1.6 $\mu\text{g}/\text{ml}$ 的 7 μl 的 polyd (T)-选择的 poly (A)⁺ 24 - 11 - 5 # 3 的 RNA 和 2.5 μl 的 20 $\text{pmol}/\mu\text{l}$ 的含一个 XhoI 限制性位点的第一链引物 ZCb172 (SEQ ID No:14)。混合物于 65 $^{\circ}\text{C}$ 下加热 4 分钟, 再于冰上冷却。通过向 RNA - 引物混合物中加入 8 μl 的第一链缓冲液 (5x SUPERSCRIPTM 缓冲液; GIBCO BRL), 4 μl 的 100mM 二硫苏糖醇和 2 μl 的包含各 10mM 的 dATP, dGTP, dTTP 和 5 - 甲基 - dCTP (Pharmacia LKB Biotechnology Inc.) 的脱氧核苷酸三磷酸溶液, 开始第一链 cDNA 合成。反应混合物于 45 $^{\circ}\text{C}$ 下温育 4 分钟, 再加 10 μl 的 200U/ μl , RNaseH - 逆转录酶 (GIBCO BRL)。通过把 10 μCi 的 ³²p - α dCTP 加入到 10 μl 任一反应混合等分样中标记反应液分析第一链合成的效率。反应液于 45 $^{\circ}\text{C}$ 下温育 1 小时, 再于 50 $^{\circ}\text{C}$ 下温育 15 分钟。标记反应中未结合的 ³²p - α dCTP 通过 400 孔径大小的凝胶过滤柱子 (Clontech Laboratories) 层析去除掉。把未标记的第一链反

应液集中起来，未结合的核苷酸在 32 μ g 的糖原载体，2.5M 醋酸铵和 2.5 倍体积的乙醇存在时通过两次 cDNA 沉淀去除掉。未标记的 cDNA 重新悬浮于 144 μ l 的水中，供第二链合成使用。通过琼脂糖凝胶电泳测定标记的第一链 cDNA 的长度。

在启动的第一链指导第二链合成产生 DNA 发夹结构的条件下，根据第一链 cDNA 进行第二链的合成。每个第二条链的反应包含 48 μ l 的未标记的第一链 cDNA，16.5 μ l 的水，20 μ l 的 5x 聚合酶 I 缓冲液（100mM Tris:HCl, pH7.4, 500mM KCl, 25mM MgCl₂, 50mM(NH₄)₂SO₄），1 μ l 的 100mM 的二硫苏糖醇，1 μ l 的含 10mM 各脱氧核苷三磷酸的溶液，3 μ l 的 5mM β -NAD，1 μ l 的 3U/ μ l 的大肠杆菌 DNA 连接酶（New England Biolabs Inc.）和 5 μ l 的 10U/ μ l 大肠杆菌 DNA 聚合酶 I（Amersham Corp.）。在室温下组装反应，于室温下温育 5 分钟后再加入 1.5 μ l 的 20U/ μ l Rnase H（GIBCO BRL.）。向任一第二链合成反应的 10 μ l 等分样中加入 10 μ Ci ³²p - α dCTP 进行标记以监测第二链合成的效率。反应液于 15 $^{\circ}$ C 下温育 2 小时，再于室温下温育 15 分钟。通过 400 孔径大小的凝胶过滤柱（Clontech Laboratories）层析以去除标记反应中未结合的 ³²p - α dCTP，然后通过琼脂糖凝胶电泳分析。收集未进行标记的反应液，用苯酚/氯仿和氯仿抽提，再在 2.5M 醋酸铵存在下用乙醇沉淀 cDNA。

用绿豆核酸酶切割发夹结构的单链 DNA。反应混合物含有 100 μ l 的第二链 cDNA，20 μ l 的 10x 的绿豆核酸酶缓冲液（Stratagene Cloning Systems）和 6 μ l 的 50U/ μ l 的绿豆核酸酶（Promega Corp.）。反应于 37 $^{\circ}$ C 下温育 30 分钟。通过加入 20 μ l 的 1M Tris:HCl, pH8.0，再经以上描述的连续的苯酚/氯仿和氯仿抽提终止该反应。抽提之后，DNA 用乙醇沉淀，重新悬浮于水中。

重新悬浮的 cDNA 用 T₄ DNA 聚合酶进行末端补平。重新悬浮于 188 μ l 水中的 cDNA 与 50 μ l 5x 的 T₄ DNA 聚合酶缓冲液（250mM Tris:HCl, pH8.0, 250mM KCl, 25mM MgCl₂），3 μ l 0.1M 的二硫苏糖醇，4 μ l 的含 10mM 的各脱氧核苷酸三磷酸的溶液和 5 μ l 的

1U/ μ l T₄ DNA 聚合酶 (Boehringer Mannheim Corp.) 相混合。在 15 °C 下温育 30 分钟后, 通过加入 10 μ l 的 0.5M EDTA, 接着经一系列以上描述的苯酚/氯仿和氯仿抽提终止该反应。DNA 通过 400 孔径大小的凝胶过滤柱子 (Clontech Laboratories Inc.) 层析以去除微量的蛋白质和不到约 400bp 长度的较短的 cDNA。在 10 μ g 糖原载体和 2.5M 的醋酸铵存在下用乙醇沉淀 DNA, 并重新悬浮于 15 μ l 的水中。根据渗入的 ³²p - α dCTP, 经估算由 40 μ g 起始 mRNA 模板合成的 cDNA 产量为大约 8 μ g。

EcoRI 接头连接至以上描述的 cDNA 的 5' 末端, 使之能够克隆至一个表达载体上。将 10 μ l 等份的 cDNA (~5 μ g) 和 21 μ l 的 65pmol/ μ l 的 EcoRI 接头 (Pharmacia LKB Biotechnology Inc.) 与 4 μ l 的 10x 连接酶缓冲液 (Promega Corp.)、3 μ l 的 10mM ATP 和 3 μ l 的 15U/ μ l T₄ DNA 连接酶 (Promega Corp.) 混合。反应于 9 °C 温育过夜 (约 48 小时)。通过 140 μ l 的水, 20 μ l 的 10x 缓冲液 H (Boehringer Mannheim Corp) 和 65 °C 下温育 40 分钟终止反应。温育之后, cDNA 用以上描述的苯酚/氯仿和氯仿抽提以及在 2.5M 醋酸铵和 1.2 倍体积的异戊醇存在时沉淀 cDNA。离心之后, 用 70% 乙醇洗涤 cDNA 沉淀带, 空气干燥和重新悬浮于 89 μ l 的水中。

为了便于 cDNA 定向克隆至一个表达载体中, cDNA 用 XhoI 消化, 产生具有 5' EcoRI 粘性末端和 3' XhoI 粘性末端的 cDNA。cDNA 的 3' 末端的 XhoI 限制性位点已事先用引物 ZC6172 (SEQ ID No:14) 引入。在以上描述的含有 89 μ l 的 cDNA、10 μ l 的 10x 缓冲液 H (Promega Corp.) 和 1.5 μ l 的 40U/ μ l XhoI (Boehringer Mannheim Corp.) 的混合物中进行限制性酶消化。消化于 37 °C 下进行 1 小时。通过一系列的苯酚/氯仿和氯仿抽提及通过 400 孔径大小的凝胶过滤柱子 (Clontech Laboratories Inc.) 层析终止此反应。

cDNA 经乙醇沉淀, 用 70% 乙醇洗涤, 空气干燥, 并重新悬浮于 20 μ l 的 1x 含凝胶的缓冲液 (10mM Tris:HCl, pH8.0, 1mM

EDTA,5%甘油和 0.125%溴酚蓝)。

将重新悬浮的 cDNA 加热至 60 ℃ 持续 5 分钟, 于冰上冷却并在 0.8 % 低熔点琼脂糖凝胶 (SEA PLAQUE GTG™ 低熔点琼脂糖; FMC Corp) 上电泳。从胶上切下污染的接头和低于 0.5Kb 长的 cDNA。颠倒电极, 电泳直至 cDNA 集中于近泳道起点处。切下含浓缩的 cDNA 的凝胶区, 置于一个微量离心管中, 测定胶条的大致体积。向管中加入大约为胶条件积 3 倍的水等份样 (300 μ l), 通过加热至 65 ℃ 融化琼脂糖 15 分钟。在样品平衡至 45 ℃ 时, 加入 5 μ l 的 1U/ μ l β - 琼脂酶 I (New England Biolabs, Inc.), 再于 45 ℃ 下温育 90 分钟, 以消化琼脂糖。温育之后, 加入 40 μ l 的 3M 醋酸钠至样品中, 混合物于冰上温育 15 分钟。样品在室温下以 14,000 \times g 离心 15 分钟, 以去除未消化的琼脂糖, 接着通过 400 孔径大小的凝胶过滤柱子 (Clontech Laboratories) 层析。cDNA 经乙醇沉淀, 用 70 % 乙醇洗涤, 空气干燥, 重新悬浮于 70 μ l 的水中, 供激酶反应使用, 酶连接的 EcoRI 接头磷酸化。

向 70 μ l 的 cDNA 溶液中加入 10 μ l 10x 连接酶缓冲液 (Stratagene Cloning Systems), 并把混合物加热至 65 ℃ 持续 5 分钟。该混合物于冰上冷却, 加入 16 μ l 的 10mM ATP 和 4 μ l 的 10U/ μ l 的 T₄ 多聚核苷酸激酶 (Stratagene Cloning Systems)。反应混合物于 37 ℃ 下温育 1 小时, 然后加热至 65 ℃ 持续 10 分钟, 再经一系列苯酚/氯仿和氯仿抽提终止该反应。磷酸化的 cDNA 在 2.5M 醋酸钠存在下经乙醇沉淀, 用 70 % 的乙醇洗涤, 空气干燥和重新悬浮于 10 μ l 的水中。磷酸化的 cDNA 的浓度估计为 ~40fmol/ml。

pDX 哺乳动物表达质粒 (在美国专利 4,959,318 中已公开) (附图) 经改造后可接受用 EcoRI - XhoI 末端合成的 cDNA。通过用 SalI 消化质粒, 和用 T₄DNA 聚合酶补平 SalI 的粘性末端后再重新环化质粒从而使 pDX 上的内源 SalI 位点消失。重新环化的质粒用 EcoRI 消化后, 用由两个互补的寡核苷酸 ZC6936 (SEQ ID No:15) 和 ZC6937 (SEQ ID No:16) 组成的较短的多接头序列

与之相连产生质粒 pDX.ES。pDX.ES 上引入的多接头序列包含 EcoRI 和 SalI 位点，便于用 EcoRI - XhoI 末端合成的 24 - 11 - 5 cDNA 的定向克隆。

通过将 EcoRI - XhoI 24 - 11 - 5 cDNA 连接到 EcoRI/SalI 消化的 pDX.ES 上制备一个质粒的 cDNA 文库。经使用基因脉冲仪/脉冲控制仪和 0.2cm 的透明小容器 (Bio-Rad Laboratories, Hercules, CA)，用 0.2KV，400 ohm 和 25 μ FAD 把混合物电击穿至大肠杆菌中 (ELECTROMAX DH10BTM 感受态细胞；GIBCO BRL, Gaithersburg, MD)。细胞用 Luria 培养基稀释至 1.5ml，于 37 $^{\circ}$ C 下温育 45 分钟，然后加入 0.75ml 的 50% 的甘油。等分转染的细胞，贮藏于 -70 $^{\circ}$ C 下备用。80fmols 的 cDNA 可产生超过 700,000 个独立的重组质粒。

实施例 VII. MPL 活性的 24 - 11 - 5 cDNA 文库的表达筛选

将 24 - 11 - 5 # 3 cDNA 文库涂布于大约 2000 个 10cm 直径的补充有 100 μ g/ml 氨基青霉素的 Luria 琼脂培养基中。涂板的密度为每个平板 200 至 250 个细菌菌落。用 MAGIC MINIPREPTM DNA 提纯树脂 (Promega Corp.)，按照厂商的说明从每个细菌平板中制备供转染至 BHK 570 细胞的质粒 DNA。质粒 DNA 保存于 -20 $^{\circ}$ C 直至转染到 BHK 570 细胞中。

将 24 - 11 - 5 # 3 cDNA 质粒库(各含有约 200 至 250 个 cDNA 克隆)用 2, 3 - 二油基 (dioleoly) - N - [2 - (精胺羧基酰胺) 乙基] - N, N - 二甲基 - 丙烷铵三氟乙酸和二油基 - 磷脂酰胆碱在水中的 3:1 脂质体配方 (LIPOFECTAMINETM; GIBCO BRL) 转染至 BHK570 细胞中。将 20 μ l 的 30ng/ μ l 的 DNA 加入到 1:10 稀释的 LIPOFECTAMINETM 溶液中，于室温下温育 30 分钟。温育之后，向 DNA/LIPOFECTAMINETM 混合物中加入 160 μ l 的无血清培养基 (Hams F12:Dulbeccos MEM (1:1)，补充有 2mM L - 谷氨酰胺，0.11mg/ml 丙酮酸钠，5 μ g/ml 胰岛素，5 μ g/ml 胎球蛋白，10 μ g/ml 的转铁蛋白，2ng/ml 硒 IV 氧化物和 25mM HEPES 缓冲液)，并转移至 24 孔的含有 - 100,000 个

BHK570 细胞的微滴平板中。细胞于 37 ℃，5 % CO₂ 下温育 4 小时，之后再加入 200 μl 的 BHK 生长培养基（Dulbecco's modified Eagles's 培养基，补充有 2mM L - 谷氨酰胺，0.11mg/ml 丙酮酸钠，5 % 热钝化小牛胎儿血清和 100x PSN 抗生素（GIBCO BRL））。细胞温育 16 小时。除去培养基，更换为 0.5ml 的新鲜培养基，并于 MPL 活性检测前使之适应 48 小时。

使用细胞增生试验检测转染 BHK570 细胞文库的条件培养基中 MPL 活性的存在。向补充有 2mM L - 谷氨酰胺，PSN 抗生素（GIBCO BRL），0.00036% 2 - 巯基乙醇及 10 % 热钝化水中胎儿血清的 RPMI 1640 培养基（JRH Bioscience Inc., Lenexa, KS）上的 100 μl 的 10⁶/ml 洗涤过的 BaF₃/MPLR1.1 细胞中加入 100 μl 的条件培养基。在测定增生之前，试验细胞于 37 ℃，5 % CO₂ 条件下温育 3 天。

根据 3 - (4, 5 - 二甲基噻唑 - 2 - 基) - 2, 5 - 二苯四唑溴（MTT）（Mosman, *J. Immunol. Meth* 65:55 - 63, 1983）的代谢分解，用比色试验定量检测 MPL 存在时的细胞增生。将 20 μl 的 10mg/ml 的 MTT 溶液（Polyscience, In., Warrington, PA）加入到 100 μl 的 BaF₃/MPL R1.1 试验细胞中，细胞于 37 ℃ 下温育。4 小时后，加入溶于异丙醇中的 0.04N HCl，混合溶液，在一个 EL320 ELISA 读数计上（Bio - Tek Instruments Inc., Highland Park, VT）读出 570nm 下样品的吸光值。

将发现为阳性的一个质粒库（指定为 T1081）转染至 BHK570 细胞中。转染子的上清液在 MTT 增生试验中显示出阳性信号。PCR 和抗体中和实验表明活性并非由于 IL - 3 或 IL - 4 引起。

从阳性库中来源的质粒用来转化大肠杆菌 DH10B，并把细胞涂布于平板上（42 个板中每板约 15~20 个菌落；10 个板中每板约 90 个菌落；8 个板中每板约 250 个菌落）。对每个板进行复制，并保存于 4 ℃ 下。对原始板上的菌落进行划线培养，并于液体培养基中快速生长几小时，然后制备 DNA。

将来源于亚库的质粒 DNA 转染至 BHK570 细胞中，收集细胞

上清液并按以上方法进行测定。约 2 小时后，通过显微镜观察一个亚库（# 22）被计划阳性（伸长的细胞形状）。几个小时之后，另外两个亚库（# 19 和 # 28）也被计为阳性。每个阳性亚库剩余的上清液与对照 BaF₃ 细胞对比检测，发现没有活性。此外，发现三个阳性亚库的活性被可溶性类型 I MPL 受体抑制。

让三个阳性亚库的复制板生长几小时，然后挑出单个菌落，用来接种 3ml 的培养基。培养物于 37 °C 下生长大约 8 小时，然后按以上描述的小量制备的方法提取质粒 DNA。质粒 DNA 转染至 BHK570 细胞中，约 10 小时后收获上清液并测定活性。一小时之后，一个克隆（指定为 T1081 - 19 - 215，相当于亚库 19）被计为阳性。此克隆重新划线培养成单个菌落。从 12 个菌落中提取 DNA 并转染至 BHK570 细胞中。在以后的试验中，所有 12 个转染子都被计为阳性。来源于 12 个阳性菌落中任一菌落的 DNA 转化大肠杆菌 DH52。该质粒被指定为 pZGmol-1081。该转化子于 1994 年 2 月 14 日贮藏于美国典型培养物保藏中心（American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD），其保藏号为 69566。

编码血细胞生成蛋白（血小板生成素）的 cDNA 的核苷酸序列已经测定（SEQ ID No:1）。其编码的氨基酸序列分析表明成熟蛋白的氨基末端位于第 45 个核苷酸。两个甲硫氨酸密码子，位于 SEQ ID No:1 的第 105 位和第 174 位置上，似乎为起始密码子，主要的起始位点预期位于第 174 个位置上。

实施例 VIII. 重组血小板生成素的血细胞生成活性

从雌性 CD - 1 怀孕后的小鼠的股骨和胫骨中收获骨髓，并置于 25ml 的 CATCH 缓冲液（99mg 茶碱，0.75g 的柠檬酸钠，75mg 的腺苷，20ml10x 的无 Ca[#] Mg[#] 的 Hank's 平衡盐溶液，溶于每 200ml 的去离子水中，pH7.4）中。用 25ml 的移液管移取把细胞悬浮为单个细胞悬浮液。用 CATCH 缓冲液使体积增加至 50ml，并于 1000rpm 离心 7 分钟使细胞沉淀。用 25ml CATCH 缓冲液重新悬浮细胞沉淀，于包有第一层塑料粘合薄膜的 T₇₄ 组织培

养烧瓶中于 37℃ 下温育 2 小时。通过于 1000rpm 离心 7 分钟，收获非粘着的细胞使之沉淀。细胞沉淀重新悬浮于 15ml 的 α -MEM + 10% FBS (+ L-谷氨酰胺，丙酮酸钠和 PSN 抗生素) 中，并在以上描述的于第一层塑料粘合薄膜上再覆盖第二层塑料粘合薄膜的 T75 烧瓶中培养。在最后离心和重新悬浮之后，对细胞计数。密度为 576,000 细胞/ml 的 0.5ml 细胞悬浮液与对照 BHK 细胞的样品培养基或与用 pZGmpl-1081 转染的 BHK 细胞的条件培养基一起涂布于 24-孔的组织培养平板中。37℃ 下培养 3 天之后，收获细胞并按以上描述的方法染色。

从用标准条件培养基处理的对照孔中收获 150 μ l 的细胞。从用 10pZGmol-1081 转染的 BHK 细胞的条件培养基处理的孔中收获 50 μ l 的细胞。这些样品经离心，制备标准的显微载玻片。

这些载玻片用 100% 的甲醇固定，然后用 1:1 的 Wright's (0.5g 的 Wright 染料溶于 300ml 的甲醇中) / 水浸泡 6 分钟，经水冲洗，再干燥。然后将载玻片用溶于 Sorensen 缓冲液 (2.28 克 KH_2PO_4 / 2.38g NaPO_4 溶于 250ml 水中) 的 Giemsa 染料浸泡，经水冲洗，再干燥。

当调整使用体积之后，BHK/pZGmpl-1081 培养基样品每 150 μ l 体积含有 120 个巨核细胞，相比之下，对照培养基每 150 μ l 体积只含有 9 个巨核细胞。此外，处理实验样品的巨核细胞比对照细胞在显微镜下观察明显的要大，对多核内含物的染色明显要深。

在缺乏血清条件下收集突变的 BaF₃/MPLR1.1 细胞系 24-11-5 # 3 的条件培养基，用一个 10KD 剪下的 Amico Inc. (Beverly, MA) 过滤装置浓缩 20 倍。从小鼠股骨收获骨髓并悬浮于 IScove's Mulbecco's Media (GIBCO BRL) + 15% 小牛胎儿血清 (FCS) 中。悬浮之后，对巨核细胞进行计数，用每板 0.9ml 密度为 75000 个细胞/ml 的悬浮液涂布于调整为含有 50% 甲基纤维素、15% FCS，10BSA 和 0.6% PSN (半固体培养基) 培养基的 1ml 组织培养平板中。加入各种条件培养基和对照样品使总体积达到 1ml。该板于 37℃ / 5% CO₂ 条件下培养 6 天，然后于显微镜下

观察粒性白细胞/巨噬细胞 (GM) 菌落数。对于 24 - 11 - 5 # 3 条件培养基中培养的平板观察, 具有较弱的类似 GMCFS 活性, 产生的菌落数为 24 个, 相比之下, 负对照样品的菌落数为 0, 用正对照 (商陆促脾细胞分裂素条件培养基) (PWMSM) 刺激的平板菌落数为 130 个; PWMSM 的制备是通过在商陆促脾细胞分裂素 (从 Boehringer Mannheim, Mannheim, Indianapolis, IN 获得) + 2 个单位/ml 的促红细胞生成素存在时温育切碎的鼠脾一周。

从鼠胫骨收获骨髓并悬浮于含有 15 % FCS 的 IScove's Modified Dulbecco's 培养基 (GIBCO - BRL) 中, 对成核细胞 (nucleated cell) 进行计数, 并涂布于以上描述的半固体培养基上。用这些细胞检验形成 pZGmpl-1081 插入片段编码的蛋白质的活性的巨核细胞菌落。

将 pZGmpl-1081 稳定转染的 BHK570 细胞库在缺乏血清时培养, 收集其条件培养基。将该条件培养基单独地或与商陆促脾细胞分裂素条件培养基, 重组的鼠 IL - 3, IL - 6 (Genzyme Corp., Cambridge, MA), IL - 11 (Genzyme Corp.) 结合或这些因子的组合一起进行试验。PWMSM 用作阳性对照。非条件培养基用作负对照。

将试验或对照样品加到骨髓培养物中使总体积达到 1ml。涂覆的平板于 37 °C、5 % CO₂ 下培养 6 天, 然后于显微镜下检验巨核细胞菌落数。结果于表 4 显示。总之, BHK/pZGmpl - 1081 条件培养基显示出巨核细胞菌落形成活性, 该活性在这些早期作用因子存在时增加的水平显著地高于任何一个早期作用因子单独作用时的水平。

测定了 BHK/pZGmpl-1081 条件培养基在小鼠体内的活性。收集无血清的培养基并用一个 10Kd 剪下的过滤装置 (Amicon, Inc., Beverly, MA) 浓缩 5 倍。对照 (非条件) 培养基以同样方式浓缩。6 只 BALB/c 小鼠 (Simonsen Laboratories, Inc., Golroy, CA) 经 7 次每天腹腔内注射 0.5ml 的对照或条件培养基。在第 0、3、7 天收集血样, 并对血小板内含物进行计数。如表 5 所示, 结果表明

来源于 BHK/pZGmpl-1081 细胞的条件培养基具有血小板生成活性。

表 4

样品	巨核细胞 菌落
负对照	0
PWMSCM	7
BHK/pZGmpl-1081	2
BHK/pZGmpl-1081+PWMSCM	15
IL-3	1
IL-3+BHK/pZGmpl-1081	8
IL-6	0
IL-6+BHK/pZGmpl-1081	6
IL-11	1
IL-11+BHK/pZGmpl-1081	6
IL-3+IL-6	2
IL-3+IL-6+BHK/pZGmpl-1081	9
IL-3+IL-11	5
IL-3+IL-11+BHK/pZGmpl-1081	15

表 5

血小板计数 ($10^4/\mu\text{l}$)

处理	第 0 天	第 3 天	第 7 天
对照	141	141	87
对照	159	149	184
BHK/pZGmpl-1081	157	160	563
BHK/pZGmpl-1081	169	154	669
BHK/pZGmpl-1081	139	136	492
BHK/pZGmpl-1081	135	187	554

实施例 IX 人血小板生成素基因的分离

利用小鼠 mpl 受体的配体 cDNA 作探针,从扩增的人肺 λ FIX[®] 基因文库 (Stratagene Cloning Systems) 筛选编码人的血小板生成素的基因。测定文库效价,用大肠杆菌菌种 LE - 392 细胞 (Stratagene Cloning Systems) 接种的 30 个 150mm 的平板用 4×10^4 个噬菌斑形成单位 (PFU) 侵染。平板于 37℃ 下温育过夜。用 HYBOND - NTM 尼龙膜 (Amersham) 按厂商推荐的方法提起过滤噬菌斑。过滤膜于含有 1.5M NaCl 和 0.5M NaOH 中于室温下变性处理 7 分钟。滤膜于滤纸上简单地吸干以去除多余的变性溶液,接着用 1M Tris - HCl (pH7.5) 和 1.5M NaCl 中和 5 分钟。用 STRATALINKER[®] 紫外线交联仪 (UV crosslinker) (Stratagene Cloning Systems) 的 1,200 μ 焦尔的紫外线能量把噬菌体 DNA 固定于滤膜上。固定之后,滤膜于 0.25 \times SSC, 0.25% SDS 和 1M EDTA 中于 65℃ 预洗三次。预洗完毕之后,滤膜于经过 0.45 μ M 滤器过滤的杂交液 (5 \times SSC, 5 \times Denhardt's 溶液, 0.2% SDS 和 1mM EDTA) 中预杂交。使用之前迅速地加入热变性、剪切的鲑鱼精 DNA (终浓度为 100 μ g/ml)。滤膜于 65℃ 下预杂交过夜。

按照厂商推荐的方法,通过 ³²p 随机引物用 MEGAPRIMETM DNA 标记系统 (Amersham) 对 pZGmpl-1081 的全长小鼠 TPO cDNA 进行标记。用含有大约 1×10^6 cpm 探针的新鲜杂交溶液更换预杂交液。杂交完毕之后,去掉杂交溶液,滤膜于含有 0.25 \times SSC, 0.25% SDS 和 1mM EDTA 的洗膜溶液中漂洗 4 至 5 次。漂洗完毕之后,滤膜于洗膜溶液中于 50℃ 下连续洗涤 8 次。最后一次洗涤之后,滤膜用一个增感屏于 -70℃ 暴露于放射自显影光片 (XAR - 5; Eastman Kodak Co., Rochester, NY) 下 4 天。

放射自显影检查表明几百个区域与标记的探针杂交。从 100 个区域挑取琼脂栓进行提纯,将每个琼脂栓浸泡于 1ml 的含 1% (v/v) 氯仿的 SM 中过夜 (Maniatis et al., eds., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, 1982)。过夜温育之后,每个琼脂栓上的噬菌体用 SM 稀释至

1:1,000。将 5 μ l 稀释液等分样涂布于大肠杆菌菌株 LE392 细胞上。平板于 37 $^{\circ}$ C 下温育过夜，按上面描述的方法制备噬菌体过滤提起物，预杂交，杂交，洗膜和放射自显影。

放射自显影检查表明，两个原来的分离物呈现较强的阳性信号，其它 18 个分离物呈现较弱的信号。从呈现 20 个信号的每个阳性区域中挑取琼脂栓。琼脂栓按上面描述的方法处理。从每个琼脂栓中洗提的噬菌体以 1:100 稀释于 SM 中，取 1 μ l 稀释液等分样涂布于大肠杆菌菌株 LE392 细胞上。如上面描述的方法一样，平板经温育，制备噬菌体过滤提起物，杂交。滤膜于 55 $^{\circ}$ C 下用洗膜缓冲液洗涤。滤膜的放射自显影显示出杂交区域，这些杂交区域对应于 3 个原来的分离物 8 - 3 - 2，10 - 1 - 1 和 29 - 2 - 1 的单一不连续的噬菌斑。

噬菌体分离物 8 - 3 - 2，10 - 1 - 1 和 29 - 2 - 1 分别指定为 λ ZGmpl - H8， λ ZGmpl-H10 和 λ ZGmpl-H29。按照厂商的说明用 LAMBDA SORBTM 噬菌体吸附法 (Promega Corp., Madison, WI) 提纯分离物 λ ZGmpl-H8， λ ZGmpl-H10 和 λ ZGmpl-H29 的 DNA。通过用 XbaI 消化和琼脂糖凝胶电泳纯化从噬菌体载体 DNA 中分离来源于噬菌体的人基因组 DNA 插入片段。Southern 印迹分析表明，所有三个噬菌体分离物含有与小鼠 mpl 受体的配体 cDNA 探针杂交的序列 (Maniatis et al. *ibid*)。对噬菌体 λ ZGmpl-H8 进行分析，发现 λ ZGmpl-H8 的杂交区域位于长度为 9.5Kb，2.5Kb 和 1Kb 的 3 个 XbaI DNA 片段上。2.5Kb 的片段亚克隆至 XbaI 消化的 BLUESCRIPT[®]SKT 的噬菌粒上 (Stratagene Cloning Systems)，产生质粒 pZGmpl - H82.5。

SEQ ID No:28 和 SEQ ID No:29 分别表示人的 TPO 基因序列和其编码的氨基酸序列。

实施例 X 全长人血小板生成素 cDNA 的分离

利用来源于经鉴定的 pZGmpl-H82.5 上的外显子序列和小鼠 TPO cDNA 的保守的 5' 非翻译序列为特定引物，从人的肝和肾 cDNA 模板通过聚合酶链反应分离全长人 TPO 编码 cDNA。

人体的肾脏、肝脏和肺的 poly d(T)选择的 poly(A)⁺ RNA_s (Cbontech, Palo Alto, CA) 被用来合成第一链 cDNA。用 4μg 的 poly(A)⁺ RNA 与 1μg 的寡聚 d(T)₁₈ (5'磷酸盐号) mRNA 引物 (New England Biolab, Beverly, MA) 混合至终体积为 19μl 进行每个反应。混合物加热至 65 ℃ 持续 5 分钟, 置于冰上冷却。通过向每个 RNA-引物混合物中加入 8μl 的 5x SUPERSCRIPTM 缓冲液 (GIBCO BRL), 2μl 的 100mM 的二硫苏糖醇, 2μl 的含有各 10mM 的 dATP, dGTP, dTTP 和 dCTP (Pharmacia LKB Biotechnology Inc., Piscataway, NJ) 的脱氧核苷酸三磷酸溶液, 2μl 的 1μCi/μl ³²p-α-dCTP (Amersham, Arlington Heights, IL) 和 8μl 的 200U/μl SUPERSCRIPTM 逆转录酶 (GIBCO BRL) 开始 cDNA 合成。反应于 45 ℃ 下温育 1 小时, 用 TE (10mM Tris:HCl, pH8.0, 1mM EDTA) 稀释至 120μl。通过加入 50μl 的 8M 醋酸铵和 160μl 的异丙醇沉淀 cDNA 两次。产生的 cDNA 沉淀带重新悬浮于 10μl 的 TE 中。由 ³²p - dCTP 渗入的量估计每个反应第一链 cDNA 的产量。

用单独的聚合酶链反应, 从肝、肺和肾 mRNA 合成的第一链 cDNA 用于产生 2 个 cDNA 区段, N-端三分之一序列和 C-端三分之二序列。使用引物 ZC7422 (SEQ ID No:20) 和 ZC7423 (SEQ ID No:21) 通过 PCR 突变从基因组序列的一个单独碱基的变化把一个 Kpn 限制性位点引入到 cDNA 区段中。产生的核苷酸变化创造了一个普通的 KpnI 限制性位点, 而不改变预计的氨基酸编码能力。

在含有 5ng 的模板 cDNA (对肾、肝和肺 cDNA 用不同的反应), 各 80pmol 的寡聚核苷酸 ZC7424 (SEQ ID No:22) 和 ZC7422 (SEQ ID No:20), 5μl 的 2.5mM 的脱氧核苷酸三磷酸溶液 (Cetus Corp., Emeryville, CA), 5μl 的 10x PCR 缓冲液 (Promega Corp., Madison, WI) 和 2.5 单位的 Taq 聚合酶 (Biehringer Mannheim) 的 50μl 的反应物中扩增 N-端片段。聚合酶链反应进行 35 个循环 (94 ℃ 1 分钟, 58 ℃ 1 分钟和 72 ℃

1.5 分钟),紧接着于 72 ℃ 下温育 7 分钟。有意引物 ZC7424 (SEQ ID No:22) 跨越小鼠 mpl 受体的配体 5'非翻译区域,包括 ATG 起始密码子。反义引物 ZC7422 (SEQ ID No:20) 包括对应于人的基因组 TPO DNA 的第 4 和第 5 个外显子区域的序列。

在含有 5ng 的模板 cDNA (如上所述的人的肾、肝或肺)。各 80pmol 的寡聚核苷酸 ZC7423 (SEQ ID No:21)和 ZC7421 (SEQ ID No:23), 5 μ l 的 2.5mM 脱氧核苷酸三磷酸溶液 (Cetus Corp.),5 μ l 的 10x PCR 缓冲液 (Promega Corp.) 和 2.5 单位的 Taq 聚合酶 (Boehringer Mannheim) 的 50 μ l 反应物中扩增 C - 端片段。聚合酶链反应进行 35 个循环,再于 37 ℃ 下温育 7 分钟。有意引物包括对应于人的基因组 TPO DNA 的第 4 和第 5 个外显子区域序列,反义引物 ZC7423 (SEQ ID No:21) 包括对应于人的基因 3'非编码序列的区域序列,且包括翻译终止密码子。

扩增的 PCR 产物直接进行 DNA 序列分析和亚克隆至 pGEM-T (Promega Corp.) 中,通过与小鼠 cDNA 序列和人的基因组序列比较进一步分析。编码人的 TPO DNA 序列表示于 SEQ ID No:18 中,其编码的氨基酸序列如 SEQ ID No:19 所示,序列分析表明信号肽切割位点位于第 22 个氨基酸 (SEQ ID No:19),成熟蛋白质从第 22 个氨基酸开始 (SEQ ID No:19)。

从 pGEM - T 上切下人的 N - 端和 C - 端 PCR 产物片段 EcoRI-KpnI,并连接于表达载体 Zem229R 的 EcoRI 位点。用 LipofectaineTM (GIBCO BRL) 把该质粒转染至 BHK 570 细胞中。转染 24 小时之后,用新鲜培养基更换培养基 (DMEM + PSV + 10 % FCS),细胞在缺乏选择因子时温育 48 小时。如以前所描述的,用 BaF₃/MPLR1.1 细胞系对条件培养基检测增生活性。结果清楚地表明培养基中人的 TPO 可刺激表达鼠 MPL 受体的 BaF₃ 细胞增生。

按照厂商的说明,用 SUPERSRIPTTM 逆转录酶 (GIBCO BRL) 从人的肝和肾 mRNA (从 Clontech Laboratories, Inc. 获得) 制备 cDNA。然后用两个 PCR 反应 (条件见表 6) 构建肝一和肾

一来源的人 TPO cDNA 克隆。反应于 94 °C 下 1 分钟， 58 °C 下 1 分钟， 72 °C 下 1.5 分钟进行 35 个循环；再于 72 °C 下温育 7 分钟。

表 6

反应 # 1:

5ng 肝或肾 cDNA

4 μ l 寡聚核苷酸 ZC7454 (20pM/ μ l) (SEQ ID No:24, 在 ATG 的 5' 引入一个 EcoRI 位点)

4 μ l 寡聚核苷酸 ZC7422 (20pM/ μ l) (SEQ ID No:20; 创造一个 Asp714 位点)

5 μ l 含有 2.5mM dATP, 2.5mM dGTP, 2.5mM dCTP 和 2.5mM dTTP 的 dNTP_s 溶液。

5 μ l 10x 的 Taq 缓冲液 (Boehringer Mannheim)

1 μ l Taq 聚合酶 (Boehringer Mannheim)

30 μ l 的水

反应 # 2:

5ng 肝或肾 cDNA

4 μ l 寡聚核苷酸 ZC 7423 (20pM/ μ l) (SEQ ID No:20, 创造一个 Asp718 位点)

4 μ l 寡聚核苷酸 ZC7453 (20pM/ μ l) (SEQ ID No:25; 在 TGA 的 3' 端引入一个 EcoRI 位点)

5 μ l 含有 2.5mM dATP, 2.5mM dGTP, 2.5mM dCTP 和 2.5mM dTTP 的 dNTP_s 溶液

5 μ l 10x 的 Taq 缓冲液 (Boehringer Mannheim)

1 μ l Taq 聚合酶 (Boehringer Mannheim)

30 μ l 的水

PCR 产物用苯酚/氯仿/异戊醇处理并用 95 % 乙醇沉淀，干燥和重新悬浮于 20 μ l 的水中。每个产物然后用限制性酶 Asp718 和 EcoRI 切割，于 1 % 的琼脂糖凝胶上电泳。从胶中切下反应 # 1 的 410bp 片段 (肝和肾) 和 699bp 的片段 (肝和肾)，通过离心厚的胶块，透过尼龙羊毛膜洗脱 DNA 片段。反应 # 1 和反应 # 2 的 PCR

产物一起用 EcoRI 切割的载体 Zem229R (贮藏于 American Type Culture Collection, 12301 Park lawn Drive, Rockville, MD, 1993 年 9 月 28 日, 保藏号 69447), 从而在创造的 Asp718 位点把两种产物连接起来。产生的质粒指定为 # 10 (含肾来源的 cDNA) 和 # 28 (含肝来源的 cDNA)。

经过对 DNA_s 序列分析, 发现单一 PCR 反应产生的误差分别位于 # 28 和 # 10 质粒中唯一的 AvrII 位点的 5' 和 3'。为了创造一个没有错误的 TPO DNA, 从 # 10 质粒中分离一个 826 bp 的 EcoRI - AvrII 5' 片段, 从 # 28 质粒中分离一个 283bp 的 AvrII - EcoRI 3' 片段。两个片段与用 EcoRI 切割的 Zem229R 载体一起连接。产生的质粒被指定为 pZGmpl - 124。该质粒于 1994 年 5 月 4 日保存于美国悬典型培养物保藏中心 (American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD) 作为大肠杆菌 DH10b 转化体, 保藏号为 69615。

实施例 XI. 巨核细胞 cDNA 文库

为了扩增活体内巨核细胞前体, 20 只小鼠每天通过腹膜注射 40,000 活性单位 (在 MTT 试验中 (实施例 VII), 50U/ml 可使 BaF₃/MPLR1.1 细胞获得一半的最大增生率) 的重组小鼠血小板生成素 (经小鼠血小板 cDNA 稳定转染的 BHK570 细胞的浓缩的无血清条件培养基)。在注射的第 5 天, 取出脾脏, 置于 CATCH 缓冲液 + HEPES (Hank's 平衡盐溶液 CHBSS), 无钙和镁, 10mM HEPES (GIBCO BRL), 1.4mM 腺苷, 2.74mM 的茶碱 (Sigma Chemical Corp., St. Louis, MO) 和 0.38% 的柠檬酸钠 (J.T.Baker Inc. Philipsburg, NJ) 用 NaOH 调整 pH 至 7.40) 中。每次处理 5 个脾脏, 切开每个脾脏, 在两个不锈钢切口处挤出细胞, 置于 CATCH 缓冲液 + HEPES 中。用一个 25ml 的吸管分散细胞凝块后, 把体积增加至 50ml, 细胞于 Sorval (TJ - 6 离心机中以 208g 离心 7 分钟。每个细胞沉淀重新悬浮于 10ml 的 CATCH 缓冲液 + HEPES 中, 通过 130 μ m 的尼龙网过滤, 获得单细胞悬浮物。用 CATCH 缓冲液 + HEPES 使体积增加至 50ml, 细胞于 33g 下离心

15 分钟。细胞用另外的 50ml 的 CATCH + Hepes 洗涤，于 33g 下离心 10 分钟。细胞沉淀物重新悬浮于 10ml 的 CATCH 缓冲液 + Hepes 中，在离心管中置于三级 Percoll (Pharmacia LKB Biotechpology AB, Uppsala, Sweden) 梯度 (65、40.27 % 的 1 × CATCH 缓冲液 + Hepes，每 50ml 的离心管中加入 12ml) 于 833g 下离心 45 分钟。收集 40 % 和 63 % Percoll 层之间的细胞，用 CATCH 缓冲液 + Hepes 使其体积增加至 50ml。细胞于 208g 下离心 7 分钟，并重新悬浮于 50ml 的巨核细胞生长培养基 (最低必需培养基 α 修饰物，无核苷酸和脱氧核苷酸的 15 % 热钝化小牛胎儿血清，2mM L-谷氨酸 (从 JRH Biosciences, Lenexa, CA 获得的培养基成分)，1mM 丙酮酸盐 (Irvine Scientific, Santa Ana, CA)，1 × PSN 抗生素混合物 (GIBCO BRL) 中并加入每毫升 1000 活性单位的重组小鼠血小板生成素 (来源于小鼠血小板 cDNA 稳定转染的 BHK 细胞的无血清条件培养基)。然后以每毫升 10^6 个单成核细胞涂布于 150mm 组织培养皿中，在 6.0 % CO_2 、37 °C 下于充分湿润的培养箱中生长。生长 3 天之后，把外粘着细胞收集于 50ml 的离心管中，置于冰上冷却。通过 4 °C 下以 33g 离心 15 分钟沉淀较大的细胞。细胞于室温下重新悬浮于 50ml 的 CATCH + Hepes 中，并于 33g 下离心 10 分钟 (所有进一步的操作步骤均在室温下进行)。此洗涤步骤再重复一次以获得高纯度的成熟巨核细胞。剩余的细胞重新悬浮于 15ml 的 CATCH + Hepes (收集的体积) 并置于小牛胎儿血清三级梯度上 (JRH Biosciences) (用 CATCH 缓冲液 + Hepes 稀释的 65 % 和 40 % 的梯度) 以 1g 离心 30 分钟沉淀剩余细胞。收集底部 5ml 的 65 % 的部分，用 CATCH 缓冲液 + Hepes 稀释至 50ml，并以 33g 离心 10 分钟，按 Burstein 等的方法 (J.Cell.Physiol. 122: 159-165,1985) 测定细胞的乙酰胆碱酯酶，并确定纯度大于 99 % 的成熟巨核细胞。沉淀的细胞用硫氰酸胍 / 2 - 巯基乙醇裂解，通过氯化铯密度梯度离心提取 RNA。如以上实施例 VI 中描述的方法，从巨核细胞 RNA 制备 cDNA。

实施例 XII. 人血小板生成素基因的荧光原位杂交的基因作图

把下列混合物加入到置于冰上的微量离心管中：含有人血小板生成素基因的 λ ZGmpl-H8, λ ZGmpl-H10 或 λ ZGmpl-H29, 5 μ l 的缺口平移缓冲液 (Nick Translation buffer, 0.5M Tris/HCl, 50mM MgCl₂, 0.5mg/ml BSA) (无核酸酶), 含有 0.5mM dATP、0.5mM dGTP 和 0.5mM dCTP 的 5 μ l 的 dNTP_s 溶液, 5 μ l 的 5mM 的 Bio-11-dUTP (5 - NIN - 生物素 - ϵ - 氨基己酰基] - 3 - 氨基烯丙基 - 2' - 脱氧尿苷 - 5' - 三磷酸, Sigma Chemical Co.), 5 μ l 的 100mM 的 DTT, 5 μ l 的 DNaseI (从 10U/ μ l 的贮存液中稀释 1000 α , Boehringer Mannheim, 无 Rnase), 2.5 μ l 的 DNA 聚合酶 I (5U/ μ l Boehringer Mannheim) 和水至终体积 50 μ l。混合之后, 于 Bockel 微型低温培养箱中于 15 $^{\circ}$ C 下温育 2 小时, 通过向反应液中加入 5 μ l 的 0.5M EDTA, pH7.4 终止反应。按照厂商的说明, 用 Sephadex[®] G-50 DNA 纯化旋转柱 (Worthington Biochemical Corporation, Freehold, NJ) 纯化探针。为了检查探针的大小, 5 - 10 μ l 的每种纯化的探针与 5 μ l 含凝胶缓冲液 (12.5% 聚蔗糖 (ficoll)、0.2% 的溴酚蓝, 0.2M Tris-乙酸, 0.1M 醋酸钠, 1mM EDTA), 于 0.7% 的琼脂糖小型凝胶上以 80v 电压电泳。用 λ - HindIV 片段 (GIBCO BRL) 和 ϕ X - Hae 片段 (GIBCO BRL) 作为碱基对 (bp) 大小标记。从 Oncor (Gaithersburg, MD) 获得一个地高辛标记的 (digoxigenin - labeled) 的染色体 3 (D3Z1) 特定的着丝粒探针。

从 HEL 细胞培养中获得中期染色体。加入 100 μ l 的 Colcemid[®] (GIBCO BRL, 100 μ g/ml 贮存液) 到用于细胞培养的 100 \times 15mm 平皿的培养基中, 于 37 $^{\circ}$ C 下温育。2.5 - 3 小时之后, 用 10ml 的无菌塑料吸管从平皿中移去培养基, 并转移到 15ml 的聚丙烯圆锥形管 (Blue MaxTM, Becton Dickinson) 中。用 5ml 的无菌塑料吸管加入 2ml 的 1 \times PBS (140mM NaCl, 3mM KCl, 8mM Na₂HPO₄, 1.5mM KH₂PO₄, pH7.2) 至平皿中冲洗, 并转移到圆锥形管中。用无菌的 5ml 的塑料吸管加入 2ml 的胰蛋白酶 (GIBCO BRL, 贮存液) 至平皿中, 轻轻摇动平皿, 置于 37 $^{\circ}$ C 的培养箱中

3 - 5 分钟。用 5ml 的无菌塑料吸管冲洗皿中的细胞，并把细胞加入至带有培养基的试管中。培养试管于 $250 \times g$ 下离心 8 分钟，取出几乎为 0.5ml 的上清液。通过轻轻拍打离心管重新悬浮沉淀，然后缓慢和轻轻地加入 0.075M KCl（预热至 $37\text{ }^{\circ}\text{C}$ ）。轻轻混合悬浮液并置于 $37\text{ }^{\circ}\text{C}$ 水浴中 10 分钟。以 $250 \times g$ 离心溶液 5 分钟，取出沉淀上面的大约 0.5ml 的上清液。轻轻拍打离心管重新悬浮沉淀。以点注方式加入 2ml 冷的甲醇：乙酸（3:1），摇动试管以固定细胞。以这种方式共加入 8ml 的固定液。试管置于冰箱中 20 分钟，接着以 $250 \times g$ 离心 5 分钟。再吸出上清，并重复固定过程二次以上。把中期分散的染色体点滴于 $25 \times 75\text{mm}$ 预先洗净的冰冻的载玻片上（VMR Scientific, Media, PA），用 $20\mu\text{l}$ 的 PipenmanTM（Gilson Medical Electronics, Inc., Middleton, WI）滴加 $5\mu\text{l}$ 的 50% 的乙酸于每个载玻片上，再加 $5\mu\text{l}$ 的细胞悬乳液。载玻片于空气中干燥，使用之前再于 $42\text{ }^{\circ}\text{C}$ 的烘箱（Boekel Industries, Inc., Philadelphia, PA）中过夜老化，用配有相衬浓缩装置的显微镜对合适的中期分散的染色体进行计数。某些中期染色体制备物用 Gurr's 改进的 R₆₆ Giemsa's 染料（BDH Ltd.）Dorset, England）染色呈现 G - 带型，对这些染色体照相，用于杂交实验之前再脱色。用人的中期染色体扩散物制备的载玻片于 $2 \times \text{SSC}$ （ 0.3M NaCl , 0.03M 柠檬酸钠， $\text{pH}7.0$ ）中温育 2 小时，简单地用水冲洗，用 Giemsa's 缓冲液， $\text{pH}6.5$ （BHK Ltd.）以 1:4 稀释的 Gurr's Giemsa's 染料染色（染料使用之前经 Whatman#1 滤纸过滤）。某些制备的载玻片先于 $90\text{ }^{\circ}\text{C}$ 的烘箱中温育 45 分钟至 1 小时，在 SSC 中温育之前冷却。这些制备的载玻片于 Giemsa's 缓冲液中使染色体变异，用水冲洗载玻片并于空气中干燥。用 Kodak EktachromeTM 400 幻灯片胶卷，在 Olympus 显微镜下对合适的呈 G - 带型的中期染色体扩散物照相，并用 Optronics（Goleta, CA）ZVS - 47E CCD RGB 彩色视象系统和 Optimus 软件（来自 Bioscan Inc., Edmonds, WA）使其数字化并贮存起来。进一步使用之前，制备的载玻片于 100% 乙醇中脱色 20 分钟，空气干燥。未用的中期染

染色体载玻片制备物保存于 $-70\text{ }^{\circ}\text{C}$ 下。

经在 1.5ml 的无菌微量离心管中组合 $2.5\mu\text{g}$ 的竞争 DNA (Cot-1DNA, GIBCO BRL), $40 - 60\text{ng}$ 的生物素标记的 $\lambda\text{ZGmpl} - \text{H8}$, $\lambda\text{ZGmpl} - \text{H10}$ 或 $\lambda\text{ZGmpl} - \text{H29}$ 噬菌体 (含人血小板生成素基因), $7\mu\text{g}$ 的载体 DNA (变性的鲑鱼精巢 DNA, Sigma Chemical Co.), 1ml 3M 的醋酸钠和 2 倍体积的乙醇制备杂交混合物, 并在真空离心蒸发浓缩器中真空干燥, 将离心沉淀溶于由葡萄糖硫酸盐, $2 \times \text{SSC}$ 和 50% 的甲酰胺 (EM Science, Gibbstwn, NJ) 组成的 $10\mu\text{l}$ 的杂交液中。使探针和竞争 DNA 于 $70\text{ }^{\circ}\text{C} - 80\text{ }^{\circ}\text{C}$ 下变性 5 分钟, 置于冰上冷却, 并于 $37\text{ }^{\circ}\text{C}$ 下预退火 1 - 2 小时。通过把每个载玻片浸入 70% 甲酰胺, $2 \times \text{SSC}$ 中于 $70 - 80\text{ }^{\circ}\text{C}$ 下 5 分钟, 接着于冰冷的 70% 乙醇中立即冷却, 然后置于 100% 乙醇中 5 至 10 分钟使染色体变性。这些载玻片再经空气干燥, 在用 $20\mu\text{l}$ 的 Gilson PipetmanTM 加入杂交混合物至载玻片之前把载玻片加热至 $42\text{ }^{\circ}\text{C}$ 。杂交混合物和染色体再用 $18 \times 18\text{mm}$ 的 1 号外套 (VWR Scientific) 罩上。于润湿的培养箱中在 $37\text{ }^{\circ}\text{C}$ 下进行杂交过夜。在某些情况下, 在杂交约 6 小时之后, 向制备物中加入 $5 - 10\text{ng}$ 的变性的地高辛标记的 D3Z1 着丝粒探针 (于 10% 的葡萄糖硫酸盐, $2 \times \text{SSC}$ 和 65% 的甲酰胺杂交液中)。

揭开外套之后, 载玻片于 50% 甲酰胺、 $2 \times \text{SSC}$ 在 $42\text{ }^{\circ}\text{C}$ 中洗涤 3×5 分钟, 于 $2 \times \text{SSC}$ 在 $42\text{ }^{\circ}\text{C}$ 下洗涤 3×5 分钟和于 $4 \times \text{SSC}$, 0.05% 多氧乙烯山梨糖醇单月桂酸 (Tween-20, Sigma Chemical Co.) 中洗涤 1×3 分钟。紧接着于湿润的培养箱中用含有 5% 脱脂干奶粉的 $4 \times \text{SSC}$ ($100\mu\text{l}$, 于 $24 \times 50\text{mm}$ 外套下) 预培养 20 分钟。对包括染色体 3 D3Z1 着丝粒探针在内的制备物, 再于 $4 \times \text{SSC}/5\%$ BSA 中与 $1:100$ 稀释的生物素一标记的小鼠抗地高辛 (Sigma Chemical Co.) 一起温育 45 分钟, 接着于 $4 \times \text{SSC}$, 0.05% Tween - 20 中每次 3 分钟共洗涤 3 次。于 $24 \times 50\text{mm}$ 外套下, 用荧光素标记的禽抗生物素蛋白 (荧光素抗生物素蛋白 DCS, Vector Laboratories, Burlingame, CA) ($100\mu\text{l}$, $5\mu\text{g}/\text{ml}$,

于 $4 \times \text{SSC}$, 5% 脱脂干奶粉中) 对所有的制备物进行杂交后的 20 分钟的温育。然后将载玻片于 $4 \times \text{SSC}$, 0.05% Tween - 20 中洗涤 3×3 分钟, 接着在 $24 \times 50\text{mm}$ 外罩下, 用生物素化的羊抗 - 抗生物素蛋白 D (亲和提纯, Vector Laboratories) ($5\mu\text{g/ml}$ 于 $4 \times \text{SSC}$, 5% 脱脂干奶粉) 再温育 20 分钟。载玻片再于 $4 \times \text{SSC}$, 0.05% Tween20 洗涤 3×3 分钟, 再于 $24 \times 50\text{mm}$ 外罩下用荧光素一标记的抗生物素蛋白 ($100\mu\text{l/ml}$ 于 $4 \times \text{SSC}$, 5% 脱脂干奶粉) 温育。在某些情况下, 信号扩增步骤可再重复一次。最后于 $4 \times \text{SSC}$, 0.05% Tween - 20 中洗涤 2×3 分钟, 于 $1 \times \text{PBS}$ 中洗涤 1×3 分钟。把载玻片固定于 antifade 培养基中, 该培养基由 9 份含有 2% 1,4 - 重氨基双环 - (2,2,2) - 辛烷 (DABCO , 于 70°C 下溶解) 的甘油和 1 份 0.2M Tris/HCl, pH7.5 和 0.25-0.5 $\mu\text{g/ml}$ Propidium 碘化物组成。在配有 BH2 - RFC 反射光荧光的 Olympus BH2 显微镜 (PM - 10 ADS 自动显微照相系统, Optronics ZVS-47E CCD RGB 彩色摄象系统和 Chroma 技术公司 (Brattlebow, VT) FITC/Texas 为观察 FITC 的红色滤光装置) 下观看载玻片。用 Optronics 视象照相系统和 Optimus 软件使中期染色体扩散物的图象数字化并存贮起来。

从物理图谱定位程序的初步结果表明人血小板基因座位于 3q26 区域的染色体 3q 臂的远端。

实施例 XIII. 小鼠 TPO 细胞因子结构域在啤酒糖酵母中的表达

质粒 pBJ3-5 含有啤酒糖酵母 TPI1 启动子, α - 因子分泌前导区, 从 237bp 至 692bp 的小鼠 TPO 编码序列, TPI1 转录终止子, 酵母中复制的 2M 序列和酵母中选择的粟酒裂殖糖酵母丙糖磷酸酯异构酶基因 (POT1 基因)。该质粒的设计是为了指导分泌含 SEQ ID No:2 的氨基酸 45 - 196 的小鼠 TPO 蛋白。

为了构建 pBJ3-5, 用 SphI 和 XbaI 消化 pMVR1 (图 2), 回收含有 TPI1 启动子 5' 部分和 PTI1 终止子的载体主要部分。然后将以下片段插入到载体的主要部分中。

1) 一个来源于 pBS114 的 SphI/HindIV 片段, 该片段含有 TPI1

的 3' 部分和 α - 因子前导区。质粒 pBS114 为酵母穿梭质粒，含有 TPI1 启动子和 α - 因子，紧接着为一个包含 HindIV 位点的多接头序列。

2) 一个 PCR 产生的 Hind III/SalI 片段，该片段含有一个设计的与 α - 因子前导区中 Hind IV 位点处于同一框架内的 Hind IV 位点，一个 Ke_x2 蛋白水解切割位点和 SEQ ID No:1 中的从 237bp 至 335 个 bp 的小鼠 TPO 序列。

3) 一个 SalI/EcoRI 片段，该片段含有 SEQ ID No:1 中的小鼠 TPO336 个碱基对到 692 个碱基对，来源于质粒 pSL - MPL - 100 (通过用引物 ZC7319 (SEQ ID No:27) 和 ZC7318 (SEQ ID No:26) 扩增 pZGmpl-1081，用 EcoRI 消化并把含有 TPO 细胞因子结构域序列和 5' 非编码序列的片段克隆至 Zem229R [ATCC 69447] 中)。通过克隆至事先用 SalI 和 EcoRI 消化的 pIC19H 中可以把该片段变化为 SaI/XbaI 片段。

产生的质粒指定为 pBJ3 (图 2)，再用 BglII 和 XhoI 消化，以释放含有启动子，前导区，TPO 编码序列和终止子的整个表达盒。该 BglII/XhoI 片段插入到用 BamHI 和 XhoI 消化的 pRPOT 中 (在美国专利 5,128,321 中已被公开，本文一并参考)，产生的质粒指定为 pBJ3-5。

通过醋酸锂方法 (已由 Ito 等作一般性描述，J.Bacteriol. 153: 163 - 168,1983)，用 pBJ3-5 和 pRPOT 转化啤酒糖酵母菌株 JG134 (MAT α ura3-52 leu2- Δ 2 pep4- Δ 1 Δ tpil: URA3 [Cir⁰])。通过在含有葡萄糖的培养基中生长选择转化体。使 JG134/pBJ3-5 和 JG134/pRPOT 于 YEPD 液体培养基中生长 3 天。通过离心把培养基与细胞分离开，并在含有 MPL 受体的 BaF₃ 细胞中通过细胞增生试验分析培养基。来源于菌株 JG134/pBJ3-5 的培养基含有 5000 - 7000 单位/ml 的 TPO 活性，而负对照 JG134/pRPOT 则没有活性。该结果表明酵母可分泌一种生物活性形式的 TPO。

实施例 XIV 重组人 TPO 的活性

通过用碱裂解细胞，接着再把 DNA 结合到高盐树脂上(用 Promega 公司的 Magic MiniprepsTM Sampler Kit)从用 pZGmpl-124 转化的 2 个 5ml 过夜细菌培养物中制备 DNA。DNA 用 75 μ l 的 10mM Tris, 1mM EDTA, pH8.0 洗提。

用 pZGmpl-124 DNA 转染 50,000 细胞/孔的 BHK570 细胞培养物。在加入到 BHK570 细胞和于 37 $^{\circ}$ C 下温育 4 小时之前，于室温下 30 分钟内把 20 μ l 的 1:10 稀释的 LIPOFECTAMINETM (GIBCO BRL) 加入到 20 μ l 的质粒 DNA 和 160 μ l 的无血清培养基 (F/DV 培养基 [DMEM 和 Ham's 以 1:1 的混合物]，补充有 10 μ g/ml 胎球蛋白，20ng/ml 硒，5 μ g/ml 胰岛素，10 μ g/ml 转铁蛋白，2mM L-谷氨酰胺，110 μ g/ml 的丙酮酸钠，25mM HEPES 和 0.1mM 的非必需氨基酸溶液 [GIBCO BRL]) 中， μ l 然后加入 200 μ l 的生长培养基 (DMEM (Biowhittaker)，该培养基补充有 2mM 的 L-谷氨酰胺，110 μ g/ml 丙酮酸钠，0.05mg/ml 青霉素，0.05mg/ml 链霉素，0.01mg/ml 新霉素，25mM HEPES，10% 小牛胎儿血清)，将细胞于 37 $^{\circ}$ C 下温育过夜。然后用含有 5% 的小牛胎儿血清的生长培养基更换培养基，并于 37 $^{\circ}$ C 下温育 4 小时。

测定 BHK570 转染体的条件培养基在表达小鼠 MPL 受体的 BaF₃ 细胞中引起细胞增生的能力。使细胞于 BaF₃ 培养基中生长 (RPMI 1640 培养基 (JRH Biosciences)，并补充有 10% 小牛胎儿血清，2mM L-谷氨酰胺，1mM 的丙酮酸钠，10mM HEPES, 57 μ M β -巯基乙醇，0.05mg/ml 青霉素，0.05mg/ml 的链霉素，0.01mg/ml 新霉素和 4% V/V 的 WEHI-3 细胞培养物的条件培养基 (小鼠白细胞介素-3，培养基补充物，Collaborative Biochemical Products))。测定之前，BaF₃ 细胞用没有 IL-3 的培养基稀释并重新悬浮至 10,000 细胞/100 μ l。加入 pZGmpl-124 转染 BHK570 细胞的 100 μ l 的条件培养基，于 37 $^{\circ}$ C 下温育培养物。通过视觉观察 30 分钟和 24 小时之后的细胞伸长情况。对由没有 IL-3 的 BaF₃ 培养基组成的负对照和用小鼠 TPO DNA 转染的

BHK570 细胞的条件培养基的正对照也进行测定。结果表明负对照中 BaF₃ 细胞没有伸长，正对照中有些伸长，在 pZGmpl - 124 转染的细胞中细胞有明显的伸长。

实施例 XI. 受体亲和沉淀

将含有产生 TPO 细胞或正常的 BHK 细胞的 150 - mm 组织培养皿用没有甲硫氨酸、含有 2mM L - 谷氨酰胺，抗生素和 Expressr 200 μ Ci 的 ³⁵S (Amersham, Arlington Heights, IL) Dulbecco's MEM 标记 18 小时。

过夜温育之后，收集营养耗尽的培养基，用 Centriprep-10TM (Amicon, Inc.) 离心 15 分钟。产生的 0.75ml 浓缩的上清液与按提供商提供的说明与 CNBr-Sepharose 4B (Pharmacia) 连接的 40 μ l 多聚组氨酸尾序的可溶性受体相混合。混合物置于冰上摇动温育 2 小时。

细胞用 PBS 洗涤一次，然后用 1ml 的 RIPA 缓冲液 (10mM Tris, pH7.4, 1% 脱氧胆酸盐, 1% Triton x-100, 0.1% SDS, 5mM EDTA, 0.15M NaCl) 裂解。裂解物经离心去除不溶解的物质，用如上所述的方法加入 40 μ l 的 MPL - Sepharose 。

然后通过低速离心使 MPL - Sepharose 沉淀，去掉耗尽的培养基和细胞裂解物。沉淀用含 0.5M NaCl 的 PBS 洗涤 4 次。最后一次冲洗之后，去除 PBS，加入含 4% β - 巯基乙醇的 40 μ l 的 2x 样品缓冲液 (10% 甘油, 4% SDS, 5mM Tris, pH7.0, 1mM 0.05% 溴酚蓝) 。

样品煮沸 5 分钟，每个样品 18 μ l 上样至 10 - 20% 的梯度微型凝胶上 (Integrated Separation System)，然后于 100v 下电泳大约 2 小时。固定凝胶 30 分钟 (于 40% 甲醇, 16% 冰醋酸的蒸馏水中)，然后于 AmplifyTM (Amersham) 中浸透 20 分钟。干燥之后，凝胶暴露于胶片下过夜。在与用 TPO cDNA 转染细胞的耗尽营养的培养基相对应的泳道中清晰可见一条 ~70KD 的带。该带不存在于 BHK 细胞的耗尽培养基中，也不是这两种细胞系的细胞裂解物。

从前面所述将意识到的是，虽然为了说明的目的，本文已对本发明的具体实施方案作了描述，但可作各种各样的修改而不偏离本发明的精神和范围。因此，除了所附权利要求书之外，本发明不受限制。

序列表

(1) 一般信息

(i) 申请人: **ZymoGenetics, Inc.**

1201 Eastlake Avenue East
Seattle
WA
US
98102

申请人: **University of Washington**

Seattle
WA
US
98195

(ii) 发明名称: **血细胞生成蛋白及其制造材料和方法**

(iii) 序列数: **29**

(iv) 通信地址:

(A) 通信人: **ZymoGenetics, Inc.**

(B) 街道: **1201 Eastlake Avenue East**

(C) 城市: **Seattle**

(D) 州: **WA**

(E) 国家: **USA**

(F) 邮政编码: **98102**

(v) 计算机可读形式

(A) 介质类型: **Floppy disk**

(B) 计算机: **IBM PC 兼容机**

(C) 操作系统: **PC - DOS/MS - DOS**

(D) 软件: **PatentIn Release #1.0, Version #1.25**

(vi) 目前的申请资料

(A) 申请号:

(B) 申请日:

(C) 分类号:

(viii) 律师/代理人信息

- (A) 姓名: Parker, Gary E
- (B) 注册号: 31 - 648
- (C) 证书/文件号: 93 - 12PC

(ix) 电讯信息

- (A) 电话: 206 - 442 - 6600 ext 6673
- (B) 电传: 206 - 442 - 6678

(2) SEQ ID NO: 1 信息:

(i) 序列特征:

- (A) 长度: 1486 个碱基对
- (B) 类型: 核酸
- (C) 链型: 双链
- (D) 拓扑结构: 线形

(ii) 分子类型: cDNA

(vii) 直接来源:

- (B) 克隆: 1081

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 105..1241

(xi) 序列描述: SEQ ID NO: 1:

CCTCGTGCCG GTCCTGAGGC CTTTCTCCAC CCGGACAGAG TCCTTGGCCC ACCTCTCTCC	60
CACCCGACTC TGCCGAAAGA AGCACAGAAG CTCAAGCCGC CTCC ATG GCC CCA GGA	116
	Met Ala Pro Gly
	1
AAG ATT CAG GGG AGA GGC CCC ATA CAG GGA GCC ACT TCA GTT AGA CAC	164
Lys Ile Gln Gly Arg Gly Pro Ile Gln Gly Ala Thr Ser Val Arg His	
5 10 15 20	
CTG GCC AGA ATG GAG CTG ACT GAT TTG CTC CTG GCG GCC ATG CTT CTT	212
Leu Ala Arg Met Glu Leu Thr Asp Leu Leu Leu Ala Ala Met Leu Leu	
25 30 35	

GCA GTG GCA AGA CTA ACT CTG TCC AGC CCC GTA GCT CCT GCC TGT GAC	260
Ala Val Ala Arg Leu Thr Leu Ser Ser Pro Val Ala Pro Ala Cys Asp	
40 45 50	
CCC AGA CTC CTA AAT AAA CTG CTG CGT GAC TCC CAC CTC CTT CAC AGC	308
Pro Arg Leu Leu Asn Lys Leu Leu Arg Asp Ser His Leu Leu His Ser	
55 60 65	
CGA CTG AGT CAG TGT CCC GAC GTC GAC CCT TTG TCT ATC CCT GTT CTG	356
Arg Leu Ser Gln Cys Pro Asp Val Asp Pro Leu Ser Ile Pro Val Leu	
70 75 80	
CTG CCT GCT GTG GAC TTT AGC CTG GGA GAA TGG AAA ACC CAG ACG GAA	404
Leu Pro Ala Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Thr Glu	
85 90 95 100	
CAG AGC AAG GCA CAG GAC ATT CTA GGG GCA GTG TCC CTT CTA CTG GAG	452
Gln Ser Lys Ala Gln Asp Ile Leu Gly Ala Val Ser Leu Leu Leu Glu	
105 110 115	
GGA GTG ATG GCA GCA CGA GGA CAG TTG GAA CCC TCC TGC CTC TCA TCC	500
Gly Val Met Ala Ala Arg Gly Gln Leu Glu Pro Ser Cys Leu Ser Ser	
120 125 130	
CTC CTG GGA CAG CTT TCT GGG CAG GTT CGC CTC CTC TTG GGG GCC CTG	548
Leu Leu Gly Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu	
135 140 145	
CAG GGC CTC CTA GGA ACC CAG CTT CCT CTA CAG GGC AGG ACC ACA GCT	596
Gln Gly Leu Leu Gly Thr Gln Leu Pro Leu Gln Gly Arg Thr Thr Ala	
150 155 160	
CAC AAG GAC CCC AAT GCC CTC TTC TTG AGC TTG CAA CAA CTG CTT CGG	644
His Lys Asp Pro Asn Ala Leu Phe Leu Ser Leu Gln Gln Leu Leu Arg	
165 170 175 180	
GGA AAG GTG CGC TTC CTG CTT CTG GTA GAA GGT CCC ACC CTC TGT GTC	692
Gly Lys Val Arg Phe Leu Leu Leu Val Glu Gly Pro Thr Leu Cys Val	
185 190 195	
AGA CGG ACC CTG CCA ACC ACA GCT GTC CCA AGC AGT ACT TCT CAA CTC	740
Arg Arg Thr Leu Pro Thr Thr Ala Val Pro Ser Ser Thr Ser Gln Leu	
200 205 210	

CTC	ACA	CTA	AAC	AAG	TTC	CCA	AAC	AGG	ACT	TCT	GGA	TTG	TTG	GAG	ACG	788
Leu	Thr	Leu	Asn	Lys	Phe	Pro	Asn	Arg	Thr	Ser	Gly	Leu	Leu	Glu	Thr	
		215					220					225				
AAC	TTC	AGT	GTC	ACA	GCC	AGA	ACT	GCT	GGC	CCT	GGA	CTT	CTG	AGC	AGG	836
Asn	Phe	Ser	Val	Thr	Ala	Arg	Thr	Ala	Gly	Pro	Gly	Leu	Leu	Ser	Arg	
		230				235					240					
CTT	CAG	GGA	TTC	AGA	GTC	AAG	ATT	ACT	CCT	GGT	CAG	CTA	AAT	CAA	ACC	884
Leu	Gln	Gly	Phe	Arg	Val	Lys	Ile	Thr	Pro	Gly	Gln	Leu	Asn	Gln	Thr	
245					250					255					260	
TCC	AGG	TCC	CCA	GTC	CAA	ATC	TCT	GGA	TAC	CTG	AAC	AGG	ACA	CAC	GGA	932
Ser	Arg	Ser	Pro	Val	Gln	Ile	Ser	Gly	Tyr	Leu	Asn	Arg	Thr	His	Gly	
				265				270						275		
CCT	GTG	AAT	GGA	ACT	CAT	GGG	CTC	TTT	GCT	GGA	ACC	TCA	CTT	CAG	ACC	980
Pro	Val	Asn	Gly	Thr	His	Gly	Leu	Phe	Ala	Gly	Thr	Ser	Leu	Gln	Thr	
			280					285						290		
CTG	GAA	GCC	TCA	GAC	ATC	TCG	CCC	GGA	GCT	TTC	AAC	AAA	GGC	TCC	CTG	1028
Leu	Glu	Ala	Ser	Asp	Ile	Ser	Pro	Gly	Ala	Phe	Asn	Lys	Gly	Ser	Leu	
		295					300					305				
GCA	TTC	AAC	CTC	CAG	GGT	GGA	CTT	CCT	CCT	TCT	CCA	AGC	CTT	GCT	CCT	1076
Ala	Phe	Asn	Leu	Gln	Gly	Gly	Leu	Pro	Pro	Ser	Pro	Ser	Leu	Ala	Pro	
		310				315					320					
GAT	GGA	CAC	ACA	CCC	TTC	CCT	CCT	TCA	CCT	GCC	TTG	CCC	ACC	ACC	CAT	1124
Asp	Gly	His	Thr	Pro	Phe	Pro	Pro	Ser	Pro	Ala	Leu	Pro	Thr	Thr	His	
325					330					335					340	
GGA	TCT	CCA	CCC	CAG	CTC	CAC	CCC	CTG	TTT	CCT	GAC	CCT	TCC	ACC	ACC	1172
Gly	Ser	Pro	Pro	Gln	Leu	His	Pro	Leu	Phe	Pro	Asp	Pro	Ser	Thr	Thr	
				345					350					355		
ATG	CCT	AAC	TCT	ACC	GCC	CCT	CAT	CCA	GTC	ACA	ATG	TAC	CCT	CAT	CCC	1220
Met	Pro	Asn	Ser	Thr	Ala	Pro	His	Pro	Val	Thr	Met	Tyr	Pro	His	Pro	
			360					365					370			
AGG	AAT	TTG	TCT	CAG	GAA	ACA	TAGCGCGGGC	ACTGGCCCAG	TGAGCGTCTG							1271
Arg	Asn	Leu	Ser	Gln	Glu	Thr										
		375														

CAGCTTCTCT CGGGGACAAG CTTCCCCAGG AAGGCTGAGA GGCAGCTGCA TCTGCTCCAG 1331
 ATGTTCTGCT TTCACCTAAA AGGCCCTGGG GAAGGGATAC ACAGCACTGG AGATTGTAAA 1391
 ATTTTAGGAG CTATTTTTTT TTAACCTATC AGCAATATTC ATCAGAGCAG CTAGCGATCT 1451
 TTGGTCTATT TTCGGTATAA ATTTGAAAAT CACTA 1486

(2) SEQ ID NO : 2 信息

(i) 序列特征:

(A) 长度: 379 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线形

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO : 2 :

Met Ala Pro Gly Lys Ile Gln Gly Arg Gly Pro Ile Gln Gly Ala Thr
 1 5 10 15
 Ser Val Arg His Leu Ala Arg Met Glu Leu Thr Asp Leu Leu Leu Ala
 20 25 30
 Ala Met Leu Leu Ala Val Ala Arg Leu Thr Leu Ser Ser Pro Val Ala
 35 40 45
 Pro Ala Cys Asp Pro Arg Leu Leu Asn Lys Leu Leu Arg Asp Ser His
 50 55 60
 Leu Leu His Ser Arg Leu Ser Gln Cys Pro Asp Val Asp Pro Leu Ser
 65 70 75 80
 Ile Pro Val Leu Leu Pro Ala Val Asp Phe Ser Leu Gly Glu Trp Lys
 85 90 95
 Thr Gln Thr Glu Gln Ser Lys Ala Gln Asp Ile Leu Gly Ala Val Ser
 100 105 110
 Leu Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln Leu Glu Pro Ser
 115 120 125

Cys Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln Val Arg Leu Leu
 130 135 140

Leu Gly Ala Leu Gln Gly Leu Leu Gly Thr Gln Leu Pro Leu Gln Gly
 145 150 155 160

Arg Thr Thr Ala His Lys Asp Pro Asn Ala Leu Phe Leu Ser Leu Gln
 165 170 175

Gln Leu Leu Arg Gly Lys Val Arg Phe Leu Leu Leu Val Glu Gly Pro
 180 185 190

Thr Leu Cys Val Arg Arg Thr Leu Pro Thr Thr Ala Val Pro Ser Ser
 195 200 205

Thr Ser Gln Leu Leu Thr Leu Asn Lys Phe Pro Asn Arg Thr Ser Gly
 210 215 220

Leu Leu Glu Thr Asn Phe Ser Val Thr Ala Arg Thr Ala Gly Pro Gly
 225 230 235 240

Leu Leu Ser Arg Leu Gln Gly Phe Arg Val Lys Ile Thr Pro Gly Gln
 245 250 255

Leu Asn Gln Thr Ser Arg Ser Pro Val Gln Ile Ser Gly Tyr Leu Asn
 260 265 270

Arg Thr His Gly Pro Val Asn Gly Thr His Gly Leu Phe Ala Gly Thr
 275 280 285

Ser Leu Gln Thr Leu Glu Ala Ser Asp Ile Ser Pro Gly Ala Phe Asn
 290 295 300

Lys Gly Ser Leu Ala Phe Asn Leu Gln Gly Gly Leu Pro Pro Ser Pro
 305 310 315 320

Ser Leu Ala Pro Asp Gly His Thr Pro Phe Pro Pro Ser Pro Ala Leu
 325 330 335

Pro Thr Thr His Gly Ser Pro Pro Gln Leu His Pro Leu Phe Pro Asp
 340 345 350

Pro Ser Thr Thr Met Pro Asn Ser Thr Ala Pro His Pro Val Thr Met
 355 360 365

Tyr Pro His Pro Arg Asn Leu Ser Gln Glu Thr
370 375

(2) SEQ ID NO : 3 信息:

(i) 序列特征:

(A) 长度: 42 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC5499

(xi) 序列描述 SEQ ID NO : 3

CGAGCCACTT^UTCTGCACTCC TCGAGTTTTT TTTTTTTTTT TT

(2) SEQ ID NO : 4 信息:

(i) 序列特征:

(A) 长度: 45 个碱基对:

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC5746

(xi) 序列描述: SEQ ID NO : 4

GAGAGAGAGA GAGAATTCAT GCCCTCCTGG GCCCTCTTCA TGGTC

(2) SEQ ID NO : 5 信息:

(i) 序列特征

(A) 长度: 52 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

- (D) 拓扑结构: 线形
- (vii) 直接来源:
 - (B) 克隆: ZC5762
- (xi) 序列描述: SEQ ID NO: 5:

AGAGAGAGAG AGAGCTCGAG TCAAGGCTGC TGCCAATAGC TTAGTGGTAG GT

52

(2) SEQ ID NO: 6 信息:

- (i) 序列特征
 - (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑结构: 线形
- (vii) 直接来源
 - (B) 克隆: ZC5742
- (xi) 序列描述: SEQ ID NO: 6:

GACCCTGGAG CTGCGCCCGC GATCTCGCTA

(2) SEQ ID NO: 7 信息:

- (i) 序列特征:
 - (A) 长度: 49 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑结构: 线形
- (vii) 直接来源:
 - (B) 克隆: ZC6091
- (xi) 序列描述: SEQ ID NO: 7:

GAGCACAGAA TTCACTACTC GAGGCGGCCG CTTTTTTTTT TTTTTTTT

(2) SEQ ID NO : 11 信息:

(i) 序列特征:

(A) 长度: 60 个碱基对

(B) 类型: 核苷酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC6703

(xi) 序列描述: SEQ ID NO : 11 :

CGACTTTACC TCGAGTGCTA CTGATGCTCT TCTGCCAGCA GTCTCGGAGC CCGTGGACAC 60

(2) SEQ ID NO : 12 信息:

(i) 序列特征:

(A) 长度: 42 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC6707

(xi) 序列描述: SEQ ID NO : 12 :

AATTCGCCAT GGGACTCGAG CATCACCATC ACCATCACTG AG 42

(2) SEQ ID NO : 13 信息:

(i) 序列特征:

(A) 长度: 42 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC6706

(xi) 序列描述: SEQ ID NO : 13 :

GATCCTCAGT GATGGTGATG GTGATGCTCG AGTCCCATGG CG

42

(2) SEQ ID NO : 14 信息:

(i) 序列特征:

(A) 长度: 47 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC6172

(xi) 序列描述: SEQ ID NO : 14 :

GTCGGTGCTC AGCATTCACT ACTCGAGGGT TTTTTTTTTT TTTTTT

47

(2) SEQ ID NO : 15 信息:

(i) 序列特征:

(A) 长度: 28 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC6936

(xi) 序列描述: SEQ ID NO : 15 :

AATTGGCGGC CGCGTCGACT CGTGGATG

28

(2) SEQ ID NO : 16 信息:

(i) 序列特征:

(A) 长度: 28 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC6937

(xi) 序列描述: SEQ ID NO : 16 :

AATTCATCCA CGAGTCGACG CGGCCGCC

28

(2) SEQ ID NO: 17 信息:

(i) 序列特征:

(A) 长度: 633 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线形

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 17:

Met Pro Ser Trp Ala Leu Phe Met Val Thr Ser Cys Leu Leu Leu Ala
1 5 10 15

Leu Pro Asn Gln Ala Gln Val Thr Ser Gln Asp Val Phe Leu Leu Ala
20 25 30

Leu Gly Thr Glu Pro Leu Asn Cys Phe Ser Gln Thr Phe Glu Asp Leu
35 40 45

Thr Cys Phe Trp Asp Glu Glu Glu Ala Ala Pro Ser Gly Thr Tyr Gln
50 55 60

Leu Leu Tyr Ala Tyr Arg Gly Glu Lys Pro Arg Ala Cys Pro Leu Tyr
65 70 75 80

Ser Gln Ser Val Pro Thr Phe Gly Thr Arg Tyr Val Cys Gln Phe Pro
85 90 95

Ala Gln Asp Glu Val Arg Leu Phe Phe Pro Leu His Leu Trp Val Lys
100 105 110

Asn Val Ser Leu Asn Gln Thr Leu Ile Gln Arg Val Leu Phe Val Asp
115 120 125

Ser Val Gly Leu Pro Ala Pro Pro Arg Val Ile Lys Ala Arg Gly Gly
130 135 140

Ser Gln Pro Gly Glu Leu Gln Ile His Trp Glu Ala Pro Ala Pro Glu
145 150 155 160

Ile Ser Asp Phe Leu Arg His Glu Leu Arg Tyr Gly Pro Thr Asp Ser
165 170 175

Ser Asn Ala Thr Ala Pro Ser Val Ile Gln Leu Leu Ser Thr Glu Thr
 180 185 190

Cys Cys Pro Thr Leu Trp Met Pro Asn Pro Val Pro Val Leu Asp Gln
 195 200 205

Pro Pro Cys Val His Pro Thr Ala Ser Gln Pro His Gly Pro Val Arg
 210 215 220

Thr Ser Pro Ala Gly Glu Ala Pro Phe Leu Thr Val Lys Gly Gly Ser
 225 230 235 240

Cys Leu Val Ser Gly Leu Gln Ala Gly Lys Ser Tyr Trp Leu Gln Leu
 245 250 255

Arg Ser Gln Pro Asp Gly Val Ser Leu Arg Gly Ser Trp Gly Pro Trp
 260 265 270

Ser Phe Pro Val Thr Val Asp Leu Pro Gly Asp Ala Val Thr Ile Gly
 275 280 285

Leu Gln Cys Phe Thr Leu Asp Leu Lys Met Val Thr Cys Gln Trp Gln
 290 295 300

Gln Gln Asp Arg Thr Ser Ser Gln Gly Phe Phe Arg His Ser Arg Thr
 305 310 315 320

Arg Cys Cys Pro Thr Asp Arg Asp Pro Thr Trp Glu Lys Cys Glu Glu
 325 330 335

Glu Glu Pro Arg Pro Gly Ser Gln Pro Ala Leu Val Ser Arg Cys His
 340 345 350

Phe Lys Ser Arg Asn Asp Ser Val Ile His Ile Leu Val Glu Val Thr
 355 360 365

Thr Ala Gln Gly Ala Val His Ser Tyr Leu Gly Ser Pro Phe Trp Ile
 370 375 380

His Gln Ala Val Leu Leu Pro Thr Pro Ser Leu His Trp Arg Glu Val
 385 390 395 400

Ser Ser Gly Arg Leu Glu Leu Glu Trp Gln His Gln Ser Ser Trp Ala
 405 410 415

Ala Gln Glu Thr Cys Tyr Gln Leu Arg Tyr Thr Gly Glu Gly Arg Glu
 420 425 430

Asp Trp Lys Val Leu Glu Pro Ser Leu Gly Ala Arg Gly Gly Thr Leu
 435 440 445

Glu Leu Arg Pro Arg Ala Arg Tyr Ser Leu Gln Leu Arg Ala Arg Leu
 450 455 460

Asn Gly Pro Thr Tyr Gln Gly Pro Trp Ser Ala Trp Ser Pro Pro Ala
 465 470 475 480

Arg Val Ser Thr Gly Ser Glu Thr Ala Trp Ile Thr Leu Val Thr Ala
 485 490 495

Leu Leu Leu Val Leu Ser Leu Ser Ala Leu Leu Gly Leu Leu Leu Leu
 500 505 510

Lys Trp Gln Phe Pro Ala His Tyr Arg Arg Leu Arg His Ala Leu Trp
 515 520 525

Pro Ser Leu Pro Asp Leu His Arg Val Leu Gly Gln Tyr Leu Arg Asp
 530 535 540

Thr Ala Ala Leu Ser Pro Ser Lys Ala Thr Val Thr Asp Ser Cys Glu
 545 550 555 560

Glu Val Glu Pro Ser Leu Leu Glu Ile Leu Pro Lys Ser Ser Glu Ser
 565 570 575

Thr Pro Leu Pro Leu Cys Pro Ser Gln Pro Gln Met Asp Tyr Arg Gly
 580 585 590

Leu Gln Pro Cys Leu Arg Thr Met Pro Leu Ser Val Cys Pro Pro Met
 595 600 605

Ala Glu Thr Gly Ser Cys Cys Thr Thr His Ile Ala Asn His Ser Tyr
 610 615 620

Leu Pro Leu Ser Tyr Trp Gln Gln Pro
 625 630

(2) SEQ ID NO: 18 信息:

(i) 序列特征:

(A) 长度: 1062 个碱基对

(B) 类型: 核酸

(C) 链型: 双链

(D) 拓扑结构: 线形

(ii) 分子类型: cDNA

(ix) 特征:

(A) 名称/关键词: CDS

(B) 位置: 1..1059

(xi) 序列描述: SEQ ID NO: 18:

ATG GAG CTG ACT GAA TTG CTC CTC GTG GTC ATG CTT CTC CTA ACT GCA	48
Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala	
1 5 10 15	
AGG CTA ACG CTG TCC AGC CCG GCT CCT CCT GCT TGT GAC CTC CGA GTC	96
Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val	
20 25 30	
CTC AGT AAA CTG CTT CGT GAC TCC CAT GTC CTT CAC AGC AGA CTG AGC	144
Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser	
35 40 45	
CAG TGC CCA GAG GTT CAC CCT TTG CCT ACA CCT GTC CTG CTG CCT GCT	192
Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala	
50 55 60	
GTG GAC TTT AGC TTG GGA GAA TGG AAA ACC CAG ATG GAG GAG ACC AAG	240
Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys	
65 70 75 80	
GCA CAG GAC ATT CTG GGA GCA GTG ACC CTT CTG CTG GAG GGA GTG ATG	288
Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met	
85 90 95	
GCA GCA CGG GGA CAA CTG GGA CCC ACT TGC CTC TCA TCC CTC CTG GGG	336
Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly	
100 105 110	

CAG CTT TCT GGA CAG GTC CGT CTC CTC CTT GGG GCC CTG CAG AGC CTC Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu 115 120 125	384
CTT GGA ACC CAG CTT CCT CCA CAG GGC AGG ACC ACA GCT CAC AAG GAT Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp 130 135 140	432
CCC AAT GCC ATC TTC CTG AGC TTC CAA CAC CTG CTC CGA GGA AAG GTG Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val 145 150 155 160	480
CGT TTC CTG ATG CTT GTA GGA GGG TCC ACC CTC TGC GTC AGG CGG GCC Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala 165 170 175	528
CCA CCC ACC ACA GCT GTC CCC AGC AGA ACC TCT CTA GTC CTC ACA CTG Pro Pro Thr Thr Ala Val Pro Ser Arg Thr Ser Leu Val Leu Thr Leu 180 185 190	576
AAC GAG CTC CCA AAC AGG ACT TCT GGA TTG TTG GAG ACA AAC TTC ACT Asn Glu Leu Pro Asn Arg Thr Ser Gly Leu Leu Glu Thr Asn Phe Thr 195 200 205	624
GCC TCA GCC AGA ACT ACT GGC TCT GGG CTT CTG AAG TGG CAG CAG GGA Ala Ser Ala Arg Thr Thr Gly Ser Gly Leu Leu Lys Trp Gln Gln Gly 210 215 220	672
TTC AGA GCC AAG ATT CCT GGT CTG CTG AAC CAA ACC TCC AGG TCC CTG Phe Arg Ala Lys Ile Pro Gly Leu Leu Asn Gln Thr Ser Arg Ser Leu 225 230 235 240	720
GAC CAA ATC CCC GGA TAC CTG AAC AGG ATA CAC GAA CTC TTG AAT GGA Asp Gln Ile Pro Gly Tyr Leu Asn Arg Ile His Glu Leu Leu Asn Gly 245 250 255	768
ACT CGT GGA CTC TTT CCT GGA CCC TCA CGC AGG ACC CTA GGA GCC CCG Thr Arg Gly Leu Phe Pro Gly Pro Ser Arg Arg Thr Leu Gly Ala Pro 260 265 270	816
GAC ATT TCC TCA GGA ACA TCA GAC ACA GGC TCC CTG CCA CCC AAC CTC Asp Ile Ser Ser Gly Thr Ser Asp Thr Gly Ser Leu Pro Pro Asn Leu 275 280 285	864

CAG CCT GGA TAT TCT CCT TCC CCA ACC CAT CCT CCT ACT GGA CAG TAT	912
Gln Pro Gly Tyr Ser Pro Ser Pro Thr His Pro Pro Thr Gly Gln Tyr	
290 295 300	
ACG CTC TTC CCT CTT CCA CCC ACC TTG CCC ACC CCT GTG GTC CAG CTC	960
Thr Leu Phe Pro Leu Pro Pro Thr Leu Pro Thr Pro Val Val Gln Leu	
305 310 315 320	
CAC CCC CTG CTT CCT GAC CCT TCT GCT CCA ACG CCC ACC CCT ACC AGC	1008
His Pro Leu Leu Pro Asp Pro Ser Ala Pro Thr Pro Thr Pro Thr Ser	
325 330 335	
CCT CTT CTA AAC ACA TCC TAC ACC CAC TCC CAG AAT CTG TCT CAG GAA	1056
Pro Leu Leu Asn Thr Ser Tyr Thr His Ser Gln Asn Leu Ser Gln Glu	
340 345 350	
GGG TAA	1062
Gly	

(2) SEQ ID NO : 19 信息:

(i) 序列特征:

(A) 长度: 353 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线形

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO : 19 :

Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu u Thr Ala	
1 5 10 15	
Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val	
20 25 30	
Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser	
35 40 45	
Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala	
50 55 60	
Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys	
65 70 75 80	

Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met
85 90 95

Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly
100 105 110

Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu
115 120 125

Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp
130 135 140

Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val
145 150 155 160

Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala
165 170 175

Pro Pro Thr Thr Ala Val Pro Ser Arg Thr Ser Leu Val Leu Thr Leu
180 185 190

Asn Glu Leu Pro Asn Arg Thr Ser Gly Leu Leu Glu Thr Asn Phe Thr
195 200 205

Ala Ser Ala Arg Thr Thr Gly Ser Gly Leu Leu Lys Trp Gln Gln Gly
210 215 220

Phe Arg Ala Lys Ile Pro Gly Leu Leu Asn Gln Thr Ser Arg Ser Leu
225 230 235 240

Asp Gln Ile Pro Gly Tyr Leu Asn Arg Ile His Glu Leu Leu Asn Gly
245 250 255

Thr Arg Gly Leu Phe Pro Gly Pro Ser Arg Arg Thr Leu Gly Ala Pro
260 265 270

Asp Ile Ser Ser Gly Thr Ser Asp Thr Gly Ser Leu Pro Pro Asn Leu
275 280 285

Gln Pro Gly Tyr Ser Pro Ser Pro Thr His Pro Pro Thr Gly Gln Tyr
290 295 300

Thr Leu Phe Pro Leu Pro Pro Thr Leu Pro Thr Pro Val Val Gln Leu
305 310 315 320

His Pro Leu Leu Pro Asp Pro Ser Ala Pro Thr Pro Thr Pro Thr Ser
325 330 335

Pro Leu Leu Asn Thr Ser Tyr Thr His Ser Gln Asn Leu Ser Gln Glu
340 345 350

Gly

(2) SEQ ID NO: 20 信息:

(i) 序列特征:

(A) 长度: 23 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7422

(xi) 序列描述: SEQ ID NO: 20:

GGAAGCTGGG TACCAAGGAG GCT

(2) SEQ ID NO: 21 信息:

(i) 序列特征:

(A) 长度: 23 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7423

(xi) 序列描述: SEQ ID NO: 21:

AGCCTCCTTG GTACCCAGCT TCC

(2) SEQ ID NO: 22 信息:

(i) 序列特征:

(A) 长度: 20 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7424

(xi) 序列描述: SEQ ID NO : 22 :

TTAGACACCT GGCCAGAATG

(2) SEQ ID NO : 23 信息:

(i) 序列特征:

(A) 长度: 24 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7421

(xi) 序列描述: SEQ ID NO : 23 :

TGATGTCGGC AGTGTCTGAG AACC

(2) SEQ ID NO : 24 信息:

(i) 序列特征:

(A) 长度: 29 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7454

(xi) 序列描述: SEQ ID NO : 24 :

CCGGAATTCT TAGACACCTG GCCAGAATG

(2) SEQ ID NO : 25 信息:

(i) 序列特征:

(A) 长度: 33 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7453

(xi) 序列描述: SEQ ID NO : 25 :

CCGGAATTCT GATGTCGGCA GTGTCTGAGA ACC

(2) SEQ ID NO : 26 信息:

(i) 序列特征:

(A) 长度: 33 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7318

(xi) 序列描述: SEQ ID NO : 26 :

TACCGAATTC TAGACACAGA GGGTGGGACC TTC

(2) SEQ ID NO : 27 信息:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线形

(vii) 直接来源:

(B) 克隆: ZC7319

(xi) 序列描述: SEQ ID NO : 27 :

ACACTGAATT CTTCTCCACC CGGACAGAGT

(2) SEQ ID NO : 28 信息:

(i) 序列特征:

(A) 长度: 4823 个碱基对

(B) 类型: 核酸

(C) 链型: 双链

(D) 拓扑结构: 线形

(ii) 分子类型: DNA (基因组)

(ix) 特征:

(A) 名称/关键词: CDS

(B) 位置: 连接 (632..644,876..1003, 1290..1376,
3309..3476, 3713..4375)

(xi) 序列描述: SEQ ID NO : 28 :

CTTCTTGCT TTCTTCTTT CTTTCTTCT TTCTTTTTT TTTTGAGAC GGAGTTTCAC	60
TCTTATTGCC CAGGCTGGAG TGCAATGGTG CGATCTCGGC TCACCACAAC CTCCGCCTCC	120
CAGGTACAAG CGATTCTCCT GTCTCAGCCT CCCAAGTAGC TTGGATTACA GGCATGAACC	180
ACCACACCCT GCTAGTTTTT TTGTATTTTCG TAGAGCCGGG GTTTCACCAT GTTAGTGAGG	240
CTGGTGGCGA ACTCCTGACC TCAGGTGATC CACCCGCCTT GGACTCCCAA AGTGCTGGGA	300
TTACAGGCAT GAGCCACTGC ACCCGGCACA CCATATGCTT TCATCACAAG AAAATGTGAG	360
AGAATTCAGG GCTTTGGCAG TTCCAGGCTG GTCAGCATCT CAAGCCCTCC CCAGCATCTG	420
TTCACCCTGC CAGGCAGTCT CTTCTAGAA ACTTGGTTAA ATGTTCACTC TTCTTGCTAC	480
TTTCAGGATA GATTCTTCAC CTTGGTCCG CCTTTGCCCC ACCCTACTCT GCCCAGAAGT	540
GCAAGAGCCT AAGCCGCCTC CATGGCCCCA GGAAGGATTC AGGGGAGAGG CCCCAAACAG	600
GGAGCCACGC CAGCCAGACA CCCCAGCCAG A ATG GAG CTG ACT G GTGAGAACAC	654
Met Glu Leu Thr	

ACCTGAGGGG CTAGGGCCAT ATGGAAACAT GACAGAAGGG GAGAGAGAAA GGAGACACGC	714
TGCAGGGGGC AGGAAGCTGG GGGAAACCCAT TCTCCCAAAA ATAAGGGGTC TGAGGGGTGG	774
ATTCCCTGGG TTTCAGGTCT GGGTCCTGAA TGGGAATTCC TGGGAATACCA GCTGACAATG	834
ATTTCCCTCCT CATCTTTCAA CCTCACCTCT CCTCATCTAA G AA TTG CTC CTC	886
	Glu Leu Leu Leu
	5
GTG GTC ATG CTT CTC CTA ACT GCA AGG CTA ACG CTG TCC AGC CCG GCT	934
Val Val Met Leu Leu Leu Thr Ala Arg Leu Thr Leu Ser Ser Pro Ala	
10 15 20	
CCT CCT GCT TGT GAC CTC CGA GTC CTC AGT AAA CTG CTT CGT GAC TCC	982
Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu Leu Arg Asp Ser	
25 30 35 40	
CAT GTC CTT CAC AGC AGA CTG GTGAGAACTC CCAACATTAT CCCCTTTATC	1033
His Val Leu His Ser Arg Leu	
45	
CGCGTAACTG GTAAGACACC CATACTCCCA GGAAGACACC ATCACTTCCT CTAACCTCCT	1093
GACCCAATGA CTATTCTTCC CATATTGTCC CCACCTACTG ATCACACTCT CTGACAAGGA	1153
TTATTCTTCA CAATACAGCC CGCATTAAA AGCTCTCGTC TAGAGATAGT ACTCATGGAG	1213
GACTAGCCTG CTTATTAGGC TACCATAGCT CTCTCTATTT CAGCTCCCTT CTCCCCCAC	1273
CAATCTTTTT CAACAG AGC CAG TGC CCA GAG GTT CAC CCT TTG CCT ACA	1322
Ser Gln Cys Pro Glu Val His Pro Leu Pro Thr	
50 55	
CCT GTC CTG CTG CCT GCT GTG GAC TTT AGC TTG GGA GAA TGG AAA ACC	1370
Pro Val Leu Leu Pro Ala Val Asp Phe Ser Leu Gly Glu Trp Lys Thr	
60 65 70	
CAG ATG GTAAGAAAGC CATCCCTAAC CTTGGCTTCC CTAAGTCCTG TCTTCAGTTT	1426
Gln Met	
75	

CCCACTGCTT CCCATGGATT CTCCAACATT CTTGAGCTTT TAAAAATAT CTCACCTTCA	1486
GCTTGGCCAC CCTAACCCAA TCTACATTCA CCTATGATGA TAGCCTGTGG ATAAGATGAT	1546
GGCTTGCAGG TCCAATATGT GAATAGATTT GAAGCTGAAC ACCATGAAAA GCTGGAGAGA	1606
AATCGCTCAT GGCCATGCCT TTGACCTATT CCCGTTCACT CTTCTTAAAT TGGCATGAAG	1666
AAGCAAGACT CATATGTCAT CCACAGATGA CACAAAGCTG GGAAGTACCA CTAAAATAAC	1726
AAAAGACTGA ATCAAGATTC AAATCACTGA AAGACTAGGT CAAAAACAAG GTGAAACAAC	1786
AGAGATATAA ACTTCTACAT GTGGGCCGGG GGCTCACGCC TGTAATCCCA GCACTTTGGG	1846
AGGCCGAGGC AGGCAGATCA CCTGAGGGCA GGAGTTTGAG AGCAGCCTGG CCAACATGGC	1906
GAAACCCCGT CTCTACTAAG AATACAGAAT TAGCCGGGCA TGGTAGTGCA TGCCTGTAAT	1966
CCCAGCTACT TGGAAGGCTG AAGCAGGAGA ATCCCTTGAA CCCAGGAGGT GGAGGTTGTA	2026
GTGAGCTGAG ATCATGCCAA TGCCTCCAG CCTGGGTGAC AAGAGCAAAA CTCCGTCTCA	2086
AAAAGAAAAA AAAATTCTAC ATGTGTAAT TAATGAGTAA AGTCCTATTC CAGCTTTCAG	2146
GCCACAATGC CCTGCTTCCA TCATTTAAGC CTCTGGCCCT AGCACTTCCT ACGAAAAGGA	2206
TCTGAGAGAA TTAAATTGCC CCCAACTTA CCATGTAACA TTAAGTGAAGC TGCTATTCTT	2266
AAAGCTAGTA ATTCTTGCTT GTTTGATGTT TAGCATCCCC ATTGTGGAAA TGCTCGTACA	2326
GAACTCTATT CCGAGTGGAC TACTTAAA TATACTGGCC TGAACACCGG ACATCCCCCT	2386
GAAGACATAT GCTAATTTAT TAAGAGGGAC CATATTAAC TAACATGTGT CTAGAAAGCA	2446
GCAGCCTGAA CAGAAAGAGA CTAGAAGCAT GTTTTATGGG CAATAGTTTA AAAAATAAA	2506
ATCTATCCTC AAGAACCCTA GCGTCCCTC TTCCTTCAGG ACTGAGTCAG GGAAGAAGGG	2566
CAGTTCCTAT GGGTCCCTC TAGTCCTTC TTTTCATCCT TATGATCATT ATGGTAGAGT	2626
CTCATACCTA CATTAGTTT ATTTATTATT ATTATTTGAG ACGGAGTCTC ACTCTATCCC	2686
CCAGGCTGGA GTGCAGTGGC ATGATCTCAA CTCCTGCAA CCTCAGCCTC CCGGATTCAA	2746

GCGATTCTCC TGTCTCAGTC TCCCAAGTAG CTGGGATTAC AGGTGCCCCAC CACCATGCCC	2806
AGCTAATTTG TGTATTTGTG GTAGAGATGG GGTTTCACCA TGTTGGGCAG GCTGATCTTG	2866
AACTCCTGAC CTCAGGTGAT CCACCTGCCT CAGCCTCCCA AAGTGCTGGG ATTACAGGCG	2926
TGAGCCACTG CACCCAGCCT TCATTTCAGTT TAAAAATCAA ATGATCCTAA GGTTTTGCAG	2986
CAGAAAGAGT AAATTTGCAG CACTAGAACC AAGAGGTAAG AGCTGTAACA GGGCAGATTT	3046
CAGCAACGTA AGAAAAAAGG AGCTCTTCTC ACTGAAACCA AGTGTAAGAC CAGGCTGGAC	3106
TAGAGGACAC GGGAGTTTTT GAAGCAGAGG CTGATGACCA GCTGTCGGGA GACTGTGAAG	3166
GAATTCCTGC CCTGGGTGGG ACCTTGGTCC TGTCAGTTC TCAGCCTGTA TGATTCACTC	3226
TGCTGGCTAC TCCTAAGGCT CCCCACCCGC TTTTAGTGTG CCCTTTGAGG CAGTGCGCTT	3286
CTCTCTTCCA TCTCTTTCTC AG GAG GAG ACC AAG GCA CAG GAC ATT CTG GGA	3338
Glu Glu Thr Lys Ala Gln Asp Ile Leu Gly	
80 85	
GCA GTG ACC CTT CTG CTG GAG GGA GTG ATG GCA GCA CGG GGA CAA CTG	3386
Ala Val Thr Leu Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln Leu	
90 95 100	
GGA CCC ACT TGC CTC TCA TCC CTC CTG GGG CAG CTT TCT GGA CAG GTC	3434
Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln Val	
105 110 115	
CGT CTC CTC CTT GGG GCC CTG CAG AGC CTC CTT GGA ACC CAG	3476
Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu Leu Gly Thr Gln	
120 125 130	
GTAAGTCCCC AGTCAAGGGA TCTGTAGAAA CTGTTCTTTT CTGACTCAGT CCCCCTAGAA	3536
GACCTGAGGG AAGAAGGGCT CTTCCAGGGA GCTCAAGGGC AGAAGAGCTG ATCTACTAAG	3596
AGTGCTCCCT GCCAGCCACA ATGCCTGGGT ACTGGCATCC TGTCTTTCCT ACTTAGACAA	3656
GGGAGGCCTG AGATCTGGCC CTGGTGTTTG GCCTCAGGAC CATCCTCTGC CCTCAG	3712
CTT CCT CCA CAG GGC AGG ACC ACA GCT CAC AAG GAT CCC AAT GCC ATC	3760
Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp Pro Asn Ala Ile	
135 140 145	

TTC CTG AGC TTC CAA CAC CTG CTC CGA GGA AAG GTG CGT TTC CTG ATG	3808
Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val Arg Phe Leu Met	
150 155 160	
CTT GTA GGA GGG TCC ACC CTC TGC GTC AGG CGG GCC CCA CCC ACC ACA	3856
Leu Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala Pro Pro Thr Thr	
165 170 175 180	
GCT GTC CCC AGC AGA ACC TCT CTA GTC CTC ACA CTG AAC GAG CTC CCA	3904
Ala Val Pro Ser Arg Thr Ser Leu Val Leu Thr Leu Asn Glu Leu Pro	
185 190 195	
AAC AGG ACT TCT GGA TTG TTG GAG ACA AAC TTC ACT GCC TCA GCC AGA	3952
Asn Arg Thr Ser Gly Leu Leu Glu Thr Asn Phe Thr Ala Ser Ala Arg	
200 205 210	
ACT ACT GGC TCT GGG CTT CTG AAG TGG CAG CAG GGA TTC AGA GCC AAG	4000
Thr Thr Gly Ser Gly Leu Leu Lys Trp Gln Gln Gly Phe Arg Ala Lys	
215 220 225	
ATT CCT GGT CTG CTG AAC CAA ACC TCC AGG TCC CTG GAC CAA ATC CCC	4048
Ile Pro Gly Leu Leu Asn Gln Thr Ser Arg Ser Leu Asp Gln Ile Pro	
230 235 240	
GGA TAC CTG AAC AGG ATA CAC GAA CTC TTG AAT GGA ACT CGT GGA CTC	4096
Gly Tyr Leu Asn Arg Ile His Glu Leu Leu Asn Gly Thr Arg Gly Leu	
245 250 255 260	
TTT CCT GGA CCC TCA CGC AGG ACC CTA GGA GCC CCG GAC ATT TCC TCA	4144
Phe Pro Gly Pro Ser Arg Arg Thr Leu Gly Ala Pro Asp Ile Ser Ser	
265 270 275	
GGA ACA TCA GAC ACA GGC TCC CTG CCA CCC AAC CTC CAG CCT GGA TAT	4192
Gly Thr Ser Asp Thr Gly Ser Leu Pro Pro Asn Leu Gln Pro Gly Tyr	
280 285 290	
TCT CCT TCC CCA ACC CAT CCT CCT ACT GGA CAG TAT ACG CTC TTC CCT	4240
Ser Pro Ser Pro Thr His Pro Pro Thr Gly Gln Tyr Thr Leu Phe Pro	
295 300 305	
CTT CCA CCC ACC TTG CCC ACC CCT GTG GTC CAG CTC CAC CCC CTG CTT	4288
Leu Pro Pro Thr Leu Pro Thr Pro Val Val Gln Leu His Pro Leu Leu	
310 315 320	

CCT GAC CCT TCT GCT CCA ACG CCC ACC CCT ACC AGC CCT CTT CTA AAC 4336
 Pro Asp Pro Ser Ala Pro Thr Pro Thr Pro Thr Ser Pro Leu Leu Asn
 325 330 335 340

ACA TCC TAC ACC CAC TCC CAG AAT CTG TCT CAG GAA GGG TAAGGTTCTC 4385
 Thr Ser Tyr Thr His Ser Gln Asn Leu Ser Gln Glu Gly
 345 350

AGACACTGCC GACATCAGCA TTGTCTCGTG TACAGCTCCC TTCCTGCAG GCGGCCCTG 4445

GGAGACAACT GGACAAGATT TCCTACTTTC TCCTGAAACC CAAAGCCCTG GTAAAAGGGA 4505

TACACAGGAC TGAAAAGGGA ATCATTTTTC ACTGTACATT ATAAACCTTC AGAAGCTATT 4565

TTTTAAAGCT ATCAGCAATA CTCATCAGAG CAGCTAGCTC TTTGGTCTAT TTTCTGCAGA 4625

AATTTGCAAC TCACTGATTC TCAACATGCT CTTTTTCTGT GATAACTCTG CAAAGACCTG 4685

GGCTGGCCTG GCAGTTGAAC AGAGGGAGAG ACTAACCTTG AGTCAGAAAA CAGAGGAAGG 4745

GTAATTCCT TTGCTTCAA TTCAAGGCCT TCCAACGCC CCATCCCCTT TACTATCATT 4805

CTCAGTGGGA CTCTGATC

(2) SEQ ID NO : 29 信息:

(i) 序列特征:

(A) 长度: 353 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线形

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO : 29 :

Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala
 1 5 10 15

Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val
 20 25 30

Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser
 35 40 45

Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala
 50 55 60

Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys
 65 70 75 80

Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met
 85 90 95

Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly
 100 105 110

Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu
 115 120 125

Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp
 130 135 140

Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val
 145 150 155 160

Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala
 165 170 175

Pro Pro Thr Thr Ala Val Pro Ser Arg Thr Ser Leu Val Leu Thr Leu
 180 185 190

Asn Glu Leu Pro Asn Arg Thr Ser Gly Leu Leu Glu Thr Asn Phe Thr
 195 200 205

Ala Ser Ala Arg Thr Thr Gly Ser Gly Leu Leu Lys Trp Gln Gln Gly
 210 215 220

Phe Arg Ala Lys Ile Pro Gly Leu Leu Asn Gln Thr Ser Arg Ser Leu
 225 230 235 240

Asp Gln Ile Pro Gly Tyr Leu Asn Arg Ile His Glu Leu Leu Asn Gly
 245 250 255

Thr Arg Gly Leu Phe Pro Gly Pro Ser Arg Arg Thr Leu Gly Ala Pro
 260 265 270

Asp Ile Ser Ser Gly Thr Ser Asp Thr Gly Ser Leu Pro Pro Asn Leu
275 280 285

Gln Pro Gly Tyr Ser Pro Ser Pro Thr His Pro Pro Thr Gly Gln Tyr
290 295 300

Thr Leu Phe Pro Leu Pro Pro Thr Leu Pro Thr Pro Val Val Gln Leu
305 310 315 320

His Pro Leu Leu Pro Asp Pro Ser Ala Pro Thr Pro Thr Pro Thr Ser
325 330 335

Pro Leu Leu Asn Thr Ser Tyr Thr His Ser Gln Asn Leu Ser Gln Glu
340 345 350

Gly

图 1

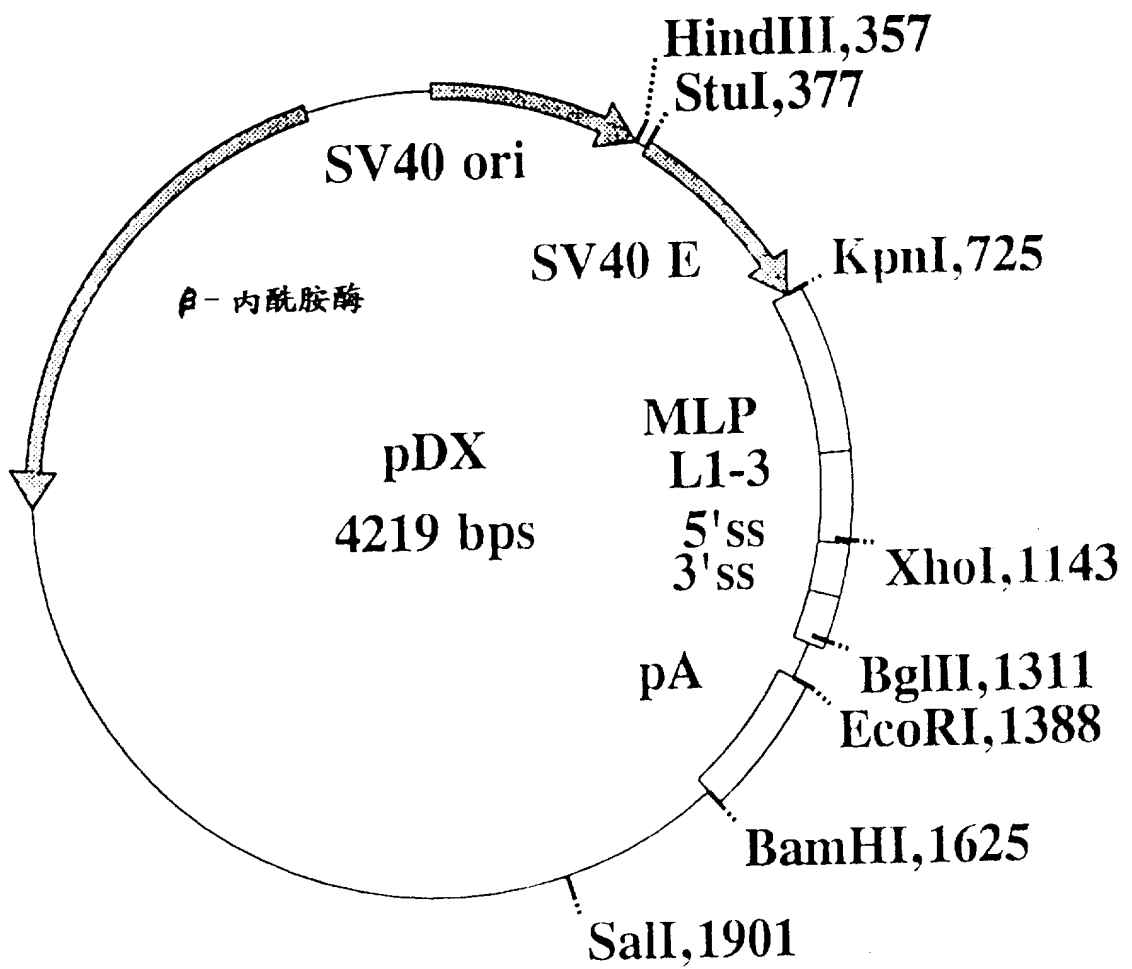


图 2

