

[19]中华人民共和国国家知识产权局

[51]Int.Cl<sup>7</sup>

C12N 15/56

C12N 15/63 C12N 1/21

C12N 9/24 C12N 15/11

A61K 38/47

## [12]发明专利申请公开说明书

[21]申请号 98809759.1

[43]公开日 2000年11月8日

[11]公开号 CN 1272886A

[22]申请日 1998.8.31 [21]申请号 98809759.1

[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

[30]优先权

代理人 黄革生

[32]1997.9.2 [33]US [31]08/922,170

[32]1998.7.2 [33]US [31]09/109,386

[86]国际申请 PCT/US98/17954 1998.8.31

[87]国际公布 WO99/11798 英 1999.3.11

[85]进入国家阶段日期 2000.3.31

[71]申请人 洞察策略与营销有限公司

地址 以色列雷霍沃特

共同申请人 哈达斯特医疗研究服务和开发有限公司

[72]发明人 I·佩克 I·弗罗达夫斯基

E·菲斯坦因

权利要求书3页 说明书54页 附图页数14页

[54]发明名称 编码具类肝素酶活性的多肽的多核苷酸及其在转导细胞中的表达

[57]摘要

编码具类肝素酶活性的多肽的多核苷酸，包括该多核苷酸的载体，表达类肝素酶的转导细胞和具类肝素酶活性的重组蛋白。

ISSN 1 0 0 8 - 4 2 7 4

## 权 利 要 求 书

---

1. 包含编码具类肝素酶催化活性之多肽的多核苷酸序列的多核苷酸片段。
2. 权利要求 1 的多核苷酸片段，其中所述多核苷酸序列包含 SEQ ID NO : 9 的核苷酸 63-1691，或 SEQ ID NO: 13 的核苷酸 139-1869。
3. 权利要求 1 的多核苷酸片段，其中所述多核苷酸序列包含 SEQ ID NO : 9 的核苷酸 63-721。
4. 权利要求 1 的多核苷酸片段，其中所述多核苷酸如 SEQ ID NO: 9 或 13 中所示。
5. 权利要求 1 的多核苷酸片段，其中所述多核苷酸序列包含 SEQ ID NO : 9 或 13 的区段，该区段编码具所述类肝素酶催化活性的所述多肽。
6. 权利要求 1 的多核苷酸片段，其中所述多肽包含如 SEQ ID NO: 10 或 14 中所示的氨基酸序列。
7. 权利要求 1 的多核苷酸片段，其中所述多肽包含 SEQ ID NO: 10 或 14 的区段，该区段具有所述类肝素酶催化活性。
8. 权利要求 1 的多核苷酸片段，其中所述多核苷酸序列选自双链 DNA、单链 DNA 或者 RNA。
9. 包含与编码具类肝素酶催化活性之多肽的多核苷酸链至少一部分互补的多核苷酸序列的单链多核苷酸片段。
10. 权利要求 9 的多核苷酸片段，其中所述多核苷酸序列包含 SEQ ID NO: 9 或 13 的至少一部分。
11. 包含编码具类肝素酶催化活性之多肽的多核苷酸序列的载体。
12. 权利要求 11 的载体，其中所述多核苷酸序列包含 SEQ ID NO: 9 的核苷酸 63-1691，或 SEQ ID NO: 13 的核苷酸 139-1869。
13. 权利要求 11 的载体，其中所述多核苷酸序列包含 SEQ ID

NO: 9 的核苷酸 63-721。

14. 权利要求 11 的载体，其中所述多核苷酸序列是 SEQ ID NO: 9 或 13 中所示序列。

15. 权利要求 11 的载体，其中所述多核苷酸序列包含 SEQ ID NO: 9 或 13 的区段，该区段编码具所述类肝素酶催化活性的所述多肽。

16. 权利要求 11 的载体，其中所述多肽包含如 SEQ ID NO: 10 或 14 中所示的氨基酸序列。

17. 权利要求 11 的载体，其中所述多肽包含 SEQ ID NO: 10 或 14 的区段，该区段具有所述类肝素酶催化活性。

18. 权利要求 11 的载体，其中所述多核苷酸序列选自双链 DNA、单链 DNA 或者 RNA。

19. 一种包含外源多核苷酸片段的宿主细胞，其中所述外源多核苷酸片段包含编码具类肝素酶催化活性的多肽的多核苷酸序列。

20. 权利要求 19 的宿主细胞，其中所述多核苷酸序列包含 SEQ ID NO: 9 的核苷酸 63-1691，或 SEQ ID NO: 13 的核苷酸 139-1869。

21. 权利要求 19 的宿主细胞，其中所述多核苷酸序列包含 SEQ ID NO: 9 的核苷酸 63-721。

22. 权利要求 19 的宿主细胞，其中所述多核苷酸序列是 SEQ ID NO: 9 或 13 中所示序列。

23. 权利要求 19 的宿主细胞，其中所述多核苷酸序列包含 SEQ ID NO: 9 或 13 的区段，该区段编码具所述类肝素酶催化活性的所述多肽。

24. 权利要求 19 的宿主细胞，其中所述多肽包含如 SEQ ID NO: 10 或 14 中所示的氨基酸序列。

25. 权利要求 19 的宿主细胞，其中所述多肽包含 SEQ ID NO: 10 或 14 的区段，该区段具有所述类肝素酶催化活性。

26. 权利要求 19 的宿主细胞，其中所述多核苷酸序列选自双链 DNA、单链 DNA 或者 RNA。

27. 一种表达重组类肝素酶的宿主细胞。

28. 包含具类肝素酶催化活性的多肽的重组蛋白。

29. 权利要求 28 的重组蛋白，其中所述多肽包含 SEQ ID NO: 10 或 14 的区段。

30. 包含能与 SEQ ID NO: 9 的核苷酸 1-721 杂交的多核苷酸序列的多核苷酸片段。

31. 如 SEQ ID NO: 9 或 13 中所示的多核苷酸序列。

32. 与 SEQ ID NO: 10 或 13 同源的多核苷酸序列。

33. 如 SEQ ID NO: 10 或 14 中所示的氨基酸序列。

34. 与 SEQ ID NO: 10 或 14 同源的氨基酸序列。

35. 一种包含具类肝素酶催化活性的重组蛋白作为活性组分的药物组合物。

36. 一种包含过量表达类肝素酶催化活性的细胞的类肝素酶过量表达系统。

37. 调节肝素结合性生长因子、对肝素结合性生长因子和细胞因子的细胞应答、与血浆脂蛋白的细胞相互作用、对病毒、原生动物和细菌感染的细胞易感性或者神经变性斑分解的调节子，其中包含具类肝素酶催化活性的重组蛋白作为活性组分。

38. 包括含有具类肝素酶催化活性的重组蛋白作为活性组分的医疗装置的医疗设备。

39. 权利要求 11 的载体，其中所述载体是杆状病毒载体。

40. 权利要求 19 的宿主细胞，其中所述细胞是昆虫细胞。

41. 权利要求 27 的宿主细胞，其中所述细胞是昆虫细胞。

42. 鉴定铺展的染色体中含人类肝素酶基因的染色体区的方法，包括以下步骤：

(a) 将铺展的染色体与编码类肝素酶的标记多核苷酸探针杂交；

(b) 洗涤铺展的染色体，从而除去过剩的未杂交探针；

(c) 搜寻与所述杂交的标记多核苷酸探针相关的信号，其中探测到的信号指示含人类肝素酶基因的染色体区。

# 说 明 书

---

## 编码具类肝素酶活性的多肽 的多核苷酸及其在转导细胞中的表达

### 发明领域及背景

本发明涉及编码具类肝素酶 (heparanase) 活性的多肽的多核苷酸(下文中称作 hpa)，包括此多核苷酸的载体和表达类肝素酶的转导细胞。本发明还进一步涉及具类肝素酶活性的重组蛋白。

**硫酸乙酰肝素蛋白多糖：**硫酸乙酰肝素蛋白多糖(HSPG)是与脊椎动物和无脊椎动物组织中的许多细胞的细胞表面和胞外基质(ECM)有关的普遍存在的大分子(1—4)。基本的 HSPG 结构包括蛋白核心以及与之共价相连的几条线性硫酸乙酰肝素链。这些多糖链通常由重复的己糖醛和 D-葡糖胺二糖单元组成，这些单元不同程度地被 N-和 O-连接的硫酸组分和 N-连接的乙酰基团所取代(1—4)。对 ECM 分子参与细胞附着、生长和分化的研究显示，HSPG 在胚形态发生、血管发生、轴突旁枝突起和组织修复中处于关键地位(1—5)。HSPG 是血管的重要组分(3)。在大血管中，它们主要集中在血管内膜和内血管中膜，而在毛细血管中，它们主要在内皮下基底膜，支持内皮细胞繁殖和迁移并稳定毛细血管壁的结构。HSPG 与 ECM 大分子如胶原、层粘连蛋白和纤连蛋白等在质膜上的不同附着位点相互作用的能力暗示了此蛋白多糖在 ECM 组分的自组装和不溶性，以及细胞粘着和细胞运动中起着重要作用。硫酸乙酰肝素(HS)链的裂解可导致内皮下 ECM 的降解，从而在血源细胞的外渗中起决定性作用。HS 分解代谢可在炎症、创伤修复、糖尿病和癌症转移中观察到，说明降解 HS 的酶在病理过程中起重要作用。类肝素酶活性曾在活化的免疫系统细胞和高度转移的癌细胞中描述过(6—8)，但由于缺乏探索疾病状况下类肝素酶潜在的病原作用的生物工具，因此阻碍了研究。

**肿瘤细胞侵入和转移中类肝素酶的参与：**停留在不同器官毛细血

管床内的循环肿瘤细胞必须侵入内皮细胞内层并降解其底层基底膜(BM)以便侵入血管外组织，从而发生转移(9, 10)。转移的肿瘤细胞经常附着在邻近内皮细胞之间的胞间连接处或附近。转移细胞附着之后，胞间连接破裂，内皮细胞边界收缩，通过内皮缺口向暴露的底层基底膜迁移(9)。一旦定位在内皮细胞和BM之间，则侵入细胞必须降解BM的内皮下糖蛋白和蛋白多糖以便迁移出血管区室。若干细胞酶(例胶原酶IV, 纤溶酶原激活物, 组织蛋白酶B, 弹性蛋白酶等)与BM的降解有关(10)。在这些酶中，有一种内切- $\beta$ -D-葡糖昔酸酶(类肝素酶)，它在特殊链内位点裂解HS(6, 8, 11)。HS降解性类肝素酶的表达与小鼠淋巴瘤(11)、纤维肉瘤和黑素瘤(8)细胞的转移潜力相关。此外，在患有转移肿瘤的动物和黑素瘤患者的血清(8)和癌症患者的肿瘤活组织(12)中检测到类肝素酶水平升高。

本发明人以前研究过通过局部微环境对细胞增殖和肿瘤发展进行控制，主要着眼于细胞与培养的角膜和血管内皮细胞产生的细胞外基质(ECM)之间的相互作用。该培养ECM在其形态外观及分子组成方面与体内的内皮下层十分类似。它包括胶原(主要是III型和IV型，及少量I型和V型)，蛋白多糖(主要为硫酸乙酰肝素和硫酸皮肤素蛋白多糖，以及少量硫酸软骨素蛋白多糖)，层粘连蛋白，纤连蛋白，巢蛋白和弹性蛋白(13, 14)。细胞降解培养ECM中的HS的能力可通过以下方法研究：使细胞与新陈代谢地硫酸盐标记的ECM相互作用，对释放到培养基中的降解产物进行凝胶过滤(琼脂糖6B)分析(11)。完整的HSPG在柱的空体积之后洗脱( $Kav < 0.2$ ,  $Mr \approx 0.5 \times 10^6$ )，而标记的HS侧链降解片段在更接近柱的 $V_t$ 时洗脱( $0.5 < Kav < 0.8$ ,  $Mr = 5-7 \times 10^3$ ) (11)。

本发明者还研究了各种非抗凝剂类型的肝素对类肝素酶的抑制作用，这些肝素有可能用于防止血源细胞的外渗。包括16个或更多糖单位并在N和O位置都具硫酸基团的肝素可最成功抑制类肝素酶。O-脱硫作用消除了肝素对类肝素酶的抑制作用，而只要N-取代分子分子量大于等于4, 000道尔顿，则O-硫酸化、N-乙酰化肝素保持高抑制活

性(7)。对实验动物用类肝素酶抑制剂(例如, 肝素的非抗凝剂类型)处理, 则显著减少(>90%) B16 黑素瘤、Lewis 肺癌和乳腺癌细胞诱导的肺部转移(7, 8, 16)。与抗凝血酶III具高和低亲合性的肝素级分表现出相似的高抗转移活性, 这表明肝素的类肝素酶抑制活性而非其抗凝活性在多糖的抗转移特性中起作用(7)。

癌症患者尿中的类肝素酶活性: 为尝试进一步阐明类肝素酶在肿瘤发展中的作用及与人类癌症的相关性, 对尿样进行类肝素酶活性检查(16a)。在一部分、而非所有癌症患者的尿样中检测到类肝素酶活性。在患有侵入性转移疾病的患者的尿液中检测到高水平的类肝素酶活性, 而在健康供体的尿液中检测不到类肝素酶活性。

在 20% 的正常人和微球蛋白尿胰岛素依赖型糖尿病 (IDDM) 患者中也发现有类肝素酶活性, 这很可能是由于糖尿病性肾病——导致成人肾衰竭的最重要的疾病所致。

类肝素酶可能参与肿瘤血管发生: 成纤维细胞生长因子为结构相关的多肽家族, 具有对肝素高亲和性的特征(17)。它们对血管内皮细胞高度促有丝分裂, 且是最有效的新血管形成的诱导物之一(17, 18)。已从体外生成的内皮下 ECM(19)和角膜的基底膜(20)中提取出了碱性成纤维细胞生长因子(bFGF), 提示 ECM 可作为 bFGF 的储存器。免疫组化染色显示 bFGF 定位在多种组织和血管的基底膜中(21)。尽管 bFGF 在正常组织中普遍存在, 但这些组织中的内皮细胞通常增殖很低, 说明 bFGF 以某种方式与其作用位点隔离。对 bFGF 与 ECM 相互作用的研究表明: bFGF 与 ECM 中 HSPG 结合, 且能通过 HS 降解酶以活化形式释放(15, 20, 22)。已证实血小板、肥大细胞、嗜中性粒细胞和淋巴瘤细胞表达的类肝素酶活性与从 ECM 和基底膜释放活性 bFGF 有关(23), 说明类肝素酶活性可能不仅参与细胞迁移和侵入, 而且还可能诱发间接新血管反应。这些结果说明 ECM HSPG 是 bFGF、可能还是其它肝素结合性生长促进因子的天然储存库(24, 25)。因此, bFGF 从其在基底膜和 ECM 内的储备中移位可能是正常和病理状况下诱导新血管形成的一种新机制。

近来的研究显示，肝素和 HS 参与 bFGF 与高亲和力细胞表面受体的结合及 bFGF 细胞信号传导（26, 27）。此外，达到最佳效果所需的 HS 的大小与类肝素酶释放的 HS 片段的大小相似（28）。对血管内皮细胞生长因子（VEGF）也得到类似结果（29），表明有一个涉及 HS 与肝素结合性生长因子的细胞相互作用的双重受体机制在起作用。因此推断 ECM 中内皮细胞生长因子的限制化避免了它们系统性地作用于血管内皮，从而维持极低速率的内皮细胞转换和血管生长。另一方面，bFGF 作为与 HS 片段的复合体从 ECM 内的储备中释放出来可能会刺激例如创伤愈合、炎症和肿瘤发展等过程中的局部内皮细胞增殖和新血管生成（24, 25）。

免疫系统细胞的类肝素酶的表达：类肝素酶的活性与免疫系统内活化细胞脱离循环并引起炎症和自身免疫反应的能力有关。血小板、粒细胞、T 和 B 淋巴细胞、巨噬细胞和肥大细胞与内皮下 ECM 的相互作用与通过特异性类肝素酶活性降解 HS 有关（6）。类肝素酶在各种活化信号（例如凝血酶，钙离子载体，免疫复合体，抗原，有丝分裂原，等）刺激下从细胞内区室（例如溶酶体，特殊颗粒，等）释放，说明其受调控地参与炎症和细胞免疫。

参考文献 6 中回顾了关于类肝素酶的观察结果，列举如下文：

第一，蛋白水解活性（纤溶酶原激活物）和类肝素酶协同参与炎性白细胞和恶性细胞对 ECM HSPG 的顺序降解。

第二，大部分血小板类肝素酶以潜伏形式存在，很可能是作为与硫酸软骨素的复合体存在。该潜伏酶被肿瘤细胞衍生因子活化，随后可在肿瘤转移过程中促进通过血管内皮的细胞侵入。

第三，强刺激剂（例如凝血酶）诱导血小板类肝素酶从  $\alpha$ -颗粒中释放，但 ECM 上血小板活化不诱导释放。

第四，因阈值激活和细胞在 ECM 上孵育，优先并很容易地释放嗜中性粒细胞类肝素酶。

第五，嗜中性粒细胞与 ECM 的接触抑制了毒性酶（蛋白水解酶，溶菌酶）和氧自由基的释放，但不抑制能使血细胞渗出的酶（类肝素酶，

明胶酶)的释放。当可溶因子而非可被吞噬的刺激剂刺激时，可观察到内皮下 ECM 的这种保护作用。

第六，细胞内类肝素酶在 T 细胞系暴露于特异抗原下几分钟内即被分泌。

第七，有丝分裂原(ConA, LPS)诱导体外维持的正常 T 和 B 淋巴细胞合成与分泌类肝素酶。T 淋巴细胞类肝素酶也可通过体内抗原免疫接种而诱导。

第八，前 B 淋巴瘤和 B 淋巴瘤表达类肝素酶活性，但浆细胞瘤和静息状态的正常 B 淋巴细胞不表达。

第九，活化的巨噬细胞在与 ECM 孵育过程中表达类肝素酶活性，但仅极少或根本没有酶释放到培养基中。诱导人骨髓白血病细胞分化为成熟巨噬细胞时，得到相似结果。

第十，低剂量的类肝素酶抑制性非抗凝剂型肝素抑制 T 细胞介导的迟发型超敏反应和实验性自身免疫(30)。

第十一，血小板、嗜中性粒细胞和转移肿瘤细胞表达的类肝素酶活性从 ECM 和基底膜释放活化的 bFGF。从 ECM 内的储备中释放 bFGF 可在如创伤愈合、炎症和肿瘤发展的过程中引发局部的新血管反应。

第十二，类肝素酶产生的 ECM 降解产物中具有能体内抑制 T 细胞介导的炎症的三硫酸化双糖(31)。此种抑制与双糖对活化的 T 细胞进行的生物活性 TNF $\alpha$  体外生产的抑制作用相关(31)。

其它潜在的治疗应用：哺乳动物的类肝素酶除了参与肿瘤细胞转移、炎症和自身免疫之外，还可用于调节：肝素结合性生长因子的生物利用率(15)；针对肝素结合性生长因子(例如 bFGF, VEGF)和细胞因子(IL-8)的细胞反应(31a, 29)；与血浆脂蛋白的细胞相互作用(32)；对某些病毒和一些细菌及原生动物感染的细胞易感性(33, 33a, 33b)；和淀粉状蛋白斑的分解(34)。由此证明类肝素酶可能在诸如创伤愈合、血管发生、再狭窄、动脉粥样硬化、炎症、神经变性疾病和病毒感染之类状况中有用。哺乳动物类肝素酶能作为血精蛋白的潜在替代物用于中和血浆肝素。抗类肝素酶的抗体可应用于活组织检查样

品、血浆样品和体液中的微转移、自身免疫损害和肾衰竭的免疫检测及诊断中。在基础研究中也有普遍用途。

编码类肝素酶的 hpa 基因的鉴别有助于在异源表达系统中制备重组酶。如能获得重组蛋白，将为解决蛋白结构功能关系铺平道路，并且为开发新抑制剂提供工具。

**病毒感染：**细胞表面上硫酸乙酰肝素的存在是单纯疱疹病毒(33)和登革病毒(33a)与细胞结合并进一步感染细胞的主要条件。因此通过类肝素酶除去细胞表面硫酸乙酰肝素可消除病毒感染(33)。事实上，用细菌类肝素酶(降解硫酸乙酰肝素)或肝素酶(降解类肝素)处理细胞，能降低两种相关的动物疱疹病毒与细胞的结合，使细胞对病毒感染至少有部分抗性。有迹象表明细胞表面硫酸乙酰肝素也参与 HIV 感染(33b)。

**神经变性疾病：**在 Genstmann-Straussler 综合症，Creutzfeldt-Jakob 疾病和 Scrapie 的朊病毒蛋白淀粉样蛋白斑中鉴别出了硫酸乙酰肝素蛋白多糖(34)。类肝素酶可分解这些被认为也在早老性痴呆发病机理中起作用的淀粉样蛋白斑。

**再狭窄和动脉粥样硬化：**因内皮损伤和富含胆固醇的脂蛋白累积造成的动脉平滑肌细胞(SMC)增殖是动脉粥样硬化和再狭窄发病机理中的基本事件(35)。HS 除了参与 SMC 增殖(即肝素结合性生长因子的低亲和性受体)外，还参与脂蛋白结合、滞留和摄取(36)。已证明 HSPG 和脂蛋白脂酶参与可使富含胆固醇的脂蛋白发生显著细胞和间质积累的新分解代谢途径(32)。后一种途径可能因促进富含 apoB 和 apoE 的脂蛋白(即 LDL, VLDL, 乳糜微粒)的积累而高度促进动脉粥样硬化，而与细胞固醇含量的反馈抑制无关。因而预计通过类肝素酶除去 SMC HS 可抑制 SMC 增殖和脂类积累，因此可以阻止再狭窄和动脉粥样硬化的发展。

因此，人们已普遍认识到，得到编码具类肝素酶活性的多肽的多核苷酸，包含此多核苷酸的载体，表达类肝素酶的转导细胞及具类肝素酶活性的重组蛋白，将是有必要的并且非常有益的。

## 发明概述

本发明提供编码具类肝素酶活性的多肽的多核苷酸(即下文所指的 hpa、hpa cDNA 或 hpa 基因)，本发明还提供包含此多核苷酸的载体，表达类肝素酶的转导细胞和具类肝素酶活性的重组蛋白。

在此报导对编码类肝素酶的人 hpa 基因的克隆和通过转染宿主细胞表达重组类肝素酶。

对从人肝癌细胞分离的类肝素酶的纯化制品进行胰蛋白酶消化和微量测序。公开的 YGPDVGQPR (SEQ ID NO: 8) 序列用于筛选 EST 数据库中相应的 DNA 序列的同源物。鉴定到两个紧密相关的 EST 序列，其后发现二者相同。两种克隆都包括 1020bp 的插入片段，此插入片段包括 973bp 的开放阅读框及随后的 27bp 3' 非翻译区和多聚腺苷酸尾。未鉴定到翻译起始位点。

使用根据 EST 克隆序列和复合体的接头选择的引物，通过对来自胎盘 Marathon RACE cDNA 复合体的 DNA 进行 PCR 扩增实施对缺失的 hpa 5' 端的克隆。获得 900bp PCR 片段，该片段与鉴定出的 EST 3' 编码克隆部分重叠。连接的 cDNA 片段(hpa)为 1721bp 长 (SEQ ID NO: 9)，其中包括编码 543 个氨基酸的多肽 (SEQ ID NO: 10) 的开放阅读框，计算分子量为 61,192 道尔顿。

运用 Marathon RACE 通过 PCR 扩增从人 SK-hep1 细胞系克隆延伸的 5' 序列。将 SK-hep1 hpa cDNA 的 5' 延伸序列与从人胎盘分离的 hpa cDNA (SEQ ID NO: 9) 序列装配起来。装配序列包括开放阅读框 SEQ ID NO: 13 和 15，该阅读框编码如 SEQ ID NO: 14 和 15 所示的 592 个氨基酸的多肽，计算分子量为 66,407 道尔顿。

运用杆状病毒表达系统，通过在昆虫细胞中表达 hpa 的整个开放阅读框，检测 hpa 基因产品在体外实验中催化硫酸乙酰肝素降解的能力。含 hpa 基因的病毒感染的细胞的提取物和条件培养基，经实验证明其对可溶性的 ECM 衍生 HSPG 和完整 ECM 均有高水平硫酸乙酰肝素降解活性。肝素是类肝素酶的另一底物，它抑制上述降解活性。用不

含 hpa 基因的相似构建体感染的细胞及未转染细胞都不具备此活性。在哺乳动物表达系统中证实从延伸的 5' 克隆表达的类肝素酶有对肝素的活性。

运用 RT-PCR 研究 hpa RNA 在不同组织和细胞系中的表达模式。发现 hpa RNA 仅在已知具类肝素酶活性的组织和细胞中表达。

运用一系列单染色体人/CHO 和人/小鼠体细胞杂合体将人类肝素酶基因定位在人第 4 号染色体上。该新分离的类肝素酶序列可用于鉴定铺展染色体中具有人类肝素酶基因的染色体区。

根据本发明以下描述的优选实施方案中进一步描绘的特征，本发明提供包括编码具类肝素酶催化活性之多肽的多核苷酸序列的多核苷酸片段。

根据描述的优选实施方案中的进一步特征，该多核苷酸片段包括 SEQ ID NO: 9 的核苷酸 63-1691 或 SEQ ID NO: 13 的核苷酸 139—1869，它们编码完整的人类肝素酶。

根据描述的优选实施方案中的进一步特征，本发明提供包括能与 hpa cDNA 杂交，尤其是与 SEQ ID NO: 9 的核苷酸 1—721 杂交的多核苷酸序列的多核苷酸片段。

根据描述的优选实施方案中的进一步特征，编码具类肝素酶活性的多肽的多核苷酸序列与 SEQ ID NO: 9 或 13 至少 60% 同源，优选至少 70% 同源，更优选至少 80% 同源，最优选至少 90% 同源。

根据描述的优选实施方案中的进一步特征，依照本发明的多核苷酸片段包括 SEQ ID NO: 9 或 13 的一部分(片段)。例如，这些片段可包括 SEQ ID NO: 9 的核苷酸 63—721 和/或 SEQ ID NO: 9 中编码具类肝素酶催化活性的多肽的区段。

根据描述的优选实施方案中的进一步特征，由该多核苷酸片段编码的多肽包括 SEQ ID NO: 10 或 14 中所示氨基酸序列或其功能性部分。

根据描述的优选实施方案中的进一步特征，该多核苷酸序列编码具类肝素酶活性的多肽，所述多肽与 SEQ ID NO: 10 或 14 至少 60%

同源，优选至少 70% 同源，更优选至少 80% 同源，最优选至少 90% 同源。

根据描述的优选实施方案中的进一步特征，该多核苷酸片段编码具类肝素酶活性的多肽，该多肽因此可能是 SEQ ID NO: 10 或 14 所示氨基酸序列的等位变体、种变体和/或诱导变体。任何此类变体也可被认作同系物。

根据描述的优选实施方案中的进一步特征，本发明提供一种单链多核苷酸片段，它包括与编码具如上述类肝素酶催化活性之多肽的多核苷酸链的至少一部分互补的多核苷酸序列。

根据描述的优选实施方案中的进一步特征，本发明提供包括编码具类肝素酶催化活性的多肽的多核苷酸序列的载体。

该载体可为任何适当形式，包括但不限于噬菌体、病毒、质粒、噬菌粒、粘粒、杆粒或甚至人工染色体。编码具类肝素酶催化活性的多肽的多核苷酸序列可包括以上描述的任何多核苷酸片段。

根据描述的优选实施方案中的进一步特征，本发明提供含有包括编码具类肝素酶催化活性之多肽的多核苷酸序列的外源多核苷酸片段的宿主细胞。

该外源多核苷酸片段可为以上描述的任何片段。宿主细胞可为任何形式，如原核细胞、真核细胞、细胞系或作为多细胞生物一部分的细胞（例如转基因生物的细胞）。

根据描述的优选实施方案中的进一步特征，本发明提供包含具类肝素酶催化活性的多肽的重组蛋白。

根据描述的优选实施方案中的进一步特征，本发明提供包含具类肝素酶催化活性的重组蛋白作为活性组分的药物组合物。

根据描述的优选实施方案中的进一步特征，本发明提供包括含有具类肝素酶催化活性的重组蛋白作为活性组分的医疗装置的医疗设备。

根据描述的优选实施方案中的进一步特征，本发明提供包括过表达类肝素酶催化活性的细胞的类肝素酶过量表达系统。

根据描述的优选实施方案中的进一步特征，本发明提供鉴定铺展

的染色体中包含人类肝素酶基因的染色体区的方法，包括以下步骤：  
(a) 将铺展的染色体与编码类肝素酶的标记多核苷酸探针杂交；(b) 洗涤铺展的染色体，从而除去过剩的未杂交探针；及(c) 搜寻与所述杂交的标记多核苷酸探针相关的信号，其中探测到信号指示包含人类肝素酶基因的染色体区。

本发明能用于开发抑制肿瘤细胞转移、炎症和自身免疫的新药。对编码类肝素酶的 hpa 基因的鉴定使在异源表达系统中生产重组酶成为可能。

### 附图简述

本发明仅通过举例的方式，参考附图在此描述，其中：

图 1 显示 hpa cDNA 的核苷酸序列和推断的氨基酸序列。在 EST(表达序列标志)和 PCR 扩增的 cDNA(反转录 RNA)之间，存在位于 799 位的单核苷酸差异(A 改为 T)，此差异及其导致的氨基酸替代(Tyr 改为 Phe)在各取代单元的上下标出。半胱氨酸残基和聚腺苷酸化共有序列下划线标出。星号表示终止密码子 TGA。

图 2 表示以 pFhpa2 病毒感染的 High Five 细胞的溶解产物降解可溶的硫酸盐标记 HSPG 底物。pFhpa2 病毒(●)或对照 pF<sub>2</sub> 病毒(□)感染的 High Five 细胞的溶解产物与硫酸盐标记的 ECM 衍生的可溶性 HSPG(峰 I )温育(18h, 37℃)。将温育培养基随后加入到 Sepharose 6B 上进行凝胶过滤。低分子量 HS 降解片段(峰 II )仅在与 pFhpa2 感染的细胞温育时产生，而 pF2 感染的细胞的溶解产物不造成 HSPG 底物(◇)的降解。

图 3a-b 表示可溶性的硫酸盐标记 HSPG 底物被 pFhpa2 和 pFhpa4 感染细胞的培养基降解。将 pFhpa2 (3a) 或 pFhpa4 (3b) 病毒(●)或对照病毒(□)感染的 High Five 细胞的培养基与硫酸盐标记的 ECM 衍生的可溶 HSPG(峰 I , ◇)温育(18h, 37℃)。温育培养基随后加至 Sepharose 6B 上进行凝胶过滤。低分子量 HS 降解片段(峰 II )仅在与含 hpa 基因的病毒温育时产生。对照病毒感染细胞的培养基不降解

HSPG 底物。

图 4 表示 pFhpa2 感染的细胞表达的类肝素酶活性的大小分级分离。将 pFhpa2 感染的 High Five 细胞的培养基加到 50kDa 的截留膜上。在高分子量 ( $> 50\text{kDa}$ ) 部分 (●) 而非低分子量 ( $< 50\text{kDa}$  = (○)) 部分中发现有类肝素酶活性 (峰 I 底物 ( $\diamond$ ) 转换为峰 II HS 降解片段)。

图 5a-b 表示肝素对 pFhpa2 和 pFhpa4 感染的 High Five 细胞表达的类肝素酶活性的影响。将 pFhpa2 (5a) 和 pFhpa4 (5b) 感染的 High Five 细胞的培养基与硫酸盐标记的 ECM 衍生的可溶 HSPG (峰 I,  $\diamond$ ) 在缺乏 (●) 或存在 ( $\triangle$ )  $10 \mu\text{g/ml}$  肝素时温育 (18h, 37°C)。在肝素 - 类肝素酶活性的一种强有力抑制剂存在时完全不产生低分子量 HS 降解片段 (6, 7)。

图 6a-b 表示病毒感染的 High Five 和 Sf21 细胞对硫酸盐标记的完整 ECM 的降解。将 High Five (6a) 和 Sf21 (6b) 细胞铺在硫酸盐标记的 ECM 之上, 以 pFhpa4 (●) 或对照 pF1 (□) 病毒感染 (48h, 28°C)。将对照未感染 Sf21 细胞 (R) 同样铺在标记的 ECM 上。培养基的 pH 调至 6.0—6.2, 37°C 温育 24 小时。通过在 Sepharose 6B 上凝胶过滤, 分析释放到温育培养基中的硫酸盐标记物质。仅有含 hpa 的病毒感染的细胞可产生 HS 降解片段。

图 7a-b 表示病毒感染的细胞对硫酸盐标记的完整 ECM 的降解。将 High Five (7a) 和 Sf21 (7b) 细胞铺在硫酸盐标记的 ECM 上, 以 pFhpa4 (●) 或对照 pF1 (□) 病毒感染 (48h, 28°C)。对照未感染的 Sf21 细胞 (R) 也铺在标记的 ECM 上。培养基的 pH 调至 6.0—6.2, 28°C 温育 48 小时。通过在 Sepharose 6B 上进行凝胶过滤, 分析释放到温育培养基中的硫酸盐标记的降解片段。仅有含 hpa 的病毒感染的细胞产生 HS 降解片段。

图 8a-b 表示 pFhpa4 感染细胞的培养基对硫酸盐标记的完整 ECM 的降解。将 pFhpa4 (●) 或对照 pF1 (□) 病毒感染的 High Five (8a) 和 Sf21 (8b) 细胞的培养基与完整的硫酸盐标记 ECM 温育 (48h, 30°C, pH 6.0)。还将 ECM 与对照未感染 Sf21 细胞 (R) 的培养基温育。对释放

到反应混合物中的硫酸盐标记物质进行凝胶过滤分析。仅在 pFhpa4 感染细胞的培养基中测得类肝素酶活性。

图 9a-b 表示肝素对 pFhpa4 感染细胞的培养基中的类肝素酶活性的影响。在缺乏(●)或存在(V) 10 μg/ml 肝素时，将硫酸盐标记的 ECM 与 pFhpa4 感染的 High Five (9a) 和 Sf21 (9b) 细胞的培养基温育(24h, 37°C, pH6.0)。将释放到温育培养基中的硫酸盐标记物质在 Sepharose 6B 上凝胶过滤。肝素存在时，类肝素酶活性(峰 II HS 降解片段的产生)被完全抑制。

图 10a-b 表示在肝素-Sepharose 上纯化重组类肝素酶。将 pFhpa4 病毒感染的 Sf21 细胞的培养基进行肝素-Sepharose 层析。用 NaCl 梯度 0.35—2M (◇) 进行各级分的洗脱。图 10a(●)表示洗脱级分中的类肝素酶活性。对级分 15—28 进行 15% SDS-聚丙烯酰胺凝胶电泳，而后硝酸银染色。证明在级分 19—24 中的主蛋白带(MW ~ 63,000)与类肝素酶活性之间有相关性。

图 11a-b 表示在 Superdex 75 凝胶过滤柱上纯化重组类肝素酶。将从肝素-Sepharose 上洗脱的活性级分(图 10a)集中，浓缩，并加到 Superdex 75 FPLC 柱上。收集各级分并检查各级分的等分样品的类肝素酶活性(C, 图 11a)，由 SDS-聚丙烯酰胺凝胶电泳而后硝酸银染色分析(图 11b)。可见级分 4—7 中的主蛋白带(MW ~ 63,000)的存在与类肝素酶活性间有相关性。

图 12a-e 表示对来自人胚组织(12a)、人胚外组织(12b)和不同起源的细胞系(12c-e)的总 RNA 进行 RT-PCR，示 hpa 基因的表达。使用 hpa 特异性引物(I)、GAPDH 持家基因的引物(II)和无逆转录酶的对照反应的 RT-PCR 产物，证明 RNA 样品(III)中无基因组 DNA 或其它污染。M-DNA 分子量标准 VI (Boehringer Mannheim)。12a：泳道 1—嗜中性细胞(成熟的)，泳道 2—肌肉，泳道 3—胸腺，泳道 4—心脏，泳道 5—肾上腺。12b：泳道 1—肾，泳道 2—胎盘(8 周)，泳道 3—胎盘(11 周)，泳道 4—7—癌(完全水泡状脂块)，泳道 8—细胞滋养层细胞(新分离)，泳道 9—细胞滋养层细胞(体外 1.5h)，泳道 10—细胞滋

养层细胞(体外 6h), 池道 11—细胞滋养层细胞(体外 18h), 池道 12—细胞滋养层细胞(体外 48h)。12c: 池道 1—JAR 膀胱细胞系, 池道 2—NCITT 睾丸肿瘤细胞系, 池道 3—SW-480 人肝癌细胞系, 池道 4—HTR (SV40 转化的细胞滋养层), 池道 5—HPTLP-I 肝细胞癌细胞系, 池道 6—EJ-28 膀胱癌细胞系。12d: 池道 1—SK-hep-1 人肝癌细胞系, 池道 2—DAMI 人巨核细胞系, 池道 3—DAMI 细胞系+PMA, 池道 4—CHRF 细胞系+PMA, 池道 5—CHRF 细胞系。12e: 池道 1—ABAE 牛主动脉内皮细胞, 池道 2—1063 人卵巢细胞系, 池道 3—人乳腺癌 MDA435 细胞系, 池道 4—人乳腺癌 MDA231 细胞系。

图 13 表示人 hpa 的核苷酸序列与小鼠 EST cDNA 片段 (SEQ ID NO: 12) 的比较, 小鼠 EST cDNA 片段与人 hpa 的 3' 末端(起始于 SEQ ID NO: 9 的核苷酸 1066) 80% 同源。在排列的终止密码子下划线。

图 14 表示 hpa 基因的染色体定位。以 hpa 特异 引物扩增后, 将从体细胞杂合体衍生的 DNA 的 PCR 产物与仓鼠、小鼠和人基因组 DNA 的 PCR 产物通过 0.7% 琼脂糖凝胶分离。池道 1—BstE II 消化的 Lambda DNA, 池道 2—无 DNA 对照, 池道 3—29, PCR 扩增产物。池道 3-5—人、小鼠和仓鼠各自的基因组 DNA。池道 6—29, 分别表示染色体 1—22 和 X 与 Y 的人单染色体体细胞杂合体。池道 30—BstE II 消化的 Lambda DNA。仅在池道 5 和 9 观察到约 2.8kb 的扩增产物, 分别代表人基因组 DNA 和衍生自携人染色体 4 的细胞杂合体的 DNA。这些结果证明 hpa 基因位于人染色体 4 上。

#### 优选实施方案的描述

本发明涉及编码具类肝素酶活性的多肽的多核苷酸, 在下文中可互换地称作 hpa、hpa cDNA 或 hpa 基因, 还涉及包括该多核苷酸的载体, 表达类肝素酶的转导细胞和具类肝素酶活性的重组蛋白。

在详细解释本发明的至少一个实施方案前, 必须了解本发明的应用不限于下文说明书或附图所示的构建细节和各成分的排列。本发明能用于其它实施方案或以不同方式实施。而且, 必须了解在此所用的术语和专门名词是为描述目的而不应认作限制。

本发明能用于开发各种疾病的治疗方法，开发这些疾病的诊断方法以及为为基础研究特别是医学和生物学领域提供新工具。

特别地，本发明可用于开发抑制肿瘤细胞转移、炎症和自身免疫的新药物。对编码类肝素酶的 hpa 基因的鉴定使在异源表达系统中生产重组酶成为可能。

进一步地，本发明能用于调节肝素结合性生长因子的生物利用率，针对肝素结合性生长因子(例如 bFGF, VEGF)和细胞因子(IL-8)的细胞反应，细胞与血浆脂蛋白的相互作用，对病毒、原生动物和一些细菌感染的细胞易感性，和神经变性斑的分解。因此重组类肝素酶是创伤愈合、血管发生、再狭窄、动脉粥样硬化、炎症、神经变性疾病(例如 Genstmann-Straussler 综合症，早老性痴呆症，Scrape)，和某些病毒、细菌和原生动物感染的潜在治疗物。重组类肝素酶可作为鱼精蛋白的潜在替代物用于中和血浆肝素。

如在此所用的，术语“调节”包括实质性地抑制、减缓或逆转一种疾病的进展，实质性地改善疾病或状况的临床症状，或实质性地阻止疾病或状况的临床症状的出现。因此“调节子”包括可调节疾病或状况的因子。对病毒、原生动物和细菌感染的调节包括实质性地中断、阻止或减少任何病毒、细菌或原生动物活性和/或病毒、细菌或原生动物生命周期的阶段，或者减少或阻止病毒、细菌或原生动物对目标，如人或低等动物的感染的任何作用。

如在此所用的，术语“创伤”包括对目标躯体的任何部分的任何伤害，包括但不仅限于急性疾病，如在紫外辐射如日灼下过度暴露造成的灼伤、化学灼伤、放射灼伤、热灼伤，因劳动和分娩而对身体组织如会阴造成的伤害，包括在医学方法如外阴切开术过程中持续的损伤，外伤诱导的损伤包括切除，在汽车和其它机械事故中持续的损伤，子弹、刀片及其它武器所致的损伤，和外科术后损伤，以及慢性疾病如褥疮，涉及糖尿病和循环不良的疾病，和所有类型的痤疮，等。

针对重组酶产生的抗类肝素酶抗体，可用于活组织检查标本、血浆样品和体液中微转移、自身免疫损害和肾衰竭的免疫检测和诊断

中。这种抗体也可用作类肝素酶活性的中和剂。

在此报导对编码类肝素酶的人 hpa 基因的克隆和通过转染细胞表达重组类肝素酶。这是首次克隆的哺乳动物类肝素酶基因。

对从人肝癌细胞分离的类肝素酶的纯化制品进行胰酶消化和微量测序。

利用公开的 YGPDVGQPR (SEQ ID NO: 8) 序列筛选 EST 数据库，寻求相应的 DNA 序列的同源物。鉴定到两个紧密相关的 EST 序列并发现二者相同。

两克隆都包括 1020bp 的插入片段，此插入片段包括 973bp 的开放阅读框及随后的 27bp 的 3' 非翻译区和聚腺苷酸尾，但未鉴定到翻译起始位点。

通过对来自胎盘 Marathon RACE cDNA 复合体的 DNA 进行 PCR 扩增，使用根据 EST 克隆序列和复合体的接头选择的引物，克隆缺失的 5' 末端。

获得 900bp 的 PCR 片段，与鉴定到的 3' 编码 EST 克隆部分重叠。连接的 cDNA 片段 (hpa) 长 1721bp (SEQ ID NO: 9)，包括编码具有如图 1 和 SEQ ID NO: 11 所示 543 个氨基酸、计算分子量为 61192 道尔顿的多肽 (SEQ ID NO: 10) 的开放阅读框。

观察到 EST 克隆和 PCR 扩增的 cDNA 之间在 799 位有一个单核苷酸差异 (A → T)。此区别导致单个氨基酸取代 (Tyr 改为 Phe) (图 1)。此外，发表的 EST 序列包括一个未鉴定的核苷酸，对两个 EST 克隆的 DNA 测序将该未鉴定核苷酸分辨为两个核苷酸 (分别为 SEQ ID NO: 9 中 1630 和 1631 位的 G 和 C)。

体外实验中 hpa 基因产物催化硫酸乙酰肝素降解的能力通过运用杆状病毒表达系统在昆虫细胞中表达完整的开放阅读框而检测。

已证明含 hpa 基因的病毒感染的细胞的提取物和条件培养基对可溶性的 ECM 衍生 HSPG 和完整 ECM 均有高水平的硫酸乙酰肝素降解活性，此活性被肝素抑制，而被不含 hpa 基因的类似构建体感染的细胞或非感染细胞则无此活性。

用 RT-PCR 检测 hpa RNA 在不同组织和细胞系中的表达模式。发现其仅在已知具类肝素酶活性的组织和细胞中表达。

使用 Marathon RACE，通过 PCR 扩增，从人 SK-hep1 细胞系克隆延伸的 5' 序列。将 SK-hep1 hpa cDNA 的 5' 延伸序列与从人胎盘分离的 hpa cDNA 序列 (SEQ ID NO: 9) 装配起来。装配的序列包括一个开放阅读框，SEQ ID NO: 13 和 15，它编码如 SEQ ID NO: 14 和 15 所示的 592 个氨基酸的多肽，计算分子量为 66407 道尔顿。此开放阅读框显示指导催化活性的类肝素酶在哺乳动物细胞表达系统中的表达。Western 印迹分析中用抗类肝素酶抗体可检测表达的类肝素酶。

运用一系列单染色体的人/CHO 和人/小鼠体细胞杂合体将人类肝素酶基因定位在人染色体 4 上。因此新分离的类肝素酶序列能用于鉴定铺展的染色体中含人类肝素酶基因的染色体区域。

因此，本发明提供包含编码具类肝素酶催化活性之多肽的多核苷酸序列的多核苷酸片段（单链或双链 DNA 或 RNA）

术语“类肝素酶催化活性”或与其相当的术语“类肝素酶活性”都指对乙酰肝素或硫酸乙酰肝素蛋白多糖底物特异的哺乳动物内切糖苷酶水解活性，这种活性与通过  $\beta$ -消除的方法降解肝素或硫酸乙酰肝素的细菌酶(肝素酶 I, II 和 III)的活性相反(37)。

在本发明优选的实施方案中，多核苷酸片段包括 SEQ ID NO: 9 的核苷酸 63—1691，或 SEQ ID NO: 13 的核苷酸 139—1869，它们编码完整的人类肝素酶。

但本发明的范围不仅限于人类肝素酶，因其是对编码任何哺乳动物类肝素酶的开放阅读框 (ORF) 的首次公开。运用 hpa cDNA、其部分或根据其序列设计的合成寡核苷酸，本领域技术人员能够鉴定到包括来自其它哺乳动物物种的同源序列的基因组和/或 cDNA 克隆。

因此本发明进一步涉及包括能与 hpa cDNA，尤其是与 SEQ ID NO: 9 的核苷酸 1-721 杂交 (严格或允许杂交条件下的碱基配对，参见如 Sambrook, J., Fritsch, E. F., Maniatis. T. (1989) 分子克隆，实验室指南，冷泉港实验室出版社，纽约) 的多核苷酸序列的多核苷

酸片段。

事实上，编码具类肝素酶活性的多肽的任何多核苷酸序列和与 SEQ ID NO: 9 或 13 至少 60% 同源、优选至少 70% 同源、更优选至少 80% 同源、最优选至少 90% 同源的任何多核苷酸序列均在本发明范围内。

根据本发明的多核苷酸片段可包括 SEQ ID NO: 9 或 13 的任何部分。例如，它可包括 SEQ ID NO: 9 的核苷酸 63—721，这是一个新序列。它可包括编码具类肝素酶催化活性之多肽的 SEQ ID NO: 9 或 13 的任何区段。

当在此和以下的权利要求部分应用术语“编码具类肝素酶催化活性的多肽”时，指的是能够指导在例如 ECM 和细胞表面相关 HS 的降解中有催化活性[如其活性所需，也可以是在翻译后修饰（包括但不限于蛋白水解，例如除去信号肽和蛋白原或前蛋白原序列）、甲硫氨酸修饰、糖基化、烷基化（如甲基化）、酰化等后有此活性]的多肽的合成。

在本发明优选的实施方案中，此多核苷酸片段编码的多肽包括 SEQ ID NO: 10 或 14 中所示的氨基酸序列或其功能性部分，即具类肝素酶催化活性的部分。

编码具类肝素酶活性的多肽的任何多核苷酸片段都在本发明范围内。因此，该多肽可为 SEQ ID NO: 10 或 14 所示氨基酸序列或其功能性部分的等位变体、物种变体和/或诱导变体。

事实上，与 SEQ ID NO: 10 或 14 至少 60% 同源、优选至少 70% 同源、更优选至少 80% 同源、最优选至少 90% 同源的任何编码具类肝素酶活性的多肽的多核苷酸序列，均在本发明的范围内。

本发明还提供单链多核苷酸片段，此片段包括与编码具如上述类肝素酶活性之多肽的多核苷酸链的至少一部分互补的多核苷酸序列。术语“互补”在此用于指碱基配对的能力。

此单链多核苷酸片段可为 DNA 或 RNA，甚至包括核苷酸类似物（例如硫代核苷酸），它可为合成的寡核苷酸或由转导的宿主细胞制备，它可为仍能提供特异碱基配对的任何希望的长度（例如长 8 或 10 个核苷酸，优选更多核苷酸），且它可包括不妨碍碱基配对的错配序列。

本发明进一步提供包括编码具类肝素酶催化活性之多肽的多核苷酸序列的载体。

载体可为任何类型。它可为感染细菌的噬菌体或感染真核细胞的病毒。它也可为质粒、噬菌粒、粘粒、杆粒或人工染色体。编码具类肝素酶催化活性的多肽的多核苷酸序列可包括上述的任何多核苷酸片段。

本发明进一步提供包括编码具类肝素酶催化活性之多肽的外源多核苷酸片段的宿主细胞。

外源多核苷酸片段可为以上描述的任何片段。宿主细胞可为任何类型。它可为原核细胞、真核细胞、细胞系或作为生物体一部分的细胞。外源多核苷酸片段可永久或瞬时存在于细胞中。换而言之，按稳定或瞬时转染、转化或转导获得的转导细胞均在本发明范围内。在此运用的术语“外源的”指多核苷酸片段是从外部引入细胞的。其中多核苷酸片段可以任何拷贝数单个存在，它可整合到一或多个染色体的任何部位或作为染色体外物质存在。

本发明进一步提供包括过量表达类肝素酶催化活性的细胞的类肝素酶过量表达系统。细胞可为被包括编码具类肝素酶活性之多肽的多核苷酸序列和指导类肝素酶过量表达的合适的启动子和增强子序列的任何适当载体瞬时或稳定转染或转化的宿主细胞。过量表达细胞也可以是在表达细胞的内源类肝素酶基因上游插入启动子和/或增强子序列(例如通过同源重组)而得到的细胞，此插入将指导从内源基因的过量表达。在以下说明书和权利要求中所用的术语“过量表达”指在其它条件相同下表达水平高于给定细胞通常具有的表达基础水平。

本发明进一步提供包括具类肝素酶催化活性的多肽的重组蛋白。

重组蛋白可通过任何常规的蛋白纯化方法纯化至几乎均质和/或与添加物混合。重组蛋白可运用以上描述的任何细胞制备。重组蛋白可以任何形式存在可为结晶形式、脱水干粉形式或存在于溶液中。重组蛋白可用于获得纯类肝素酶，纯类肝素酶又可用于产生单克隆或多克隆的抗类肝素酶抗体，并作为抗类肝素酶抑制剂或药物筛选试验或

系统中的筛选活性组分。

本发明进一步提供包括具类肝素酶催化活性的重组蛋白作为活性组分的药物组合物。

局部给药的制剂可包括，但不限于：洗液，药膏，凝胶剂，膏状物，栓剂，滴剂，液体，喷雾剂和粉末。常规的水性、粉状或油基药学载体、增稠剂及其它是必要的或所希望的。外被的保险套、斯坦特固定膜、活动的垫及其它医疗装置也有用。事实上，本发明的范围包括含有具类肝素酶催化活性的重组蛋白作为活性组分的任何医疗设备，如医疗装置。

口服的组合物包括粉末或颗粒，水或非水介质中的悬浮液或溶液，粉袋、胶囊或片剂。增稠剂、稀释液、调味品、分散辅助剂、乳化剂或粘合剂也可能合乎需要。

用于胃肠外给药的制剂可包括，但不限于无菌水溶液，其中还可包含缓冲液、稀释液和其它合适添加物。

所用剂量取决于待治疗疾病的严重性和反应性，但通常为每天 1 剂或更多，疗程为几天至几个月或者直到治疗有效或病状成功减低。本领域的普通技术人员能很容易确定最佳剂量、给药方法和重复率。

本发明还提供鉴定铺展的染色体上包含人类肝素酶基因的染色体区的方法。该方法包括以下方法步骤，其中第一步将铺展的染色体(细胞间期或中期)与编码类肝素酶的标记多核苷酸探针杂交。此标记优选荧光标记。第二步中，洗涤铺展的染色体，从而除去过剩的未杂交探针。最后查找与杂交的标记多核苷酸探针相关的信号，其中探测到的信号指示包含人类肝素酶基因的染色体区。本领域的普通技术人员会懂得如何在合适的标记反应中运用在此公开的序列以及如何运用标记的探针在原位杂交中探测包含人类肝素酶基因的染色体区。

现参照下述实施例，这些实施例及上述说明书均以非限制方式一起阐明本发明。

### 实施例

以下方法和实验细节在随后的实施例中引用。

从人肝癌细胞系和人胎盘纯化类肝素酶并进行特性鉴定：选择人肝癌细胞系(SK-hep-1)作为纯化人肿瘤衍生的类肝素酶的来源。基本上如授予 Fuks 的美国专利 No. 5, 362, 641 中所描述的进行纯化，所述的专利全文收入参考文献。简要的，悬浮培养 500 升  $5 \times 10^{11}$  细胞，类肝素酶通过应用以下步骤提纯 240, 000 倍：(i) 在 pH6.0, NaCl 梯度 0.3—1.4M 进行阳离子交换(CM-Sephadex)层析，(ii) 在 pH7.4, 0.1% CHAPS 存在时，NaCl 梯度 0.3—1.1M，进行阳离子交换(CM-Sephadex)层析；(iii) 在 pH7.4, 0.1% CHAPS 存在时，NaCl 梯度 0.35—1.1M，进行肝素-Sepharose 层析；(iv) 在 pH6.0，包含 0.1% CHAPS 和 1M NaCl 的缓冲液中进行 Con A-琼脂糖凝胶层析，用 0.25M  $\alpha$ -甲基甘露糖苷洗脱；和(v)在 pH7.4, 0.1% CHAPS 存在时，用 NaCl 梯度 0.25—1M 进行 HPLC 阳离子交换(Mono-S)层析。

合并活性级分，TCA 沉淀，对沉淀物进行 SDS 聚丙烯酰胺凝胶电泳和/或胰酶消化和反相 HPLC。纯化蛋白的胰酶消化肽通过反向 HPLC (C8 柱)分离，对均质的峰进行氨基酸序列分析。

对纯化的酶进行反相 HPLC，并用氨基酸顺序分析仪(Applied Biosystems)进行 N-末端氨基酸测序。

细胞：如前述由牛眼建立牛角膜内皮细胞(BCEC)的培养物(19, 38)。培养物原代维持在补充 10% 新生小牛血清和 5% FCS 的 DMEM (1g 葡萄糖/升)中。活跃细胞生长阶段，每隔一天加 bFGF (1ng/ml) (13, 14)。

铺有 ECM 的皿的制备：将 BCEC (2—50 代)以  $2 \times 10^5$  细胞/毫升的初始密度铺在 4 孔板中，在无硫酸盐的 Fisher 培养基加 5% 葡聚糖 T-40 中培养 12 天。在接种后第一天和第 5 天加入  $\text{Na}_2^{35}\text{SO}_4$  ( $25 \mu\text{Ci}/\text{ml}$ )，将培养物与标记物温育，不换培养基。用含 0.5% Triton X-100 和 20mM  $\text{NH}_4\text{OH}$  的 PBS 溶解(5 分钟，室温)细胞层，随后 PBS 洗 4 次，暴露内皮下 ECM。ECM 保持完整，无细胞碎片，并牢固附着于组织培养皿的全部区域上(19, 22)。

为制备可溶的硫酸盐标记的蛋白多糖(峰 I 物质)，用胰酶消化

ECM(胰酶  $25\mu\text{g/ml}$ , 6h,  $37^\circ\text{C}$ ), 消化液通过反透析浓缩, 将浓缩物质加到 Sepharose 6B 凝胶过滤柱上。收集得到的高分子量物质( $\text{Kav} < 0.2$ , 峰 I)。结果显示多于 80%的标记物质由硫酸乙酰肝素蛋白多糖组成(11, 39)。

类肝素酶活性: 将细胞( $1 \times 10^6/35\text{mm}$  培养皿)、细胞溶解产物或条件培养基在  $20\text{mM}$  磷酸盐缓冲液( $\text{pH}6.2$ )存在时于  $^{35}\text{S}$  标记的 ECM 上温育(18h,  $37^\circ\text{C}$ )。也将细胞溶解产物和条件培养基与硫酸盐标记的峰 I 物质( $10-20\mu\text{l}$ )温育。收集温育培养基, 离心( $18000\times g$ ,  $4^\circ\text{C}$ , 3 分钟), 在 Sepharose CL-6B 柱( $0.9 \times 30\text{cm}$ )上凝胶过滤分析硫酸盐标记的物质。用 PBS 洗脱分级( $0.2\text{ml}$ ), 流速为  $5\text{ml/h}$ , 用 Bio-fluor 液闪计数放射性。兰色葡聚糖表示已占体积( $V_o$ ), 酚红表示总内含体积( $V_t$ )。后者与自由硫酸根共迁移(7, 11, 23)。HS 侧链的裂解片段在  $0.5 < \text{Kav} < 0.8$ (峰 II)从 Sepharose 6B 洗脱(7, 11, 23)。在  $V_o$  后洗脱( $\text{Kav} < 0.2$ , 峰 I)由胰酶从 ECM 释放的近乎完整的 HSPG, 以及与 PBS 单独温育时较低程度释放的 HSPG。加到柱上的标记物质的回收在不同实验中范围从 85%—95%(11)。每个实验至少进行 3 次, 洗脱位置( $\text{Kav}$  值)变化不超过 $+/- 15\%$ 。

hpa cDNA 的克隆: 从 I. M. A. G. E 协会(2130 Memorial Parkway SW, Hunstville, AL 35801)得到 cDNA 克隆 257548 和 260138。cDNA 最初克隆到质粒载体 pT3T7D-Pac 的 EcoR I 和 Not I 克隆位点。尽管这两个克隆被报导略有区别, 但 DNA 序列分析测定证明它们彼此相同。Marathon RACE(cDNA 末端的快速扩增)人胎盘(多聚腺苷酸)cDNA 复合体由 Tel Aviv 大学 Yossi Shiloh 教授赠送。此复合体不含载体, 包括反转录的 cDNA 片段, 所述片段的两末端均连接了双链和部分单链的衔接子。所用特异复合体的结构在参考文献 39a 中描述。

根据 Clontech 实验室提供的方法, 进行 hp3 PCR 片段的扩增。扩增所用的模板是来自以上复合体的样品。扩增所用的引物为:

第一步: 5' -引物: AP1: 5' -CCATCCTAATACGACTCACTATAGGGC-3' ,  
SEQ ID NO: 1; 3' -引物: HPL229: 5' -GTAGTGATGCCATGTAAGTGAATC-3' ,

SEQ ID NO: 2.

第二步：嵌套 5' -引物: AP2: 5' -ACTCACTATAGGGCTCGAGCGGC -3' , SEQ ID NO: 3; 嵌套 3' -引物: HPL171: 5' -GCATCTTAGCCG TCTTC TTG-3' , SEQ ID NO: 4。HPL229 和 HPL171 是根据 EST 克隆的序列选择的。它们各自包括 SEQ ID NO: 9 的核苷酸 933—956 和 876—897。

PCR 程序为 94℃ 4 分钟, 而后 94℃ 40 秒、62℃ 1 分钟、72℃ 2.5 分钟 30 个循环。用 Expand High Fidelity (Boehringer Mannheim) 进行扩增。将所得约 900bp 的 hp3 PCR 产物用 Bfr I 和 Pvu II 消化。用 EcoRI 消化克隆 257548 (phpa1), 末端补平, 然后 Bfr I 进一步消化。此后将 hp3 PCR 产物的 Pvu II -Bfr I 片段克隆至 phpa1 克隆的平端-Bfr I 端, 得到完整 cDNA 克隆在 pT3T7-pac 中的载体, 称为 phpa2。

DNA 测序: 序列测定用载体特异性和基因特异性引物, 使用自动 DNA 测序仪 (Applied Biosystems 373A 型) 进行。每个核苷酸读自至少两个独立的引物。

序列的计算机分析: 运用 Blast 网络服务器进行数据库检索分析序列相似性。对 DNA 和蛋白序列的序列分析和对比用 Wisconsin 大学 Genetic Computer Group (GCG) 开发的 DNA 序列分析软件包完成。

RT-PCR: 用 TRI-Reagent (Molecular research center Inc.) 根据生产制造商指导制备 RNA。取 1.25 μg RNA 进行逆转录反应, 采用 MuMLV 反转录酶 (Gibco BRL) 和 Oligo (dT) 15 引物, SEQ ID NO: 5, (Promega)。用 Taq 聚合酶 (Promega) 进行所得第一链 cDNA 的扩增。运用以下引物:

HPU-355: 5' -TTCGATCCCAAGAAGGAATCAAC-3' , SEQ ID NO: 6, SEQ ID NO: 9 或 11 中的核苷酸 372—394。

HPL-229: 5' -GTAGTGATGCCATGTAATGAAATC-3' , SEQ ID NO: 7, SEQ ID NO: 9 或 11 中的核苷酸 933—956。

PCR 程序: 94℃ 4 分钟, 然后 94℃ 40 秒、62℃ 1 分钟、72℃ 1 分钟 30 个循环。

重组类肝素酶在昆虫细胞中的表达：将细胞、High Five 和 Sf21 昆虫细胞系在 SF900 II -SFM 培养基 (Gibco BRL) 中作为单层培养物维持培养。

重组的杆状病毒：用 Bac to Bac 系统 (GibcoBRL) 构建包含 hpa 基因的重组病毒。转移载体 pFastBac 用 Sal I 和 Not I 消化，并与 Xho I 和 Not I 消化的 phpa2 的 1.7kb 片段连接。所得质粒称为 pFasthpa2。重复制备命名为 pFasthpa4 的相同质粒，二者独立地应用于进一步的实验中。根据生产制造商的说明，用 pFasthpa2、pFasthpa4 和 pFastBac 产生重组杆粒。后者作为阴性对照。转染重组杆粒 DNA 到 Sf21 昆虫细胞中。转染 5 天收获重组病毒并用其感染 T-25 瓶中的  $3 \times 10^6$  个 High Five 昆虫细胞。感染后 2—3 天后收获细胞。将  $4 \times 10^6$  个细胞离心，重悬于包含 20mM 磷酸柠檬酸盐缓冲液、50mM NaCl 的反应缓冲液中。细胞经过三个冻融循环，溶解产物保存在 -80 °C。条件培养基保存在 4°C。

重组类肝素酶的部分纯化：通过肝素-Sepharose 柱层析，随后通过 Superdex 75 柱凝胶过滤进行重组类肝素酶的部分纯化。对 pFhpa4 病毒感染的 Sf21 细胞的培养基 (150ml) 进行肝素-Sepharose 层析。在含 0.1% CHAPS 和 1mM DTT 的 10mM 乙酸钠缓冲液 (pH5.0) 存在时用 0.35—2M NaCl 梯度洗脱 1ml 级分。检测每级分 25 μl 样品的类肝素酶活性。类肝素酶活性在 0.65—1.1M NaCl 范围洗脱 (级分 18—26, 图 10a)。对每级分中 5 μl 进行 15% SDS 聚丙烯酰胺凝胶电泳，硝酸银染色。合并从肝素-Sepharose 洗脱的活性级分 (图 10a) 并在 YM3 截留膜上浓缩 (X6)。将 0.5ml 的浓缩物质加到 30ml Superdex 75 FPLC 柱上，该柱用含 0.8M NaCl、1mM DTT 和 0.1% CHAPS 的 10mM 乙酸钠缓冲液 (pH5.0) 平衡过。以流速 0.75ml/分钟收集各级分 (0.56ml)。从每级分中取等分检测类肝素酶活性，并进行 SDS-聚丙烯酰胺凝胶电泳，硝酸银染色 (图 11b)。

## 实施例 1

## hpa 基因的克隆

将从人肝癌细胞(SK-hep1)分离的类肝素酶的纯化级分进行胰酶消化和微量测序。筛选 EST(表达序列标记)数据库以寻找相应于所获肽的 DNA 序列的同源物。两个 EST 序列(登记号 N41349 和 N45367)中包含编码肽 YGPDVGQPR (SEQ ID NO: 8) 的 DNA 序列。这两个序列衍生自从 8—9 周胎盘的 cDNA 文库(Soares)制备的克隆 257548 和 260138 (I. M. A. G. E. 协会)。发现两个克隆是相同的，它们包括 1020bp 的一个插入片段，此插入片段包含 973bp 的开放阅读框(ORF)及随后的 27bp 3' 非翻译区和聚腺苷酸尾。在这些克隆的 5' 末端未鉴定到翻译起始位点(AUG)。

通过对胎盘 Marathon RACE cDNA 复合体的 DNA 进行 PCR 扩增，进行对缺失的 5' 末端的克隆。获得 900bp 的片段(命定为 hp3)，它与已鉴定的 3' 编码 EST 克隆部分重叠。

连接的 cDNA 片段长 1721bp (SEQ ID NO: 9)，它包括编码 543 个氨基酸、计算分子量为 61192 道尔顿的多肽(SEQ ID NO: 10)的开放阅读框，如图 1 和 SEQ ID NO: 11 所示。克隆 257548 和 260138 中包括的部分 cDNA 插入片段的 3' 末端起始于 SEQ ID NO: 9 和图 1 的核苷酸 G<sup>721</sup>。

如图 1 中进一步显示，在 EST 克隆和 PCR 扩增序列之间有单个序列差异，导致 EST 中的 Tyr<sup>246</sup> 替换为扩增的 cDNA 中的 Phe<sup>246</sup>。从两个独立的扩增产物证实 PCR 扩增的 cDNA 片段的核苷酸序列。该新基因命名为 hpa。

如上所述，EST 克隆 257548 和 260138 中所含的部分 cDNA 插入片段的 3' 末端起始于 hpa (SEQ ID NO: 9) 的核苷酸 721。运用计算机模拟分析 hpa cDNA 形成稳定二级结构、如包括在 721 位附近的核苷酸序列的茎和环结构的能力。研究发现，在核苷酸 698—724 (SEQ ID NO: 9) 之间很可能形成稳定的茎环结构。此外，hpa 基因的 5' 末端区具高 GC 含量的特点，GC 含量高达 70%，而在 3' 区，GC 含量仅为约 40%。这些发现可解释 EST 克隆的不成熟终止及因此引起的 5' 末端

缺失。

为检测 hpa 基因产物在体外试验中催化硫酸乙酰肝素降解的能力，用杆状病毒表达系统在昆虫细胞中表达完整的开放阅读框。包含 hpa 基因的病毒所感染的细胞的提取物表现有高水平的硫酸乙酰肝素降解活性，而在以不含 hpa 基因的类似构建体感染的细胞及未感染细胞的提取物中则不含此活性。这些结果在随后的实施例中进一步说明。

## 实施例 2

### 可溶的 ECM 衍生的 HSPG 的降解

以包含 pFasthpa 质粒的重组杆状病毒或含无插入片段的质粒的对照病毒感染 High Five 细胞的单层培养物 (72h, 28°C)。收获细胞，在类肝素酶反应缓冲液中经过三个冻融循环使之裂解。随后将细胞裂解物与硫酸盐标记的 ECM 衍生 HSPG(峰 I )温育 (18h, 37°C)，再对反应混和物进行凝胶过滤分析(Sepharose 6B)。

如图 2 所示，底物本身包括在  $V_0$  之后洗脱的几乎所有高分子量 (Mr) 物质(峰 I , 级分 5—20,  $K_{av} < 0.35$ )。当 HSPG 底物与以对照病毒感染的细胞的裂解物温育时，得到相似的洗脱图。与之形成对照，HSPG 底物与以包含 hpa 的病毒感染的细胞的裂解物温育，导致高分子量底物完全转换为低分子量的标记降解片段(峰 II , 级分 22—35,  $0.5 < K_{av} < 0.75$ )。

在峰 II 洗脱的片段显示是硫酸乙酰肝素的降解产物，因为它们 (i) 比用碱性氢硼化物或木瓜蛋白酶处理而从 ECM 释放的完整硫酸乙酰肝素侧链 ( $K_{av}$  近似 0.33) 小 5—6 倍；且 (ii) 抗木瓜蛋白酶或软骨素酶 ABC 的进一步消化，且对亚硝酸脱氨基作用敏感 (6, 11)。

在 Sf21 细胞中获得相似的结果(未给出)。同样，在以含 hpa 的病毒 (pFhpa) 感染的细胞中探测到类肝素酶活性，而在以对照病毒 (pF) 感染的细胞中检测不到。此结果是从两个独立产生的重组病毒获得的。未感染的 High Five 细胞对照的裂解物无法降解 HSPG 底物。

在随后的实验中，将标记的 HSPG 底物与感染的 High Five 或 Sf21 细胞的条件培养基温育。

如图 3a-b 所示，由高 Mr 峰 I 底物向代表 HS 降解片段的低 Mr 峰 II 的转变所反映的类肝素酶活性，发现存在于 pFhpa2 或 pFhpa4 病毒而非对照 pF1 或 pF2 病毒感染的细胞的培养基中。在对照未感染的 High Five 或 Sf21 细胞的培养基中未探测到类肝素酶活性。

将以 pFhpa4 病毒感染的细胞的培养基通过 50kDa 截留膜，从而获得对重组类肝素酶的分子量的粗略估定。如在图 4 中说明，所有的酶活性保留在上层部分，而流过 ( $< 50\text{kDa}$ ) 物质中无活性。此结果与 hpa 基因产物的预期分子量一致。

为进一步表征肝素对 hpa 产物的抑制作用，检查类肝素酶介导的 HS 降解的强抑制剂 (40)。

如图 5a-b 所示，由峰 I 底物向峰 II HS 降解片段的转变在肝素存在时被彻底消除。

综上所述，这些结果表明：被包含新鉴定的人 hpa 基因的杆状病毒感染的昆虫细胞以活性形式表达类肝素酶。

### 实施例 3

#### 完整 ECM 中 HSPG 的降解

接着研究完整的感染昆虫细胞降解完整、天然产生的 ECM 中的 HS 的能力。为此目的，将 High Five 或 Sf21 细胞接种在代谢性地硫酸盐标记的 ECM 上，随后用 pFhpa4 或对照 pF2 病毒感染 (48h, 28°C)。而后将培养基 pH 调到 6.2—6.4，细胞在标记 ECM 中进一步培养 48h (28°C) 或 24h (37°C)。对释放到温育培养基中的硫酸盐标记物质进行 Sepharose 6B 上的凝胶过滤分析。

如图 6a-b 和 7a-b 所示，ECM 与对照 pF2 病毒感染的细胞温育造成标记物质的恒定释放，此标记物几乎全部 (> 90%) 由在 V<sub>o</sub> 之后或与之同洗脱的高 Mr 片段 (峰 I) 组成。以前的研究显示，存在于 ECM 本身和/或由细胞表达的蛋白水解活性负责高 Mr 物质的释放 (6)。此近乎完

整的 HSPG 为类肝素酶进行的随后降解提供了可溶底物，类肝素酶被肝素抑制时相对大量的峰 I 物质积累也说明了这一点(6, 7, 12, 图 9)。另一方面，标记的 ECM 与 pFhpa4 病毒感染的细胞温育导致 60%—70% 的 ECM 相关放射性以低 Mr 的硫酸盐标记片段的形式(峰 II,  $0.5 < Kav < 0.75$ )释放，而不论感染细胞与 ECM 是在 28°C 还是在 37°C 温育。对照完整未感染 Sf21 或 High Five 细胞无法降解 ECM HS 侧链。

在随后的实验中，如图 8a-b 所示，以 pFhpa4 或对照 pF1 病毒感染 High Five 细胞和 Sf21 细胞(96h, 28°C)，将培养基与硫酸盐标记的 ECM 一起温育。仅在与 pFhpa4 感染细胞的条件培养基温育时，才从 ECM 释放低 Mr HS 降解片段。如图 9 所示，肝素存在时，这些片段不再产生。在对照未感染细胞的培养基中，未检测到类肝素酶活性。这些结果说明 pFhpa4 病毒感染的细胞所表达的类肝素酶，当其与天然产生的完整 ECM 的其它大分子成分(即纤连蛋白，层粘连蛋白，胶原蛋白)复合时，能以类似于报导的高度转移肿瘤细胞或免疫系统活化细胞中的方式降解 HS (6, 7)。

#### 实施例 4 重组类肝素酶的纯化

通过肝素-Sepharose 层析从 pFhpa4 感染的 Sf21 细胞的培养基中部分纯化重组类肝素酶(图 10a)，随后对合并活性级分通过 FPLC Superdex 75 柱凝胶过滤(图 11a)。观察到 1 个 63kDa 蛋白，如银染 SDS-聚丙烯酰胺凝胶电泳所探测的，其量与有关柱级分中的类肝素酶活性相关(分别为图 10b 和 11b)。此蛋白在对照 pF1 病毒感染细胞的培养基中未探测到，将其加到肝素-Sepharose 上进行相似分级分离(未示)。

#### 实施例 5 各种细胞类型、器官和组织中 hpa 基因的表达 参考图 12a-e，应用 RT-PCR 估计各种细胞类型和组织中 hpa 基因

的表达。为此目的，反转录总 RNA 并扩增之。在人肾、胎盘(8 和 11 周)和癌组织中，以及在新鲜分离的和短期(1.5—48h)培养的人胎盘细胞滋养层细胞中，清楚显示有预期的 585bp 长的 cDNA(图 12a)，以上细胞均已知表达高类肝素酶活性(41)。正常人嗜中性粒细胞也表达 hpa 转录物(图 12b)。与之形成对照，在胚胎人肌肉组织、胸腺、心脏和肾上腺中，无可检测水平的 hpa mRNA 表达(图 12b)。部分而非所有人类膀胱癌细胞系(图 12c)、SK 肝癌(SK-hep-1)，卵巢癌(OV1063)、乳腺癌(435, 231)、黑素瘤和巨核细胞(DAMI, CHRF)人细胞系(图 12d-e)表达 hpa 基因。

以上描述的 hpa 转录本的表达模式经确定与各种组织和细胞类型中测定的类肝素酶活性水平有极好相关性(未示)。

### 实施例 6 hpa 同源基因

筛选 EST 数据库，寻找与 hpa 基因同源的序列。鉴定到三种小鼠 EST(登记号 Aa177901, 源于小鼠脾, Aa067997 源于小鼠皮肤, Aa47943 源于小鼠胚)，将其装配为包含 629bp 的部分开放阅读框(缺失 5' 末端)和 195bp 的 3' 非翻译区的 824bp cDNA 片段(SEQ ID NO: 12)。如图 13 所示，编码区与 hpa cDNA 序列的 3' 末端 80% 相似。这些 EST 很可能是编码小鼠类肝素酶的小鼠 hpa 同源物的 cDNA 片段。

搜寻共有序列蛋白区，发现类肝素酶和几种前体蛋白如前胶原蛋白 Alpha1 前体，酪氨酸蛋白激酶 RYK, Fibulin-1, 胰岛素样生长因子结合蛋白和几种其它蛋白间在氨基末端有同源性。此氨基末端高度疏水并包含潜在的跨膜区。与已知信号肽序列的同源性说明它能作为信号肽为蛋白定位发挥功能。

### 实施例 7 从人 SK-hep1 细胞系分离 hpa cDNA 的延伸 5' 末端 使用 Marathon RACE (cDNA 末端快速扩增)试剂盒(Clontech)，

通过 PCR 扩增从人 SK-hep1 细胞系分离 hpa cDNA 的 5' 末端。按照生产制造商说明使用 TRI-Reagent (Molecular research center Inc.) 从 SK-hep1 细胞制备总 RNA。使用 mRNA 分离试剂盒(Clonetech)分离 PolyA+ RNA。

按照生产制造商推荐构建 Marathon RACE SK-hep1 cDNA 复合体。使用衔接子特异引物 AP1: 5' -CCATCCTAACATACGACTCACTATAGGGC-3' (SEQ ID NO: 1), 和相应于 SEQ ID NO: 9 核苷酸 119-99 的 hpa 特异性反义引物 hp1-629: 5' -CCCCAGGAGCAGCAGCATCAG-3' (SEQ ID NO: 17), 进行第一轮扩增。所得 PCR 产物进行第二循环扩增, 使用衔接子特异性嵌套引物 AP2: 5' -ACTCACTATAGGGCTC GAGCGGC-3' , SEQ ID NO: 3, 和相应于 SEQ ID NO: 9 核苷酸 83-63 的 hpa 特异性反义嵌套引物 hp1-666 5' -AGGCTTCGAGCGCAGCAGCAT-3' , SEQ ID NO: 18. PCR 过程如下: 热起始 94°C 1 分钟, 随后 30 个循环: 90°C 30 秒, 68°C 4 分钟。从琼脂糖胶提取所获的 300bp DNA 片段, 将其克隆到载体 pGEM-T Easy (Promega) 中。获得的重组质粒命名为 pHPSK1.

确定 pHPSK1 插入片段的核苷酸序列, 发现其包括胎盘 hpa cDNA (SEQ ID NO: 9) 5' 末端的 62 个核苷酸和上游的额外 178 个核苷酸, 即 SEQ ID NO: 13 和 15 的头 178 个核苷酸。

鉴定到 SK-hep1 cDNA 和胎盘 cDNA 之间有单个核苷酸差异。胎盘 cDNA (SEQ ID NO: 9) 第 9 位的 “T” 衍生物, 在 SK-hep1 cDNA (SEQ ID NO: 13) 的相应第 187 位被 “C” 衍生物替换。

此差异可能是由于在胎盘 cDNA 克隆的 5' 末端有突变, 通过对从胎盘分离的几个额外 cDNA 克隆(它们如 SK-hep1 cDNA 一样在 SEQ ID NO: 9 的第 9 位包括 C)进行序列分析证实了这一点。

将 SK-hep1 hpa cDNA 的 5' 延伸序列与从人胎盘分离的 hpa cDNA 序列 (SEQ ID NO: 9) 装配起来。装配后序列包括编码如 SEQ ID NO: 14 和 15 中所示计算分子量为 66407 道尔顿的 592 个氨基酸的多肽的开放阅读框。该开放阅读框侧接有 93bp 的 5' 非翻译区 (UTR)。

## 实施例 8

### hpa 基因的上游基因组区的分离

使用 Genome Walker 试剂盒 (Clontech)，按照生产制造商说明分离 hpa 基因的上游区。此试剂盒包括 5 种人基因组 DNA 样品，每个样品用不同限制性核酸内切酶消化产生平末端：EcoRV, Sca I, Dra I, Pvu II 和 Ssp I。

将平末端 DNA 片端与部分单链的衔接子连接。使用衔接子特异性引物和基因特异性引物，对基因组 DNA 样品进行 PCR 扩增。使用 Expand High Fidelity (Boehringer Mannheim) 扩增。

第一循环扩增使用 ap1 引物：5' -GTAATACGACTCACTATAGGGC-3'，SEQ ID NO: 19，和相应于 SEQ ID NO: 9 核苷酸 83—63 的 hpa 特异性反义引物 hpl-666: 5' -AGGCTTCG AGCGCAGCAGCAT-3'，SEQ ID NO: 18。PCR 程序如下：热起始 94℃ 3 分钟，然后 36 个循环，每循环为 94℃ 40 秒 67℃ 4 分钟。

将第一次扩增的 PCR 产物按 1:50 稀释。将稀释样品的 1 μl 用作第二次扩增的模板，其中使用嵌套的衔接子特异性引物 ap2: 5' -ACTATA GGGCACGCGTGGT-3'，SEQ ID NO: 20，和相应于 SEQ ID NO: 9 核苷酸 62—42 的 hpa 特异性反义引物 hpl-690，5' -CTTGGGCTCACCTGGCTGCTC-3'，SEQ ID NO: 21。所得的扩增产物用琼脂糖凝胶电泳分析。从 5 个扩增反应获得 5 种不同的 PCR 产物。从凝胶中提取 Ssp I 消化 DNA 样品所得的约 750bp 的 DNA 片段。将纯化的片段连接到质粒载体 pGEM-T Easy (Promega) 上。所获的重组质粒命名为 pGHP6905，确定 hpa 插入片段的核苷酸序列。

SEQ ID NO: 16 中给出了 594 个核苷酸的部分序列。SEQ ID NO: 16 的最后一位核苷酸相应于 SEQ ID NO: 13 的核苷酸 93。SEQ ID NO: 16 中的 DNA 序列包括 hpa cDNA 的 5' 区和被预测为包括 hpa 基因启动子区的基因组上游区 501 个核苷酸。

## 实施例 9

## 592 个氨基酸的 HPA 多肽在人 293 细胞系中的表达

通过将相当于 SK-hep1 hpa cDNA 5' 末端的 110bp 与胎盘 cDNA 相连，构建 592 个氨基酸的开放阅读框 (SEQ ID NO: 13 和 15)。更具体而言，将胎盘 hpa DNA 的 Marathon RACE-PCR 扩增产物用 Sac I 消化，并将近似 1kb 片段连接到 Sac I 消化的 pGHP6905 质粒中。所得质粒用 Ear I 和 Aat II 消化，钝化 Ear I 粘性末端，分离出约 280bp 的 Ear I / 已钝化 Aat II 片段。使用 Klenow 片段补平 EcoR I 消化的 pFasthpa，并进一步用 Aat II 消化，将所得 pFasthpa 与上述片段连接。所获质粒包括 1827bp 的插入片段，此插入片段包括 1776bp 的开放阅读框，31bp 的 3' UTR 和 21bp 的 5' UTR。此质粒命名为 pFastLhpa。

构建哺乳动物表达载体以驱动 592 个氨基酸的类肝素酶多肽在人细胞中表达。用 BssHII 和 Not I 从 pFastLhpa 切下 hpa cDNA。将所得 1850bp 的 BssHII I -Not I 片段连接到 Mlu I 和 Not I 消化的哺乳动物表达载体 pSI (Promega) 中。将所得重组质粒 pSIhpaMet2 转染到人 293 胚胎肾细胞系中。

592 个氨基酸的类肝素酶的瞬时表达通过 Western 印迹分析检测，使用凝胶移位试验检查酶活性。这些操作在 1998 年 5 月 1 日提交的美国专利申请 NO. 09/071,739 中有详细描述，此申请在此全文纳入参考文献。转染后 3 天收获细胞。将收获的细胞重悬于包括 150mM NaCl, 50mM Tris pH7.5, 1% Triton X-100, 1mM PMSF 和蛋白酶抑制剂混和物 (Boehringer Mannheim) 的裂解缓冲液中。对 40 μg 蛋白提取物样品进行 SDS-PAGE 分离。将蛋白转到 PVDF Hybond-P 膜 (Amersham) 上。将膜与亲合纯化的多克隆抗类肝素酶抗体温育 (如美国专利申请号 09/071,739 中所述)。在转染细胞中观察到约 50kDa 的主带及约 65kDa 的次带。在 pShpa 转染的细胞的提取物中观察到相似的图 (如美国专利申请号 09/071,739 中所述)。这两条带很可能代表转染细胞产生的重组类肝素酶蛋白的两种形式。65kDa 蛋白很可能代表类肝素酶前体，而 50kDa 蛋白在此推测为已加工或成熟形式。

pShpaMet2 转染细胞中表达的重组蛋白的催化活性通过凝胶移位试验检测。转染细胞和模拟转染细胞的提取物在含 20mM 磷酸柠檬酸盐缓冲液 (pH5.4), 1mM CaCl<sub>2</sub>, 1mM DTT 和 50mM NaCl 的反应体系中于 37°C 与肝素 (每反应 6 μg) 温育过夜。然后将反应混合物在 10% 聚丙烯酰胺凝胶上分离。与转染细胞提取物温育的肝素分子与对照相比快速迁移，这清楚表明了重组类肝素酶的催化活性。快速迁移说明高分子量肝素分子的消失和低分子量降解产物的产生。

### 实施例 10 hpa 基因的染色体定位

利用从 UK HGMP Resource Center (Cambridge, England) 获得的单染色体人/CHO 和人/小鼠体细胞杂合体板对 hpa 基因进行染色体作图。

对 40ng 每种体细胞杂合体 DNA 样品进行 PCR 扩增，使用相应于 SEQ ID NO: 9 核苷酸 564—586 的 hpa 引物：hpa565 5' - AGCTCTGTAGATGTGCTATAACAC-3' (SEQ ID NO: 22)，和相应于 SEQ ID NO: 9 核苷酸 897—876 的反义引物 hpl 171 5' - GCATCTTAGCCGTCTTCTTCG-3' (SEQ ID NO: 23)。

PCR 程序如下：94°C 3 分钟热起始，然后 7 个循环：94°C 45 秒、66°C 1 分钟、68°C 5 分钟，随后 30 个循环：94°C 45 秒、62°C 1 分钟、68°C 5 分钟、最后 72°C 10 分钟延伸。

反应用 Expand long PCR (Boehringer Mannheim) 进行。所得扩增产物用琼脂糖凝胶电泳分析。如图 14 所述，从染色体 4 及对照人基因组 DNA 获一条约 2.8kb 的带。2.8kb 的扩增产物被认为是基于对基因组 hpa 克隆的扩增 (未示数据)。在小鼠和仓鼠的对照 DNA 样品和其他人染色体的体细胞杂合体中，都未获得扩增产物。

尽管结合具体实施方案描述了本发明，但对本领域技术人员而言，显然还有许多变换方式、修改和变动。所以，本发明包括落入所附权利要求精神和宽范围内的所有这些变换方式、修改和变动。

## 上文按数字引用的参考文献列表

1. Wight, T. N., Kinsella, M. G., 和 Qawrnstromn. E. E. (1992), 蛋白多糖在细胞粘附, 迁移和增殖中的作用。 *Curr. Opin. Cell Biol.* 4, 793, 801
2. Jackson, R. L., Busch, S. J 和 Cardin, A. L. (1991), 糖胺聚糖: 分子特性, 蛋白相互作用及在生理过程中的作用。 *Physiol. Rev.*, 71, 481-539.
3. Wight, T. N. (1989) 动脉蛋白多糖的细胞生物学, *Arteriosclerosis*, 9, 1-20
4. Kjellen, L., 和 Lindahl, U. (1991) 蛋白多糖: 结构和相互作用。生物化学年度综述 (*Annu. Rev. Biochem.* ), 60:443-475
5. Ruoslahti, E., 和 Yamaguchi, Y. (1991) 蛋白多糖作为生长因子活性的调节子。 *细胞 (Cell)* , 64, 867-869
6. Vlodavsky, I., Eldor, A., Haimovitz-Friedman, A., Matzner, Y., Ishai-Michaeli, R., Levi, E., Basbkin, P., Lider, O., Naparstek, Y., Cohen, I. R., 和 Fuks, Z. (1992) 类肝素酶通过血小板和免疫系统的循环细胞的表达: 在血细胞渗出和外渗中的可能参与 *Invasion & Metastasis*, 12:112-127
7. Vlodasky, I., Mohsen, M., Lider, O., Ishai-Michaeli, R., Ekre, H.-P., Svahn, C. M., Vigoda, M., 和 Peretz, T. (1995). 类肝素酶抑制型肝素对肿瘤转移的抑制。 *Invasion & Metastasis*, 14, 290-302.
8. Nakajima, M., Irimura, T., 和 Nicolson, G. L. (1988). 类肝素酶和肿瘤转移, *细胞与生物化学杂志 (J. Cell. Biochem.)* , 36:157-167
9. Nicolson, G. L (1988). 肿瘤转移的器官特异性: 在特异二级位点恶性细胞优先粘附, 侵入和生长的作用, *Cancer Met. Rev.* 7, 143-188.

10. Liotta, L. A., Rao, C. N., Barsky, S. H. (1983). 肿瘤侵入和细胞外基质, *Lab. Invest.*, 49, 639-649.
11. Vlodavsky, I., Fuks, Z., Bar-Ner, M., Ariav, Y., 和 Schirrmacher, V. (1983). 内皮下细胞外基质中淋巴瘤细胞调节的硫酸盐蛋白多糖的降解: 与肿瘤细胞转移的关系。*癌症研究 (Cancer Res.)* 43, 2704-2711.
12. Vlodavsky, I., Ishai-Michaeli, R., Bor-Ner, M., Fridman, R., Horowitz, A. T., Fuks, Z. 和 Biran, S. (1988). 类肝素酶参与肿瘤转移和血管发生, *Isr. J. Med.*, 24, 464-470.
13. Vlodavsky, I., Lin, G. M., 和 Gospodarowicz D. (1980). 在各种形态的细胞外基质上维持的人肿瘤细胞的形态外观, 生长行为和迁移活性。*细胞*, 19, 607-616.
14. Gospodarowicz, D., Delgado, D., 和 Vlodavsky, I. (1980). 细胞外基质对体外细胞增殖的允许作用。*美国国家科学院院报 (Proc. Natl. Acad. Sci. USA.)*, 77, 4094-4098.
15. Bashkin, P., Doctrow, S., Klagsbun, M., Svahn, C. M., Folkman, J., 和 Vlodavsky, I. (1989) 碱性成纤维细胞生长因子与内皮下细胞外基质结合并被类肝素酶和肝素样分子释放。*生物化学 (Biochemistry)*, 28, 1737-1743.
16. Parish, C. R., Coombe, D. R., Jakobsen, K. B., 和 Underwood, P. A. (1987). 硫酸盐多糖通过阻断肿瘤细胞衍生的类肝素酶抑制肿瘤转移的证据。*Int. J. Cancer*, 40:511-517.
- 16a. Vlodavsky, I., Hua-Quan Miao., Benezra, M., Lider, O., Bar-Shavit, R., Schmidt, A., 和 Peretz, T. (1997). 细胞外基质, 硫酸乙酰肝素蛋白多糖和硫酸乙酰肝素降解酶在血管发生和转移中的参与, 《肿瘤血管发生》。C. E. Lewis, R. Bickwell & N. Ferrara 编. Oxford University Press. Oxford UK, pp. 125-140
17. Burgess, W. H., 和 Maciag, T. (1989) 肝素结合的(成纤维细胞)生长因子蛋白家族, *生物化学年度综述*, 58, 575-606.

18. Folkman, J., 和 Klagsbrun, M. (1987). 血管生成因子。科学 (Science), 235, 442-447
9. Vlodavsky, I., Folkman, J., Sullivan, R., Friolman, R., Ishai-Michaelli, R., Sasse, J., 和 Klagsburn, M. (1987). 内皮细胞衍生的碱性成纤维细胞生长因子: 合成并沉积到内皮下细胞外基质。美国国家科学院院报, 84, 2292-2296
20. Folkman, J., Klagsbrun, M., Sasse, J., Wadzinski, M., Ingber, D., 和 Vlodavsky, I. (1980). 肝素结合的血管形成蛋白—碱性成纤维细胞生成因子贮存在基底膜。Am. J. Pathol., 130, 393-400.
21. Cardon-Cardo, C., Vlodavsky, I., Haimovitz-Friedman, A., Hicklin, D., 和 Fuks, Z (1990). 在正常人组织中碱性成纤维细胞生长因子的表达。Lab. Invest., 63, 832-840.
22. Ishai-Michaeli, R., Svahn, C.-M., Chajek-Shaul, T., Korner, G., Ekre, H.-P., 和 Vlodavsky, I. (1992). 从血管内皮和细胞外基质释放碱性成纤维细胞因子过程中肝素的大小和硫酸盐化的重要性。生物化学, 31, 2080-2088.
23. Ishai-Michaeli, R., Eldor, A., 和 Vlodavsky, I. (1990). 血小板, 嗜中性粒细胞和淋巴瘤细胞表达的类肝素酶活性从细胞外基质释放活化的成纤维细胞生长因子。Cell Reg., 1, 833-842.
24. Vlodavsky, I., Bar-Shavit, R., Ishai-Michaeli, R., Bashkin, P., 和 Fuks, Z. (1991). 成纤维细胞生长因子的细胞外鳌合和释放: 一种调控机制? Trends Biochem. Sci. 16, 268-271
25. Vlodavsky, I., Bar-Shavit, R., Korner, G., 和 Fuks, Z. (1993). 细胞外基质结合的生长因子, 酶和血浆蛋白。《基底膜: 细胞和分子方面》(D. H. Rohrbach 和 R. Timpl 编), pp327-343. Academic Press Inc. Orlando. Fl.
26. Yayon, A., Klagsbrun, M., Esko, J. D. Leder, P., 和 Ornitz, D. M. (1991). 细胞表面碱性成纤维细胞生长因子与其高亲和性受体

- 结合需肝素样分子。细胞 64, 841-848.
27. Spivak-Kroizman, T., Lemmon, M. A., Dikic, I., Ladbury, J. E., Pinchasi, D., Huang, J., Jaye, M., Crumley, G., Schlessinger, J., 和 Lax, I. (1994) 肝素诱导的 FGF 分子的寡聚化是 FGF 受体二聚化, 活化和细胞增殖的根源, 细胞 79, 1015-1024.
28. Ornitz, D. M., Herr, A. B., Nilsson, M., West, a., J., Svahn, C.-M., 和 Waksman, G. (1995). 合成的类肝素衍生的二糖和三糖引起的 FGF 结合及 FGF 受体活化。科学, 268, 432-436.
29. Gitay-Goren, H., Soker, S., Vlodavsky, I., 和 Neufeld, G. (1992) 血管内皮生长因子(VEGF)与细胞表面受体结合需要细胞表面相关的肝素样分子。生物化学杂志, 267, 6093-6098
30. Lider, O., Baharav, E., Mekori, Y., Miller, T., Naparstek, Y., Vlodavsky, I., 和 Cohen, I. R. (1989). T 淋巴细胞类肝素酶的类肝素抑制剂处理动物造成的对实验性自身免疫疾病的抑制和异体移植存活期延长。临床研究杂志(J. Clin. Invest.), 83, 752-756
31. Lider, O., Cahalon, L., Gilat, D., Hershkovitz, R., Siegel, D., Margalit, R., Shoseyov, O., 和 Cohn, I. R. (1995). 抑制肿瘤坏死因子 $\alpha$  的二糖由类肝素酶从细胞外基质形成。美国国家科学院院报, 92, 5037-5041
- 31a. Rapraeger, A., Krufka, A., 和 Olwin, B. R. (1991). bFGF 介导的成纤维细胞生长和成肌细胞分化中对硫酸乙酰肝素的需求。科学, 252, 1705-1708.
32. Eisenberg, S., Sehayek, E., Olivecrona, T., 和 Vlodavsky, I. (1992) 脂蛋白脂肪酶增强脂蛋白对细胞表面和细胞外基质的硫酸乙酰肝素的结合。临床研究杂志 (J. Clin. Invest.), 90, 2013-2021.
33. Shieh, M-T., Wundunn, D., Montgomery, R. I., Esko, J. D 和 Spear, P. G. J (1992). 单纯疱疹病毒的细胞表面受体是硫酸乙酰肝素蛋白多糖。细胞生物学杂志 116, 1273-1283.

- 33a. Chen, Y., Maguire, T., Hileman, R. E., Fromm, J. R., Esko, J. D., Linhardt, R. J., 和 Marks, R. M. (1997). 登革热病毒的侵染性取决于包膜蛋白与靶细胞硫酸乙酰肝素的结合。Nature Medicine 3, 866–871
- 33b. Putnak, J. R., Kanessa-Thasan, N., 和 Innis, B. L. (1997). 登革热病毒推定的细胞受体。Nature Medicine 3, 828–829.
34. Narindrasorasak, S., Lowery, D., Gonzalez-DeWhitt, P., Poorman, R. A., Greenberg, B., Kisilevsky, R. (1991) 阿尔茨海默 $\beta$ -淀粉样前体蛋白和硫酸乙酰肝素蛋白多糖的基底膜型间的高亲合性相互作用。生物化学杂志, 266, 12878–83
35. Ross, R. (1993) 动脉粥样硬化的发病机理：对二十世纪九十年代的透视。自然 (Lond.) ., 362:801–809
36. Zhong-Sheng, J., Walter, J., Brecht, R., Miranda, D., Mahmood Hussain, M., Innerarity, T. L. 和 Mahley, W. R. (1993) 硫酸乙酰肝素蛋白多糖在载脂蛋白 E 丰富的残迹脂蛋白被培养细胞结合和摄取中的作用。生物化学杂志, 268, 10160–10167.
37. Ernst, S., Langer, R., Cooney, Ch. L., 和 Sasisekharan, R. (1995). 糖胺聚糖的酶降解. Critical Reviews in Biochemistry and Molecular Biology, 30(5), 387–444
38. Gospodarowicz, D., Mescher, AL., Birdwell, CR. (1977). 成纤维细胞和表皮生长因子体外对角膜内皮细胞增殖的刺激。Exp Eye Res 25, 75–89.
39. Haimovitz-Friedman, A., Falcone, D. J., Elder, A., Schirrmacher, V., Vlodavsky, I., 和 Fuks, Z (1991) 肿瘤细胞衍生因子造成的血小板类肝素酶的活化。血液, 78, 789–796.
- 39a. Savitsky, K., Platzer, M., Uziel, T., Gilad, S., Sartiel, A., Rosenthal, A., Elroy-Stein, D., Siloh, Y. 和 Rotman, G. (1997). 共济失调一毛细血管扩张: 非翻译序列的结构多样性表明 ATM 基表达的复杂的转录后调节。核酸研究 25(9), 1678–1684.

40. Bar-Ner, M., Eldor, A., Wasserman, L., Matzuer, Y., 和 Vlodavsky, I. (1987) 修饰的和非抗凝肝素种对类肝素酶介导的细胞外基质硫酸乙酰肝素的降解的抑制。血液, 70, 551-557.
- 41 . Goshen, R., Hochberg, A., Korner, G., Levi, E., Ishai-Michaeli, R., Elkin, M., de Grot, N., 和 Vlodavsky, I. (1996) 胎盘类肝素酶的纯化和鉴定及其在培养的细胞滋养层的表达。Mol. Human Reprod, 2, 679-684.

## 序列表

### (1) 一般信息:

(i) 申请人 Iris Pecker, Israel Vlodavsky 和 Elena Freinstein

(ii) 发明名称 编码具类肝素酶活性的多肽的多核苷酸及其在转导细胞中的表达

(iii) 序列数 23

(iv) 通信地址

(A) 地址: Mark M. Friedman c/o Robert Sheinbein

(B) 街: 2940 Birchtree Lane

(C) 城市: Silver Spring

(D) 州: Maryland

(E) 国家: 美国

(F) 邮政编码: 20906

(v) 机读形式:

(A) 介质类型: 1.44M, 3.5 吋软盘

(B) 计算机: Twinhead\* Slimnote-890TX

(C) 操作系统: MS DOS6.2 版

视窗 3.11 版

(D) 软件: 相应视窗 2.0 的 Word, 转化为 ASCII 文件

(vi) 当前申请信息:

(A) 申请号:

(B) 申请日:

(C) 分类:

(vii) 在先申请信息:

(A) 申请号: 08/922, 170

(B) 申请日: 1997 年 9 月 2 日

(viii) 律师/代理人信息:

(A) 名字: Friedmam, Mark M.

(B) 登记号: 33, 883

(C) 参考/摘要号: 910/1

(ix) 电信信息:

(A) 电话: 972-3-5625553

(B) 传真: 972-3-5625554

(C) 电报:

(2) SEQ ID NO: 1 的信息:

(i) 序列特征:

(A) 长度: 27

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 1

CCATCCTAAT ACGACTCACT ATAGGGC 27

(2) SEQ ID NO: 2 的信息:

(i) 序列特征:

(A) 长度: 24

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 2

GTAGTGATGC CATGTAAC TG AAC 24

(2) SEQ ID NO: 3 的信息:

(i) 序列特征:

(A) 长度: 23

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 3

ACTCACTATA GGGCTCGAGC GGC 23

(2) SEQ ID NO: 4 的信息:

(i) 序列特征:

- (A) 长度: 22
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 4

GCATCTTAGC CGTCTTCCTT CG 22

(2) SEQ ID NO: 5 的信息:

(i) 序列特征:

- (A) 长度: 15
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 5

TTTTTTTTTTTT TTTTTT 15

(2) SEQ ID NO: 6 的信息:

(i) 序列特征:

- (A) 长度: 23
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 6

TTCGATCCCA AGAAGGAATC AAC 23

(2) SEQ ID NO: 7 的信息:

(i) 序列特征:

- (A) 长度: 24
- (B) 类型: 核酸

(C) 链型: 单链

(D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 7

GTAGTGATGC CATGTAAC TG AATC 24

(2) SEQ ID NO: 8 的信息:

(i) 序列特征:

(A) 长度: 9

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 8

Tyr Gly Pro Asp Val Gly Gln Pro Arg  
5 9

(2) SEQ ID NO: 9 的信息:

(i) 序列特征:

(A) 长度: 1721

(B) 类型: 核酸

(C) 链型: 双链

(D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 9

CTAGAGCTT CGACTCTCCG CTGCGCGGCA GCTGGCGGGG GGAGCAGCCA GGTGAGCCCCA 60  
AGATGCTGCT GCGCTCGAAG CCTGCGCTGC CGCCGCCGCT GATGCTGCTG CTCCCTGGGC 120  
CGCTGGGTCT CCTCTCCCCCT GGGCCCTGTC CCCGACCTGTC GCAAGCACAG GACGTCGTGG 180  
ACCTGGCTT CCTCACCCAG GAGCCCTGTC ACCTGGTGAG CCCCCTCGTTC CTGTCCTGTC 240  
CCATTGACGC CAACCTGAGC AGGACCCGCG GGTTCCTCAT CCTCCTGGGT TCTCCAAGC 300  
TTCTGACCTT GGCCAGAGGC TTGCTCTGT CGTACCTGAG GTTGGTGGC ACCAAAGACAG 360  
ACTTCCTAAAT TTTCGATCCC AAGAAGGAAT CAACCTTTGA AGAGAGAAAGT TACTGGCAAT 420  
CTCAAGTCAA CCAGGATATT TGCAAATATG GATCCATCCC TCCTGATGTG GAGGAGAAAGT 480  
TACGGTTGGA ATGCCCTAC CAGGAGCAAT TGCTACTCCG AGAACACTAC CAGAAAAAAAGT 540  
TCAAGAACAG CACCTACTCA AGAACGCTCG TAGATGTGCT ATACACTTTT GCAAACATGCT 600  
CAGGACTGGG CTTGATCTT GGCTAAATG CGTTTATAAG AACAGCAGAT TTGCACTGGGA 660  
ACAGTTCTAA TGCTCAGTTC CTCTCTGACT ACTGCTCTTC CAAGGGGTAT AACATTTCTT 720  
GGGAACCTAGG CAATGAACCT AACAGTTCC TAAAGAAGGC TGATATTTTC ATCAATGGGT 780  
CGCAGTTAGG AGAAAGATTAT ATTCATTTGC ATAACACTCTT AAGAAAAGTC ACCTTCAAAAA 840  
ATGCAAAACT CTATGGCTCT GATGTTGGTC AGGCTCGAAG AAAGACGGCT AAGATGCTGA 900  
AGACGTTCTT GAAGGCTGGT GGAGAAGTGA TTGATTCACT TACATGGCAT CACTACTATT 960  
TGAATGGACG GACTGCTACC AGGGAAAGTT TTCTAAACCC TGATGTATTG GACATTTTTA 1020  
TTTCATCTGT GCAGAAAGTT TTCCAGGTGG TTGAGAGGCAC CAGGCTGCGC AAAAGGGTCT 1080  
GGTTAGGAGA AACAAAGCTCT GCATATGGAG GCGGAGGCC CTTGCTATCC GACACCTTTG 1140  
CAGCTGGCTT TATGTGGCTG GATAAAATTGG GCCTGTCAGC CGGAATGGGA ATAGAAGTGG 1200  
TGATGAGGCA AGTATTCCTT GGAGCAGGAA ACTACCAATT AGTGGATGAA AACCTTCGATC 1260  
CTTACCTGA TTATGGCTA TCTCTCTGT TCAAGAAATT GGTGGGCACC AAGGTGTTAA 1320  
TGGCAAGCGT GCAAGGTTCA AAGAGAAGGA AGCTTCGAGT ATACCTTCAT TGCACAAACA 1380  
CTGACAATCC AAGGTATAAA GAAGGAGATT TAACTCTGT A TCCATAAAC CTCCATAACG 1440

TCACCAAGTA CTTGGGTTA CCCTATCCTT TTYCTAACAA GCAAGTGGAT AAATACCTTC 1500  
 TAAGACCTTT GGGACCTCAT GGATTACTTT CCAAATCTGT CCAACTCAAT GGCTAACTC 1560  
 TAAAGATGGT GGATGATCAA ACCTTGCCAC CTTTAATGGA AAAACCTCTC CGGCCAGGAA 1620  
 GTTCACTGGG CTTGCCAGCT TTCTCATATA GTTTTTTGT GATAAGAAAT GCCAAAGTTG 1680  
 CTGCTTGCAT CTGAAAATAA AATATACTAG TCCTGACACT G 1721

(2) SEQ ID NO: 10 的信息:

(i) 序列特征:

- (A) 长度: 543
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 10

Met	Leu	Leu	Arg	Ser	Lys	Pro	Ala	Leu	Pro	Pro	Pro	Leu	Met	Leu	Leu
5								10							15
Leu	Leu	Gly	Pro	Leu	Gly	Pro	Leu	Ser	Pro	Gly	Ala	Leu	Pro	Arg	Pro
20								25							30
Ala	Gln	Ala	Gln	Asp	Val	Val	Asp	Leu	Asp	Phe	Phe	Thr	Gln	Glu	Pro
35								40							45
Leu	His	Leu	Val	Ser	Pro	Ser	Phe	Leu	Ser	Val	Thr	Ile	Asp	Ala	Asn
50								55							60
Leu	Ala	Thr	Asp	Pro	Arg	Phe	Leu	Ile	Leu	Leu	Gly	Ser	Pro	Lys	Leu
65								70							80
Arg	Thr	Leu	Ala	Arg	Gly	Leu	Ser	Pro	Ala	Tyr	Leu	Arg	Phe	Gly	Gly
85								90							95
Thr	Lys	Thr	Asp	Leu	Ile	Phe	Asp	Pro	Lys	Lys	Glu	Ser	Thr	Phe	
100								105							110
Glu	Glu	Arg	Ser	Tyr	Trp	Gln	Ser	Gln	Val	Asn	Gln	Asp	Ile	Cys	Lys
115								120							125
Tyr	Gly	Ser	Ile	Pro	Pro	Asp	Val	Glu	Glu	Lys	Leu	Arg	Leu	Glu	Trp
130								135							140
Pro	Tyr	Gln	Glu	Gln	Leu	Leu	Arg	Glu	His	Tyr	Gln	Lys	Lys	Phe	
145								150							160
Lys	Asn	Ser	Thr	Tyr	Ser	Arg	Ser	Ser	Val	Asp	Val	Leu	Tyr	Thr	Phe
165								170							175
Ala	Asn	Cys	Ser	Gly	Leu	Asp	Leu	Ile	Phe	Gly	Leu	Asn	Ala	Leu	Leu
180								185							190
Arg	Thr	Ala	Asp	Leu	Gln	Trp	Asn	Ser	Ser	Asn	Ala	Gln	Leu	Leu	Leu
195								200							205
Asp	Tyr	Cys	Ser	Ser	Lys	Gly	Tyr	Asn	Ile	Ser	Trp	Glu	Leu	Gly	Asn
210								215							220
Glu	Pro	Asn	Ser	Phe	Leu	Lys	Ala	Asp	Ile	Phe	Ile	Asn	Gly	Ser	
225								230							240
Gln	Leu	Gly	Glu	Asp	Tyr	Ile	Gln	Leu	His	Lys	Leu	Leu	Arg	Lys	Ser
245								250							255
Thr	Phe	Lys	Asn	Ala	Lys	Leu	Tyr	Gly	Pro	Asp	Val	Gly	Gln	Pro	Arg
Arg	Lys	Thr	Ala	Lys	Met	Leu	Lys	Ser	Phe	Leu	Lys	Ala	Gly	Gly	Glu
275								280							285

Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr  
 290 295 300  
 Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile  
 305 310 315 320  
 Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly  
 325 330 335  
 Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Ala  
 340 345 350  
 Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys  
 355 360 365  
 Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val  
 370 375 380  
 Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro  
 385 390 395 400  
 Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
 405 410 415  
 Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg  
 420 425 430  
 Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly  
 435 440 445  
 Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu  
 450 455 460  
 Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu  
 465 470 475 480  
 Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn  
 485 490 495  
 Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met  
 500 505 510  
 Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser  
 515 520 525  
 Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile  
 530 535 540 543

## (2) SEQ ID NO: 11 的信息:

### (i) 序列特征:

- (A) 长度: 1718
- (B) 类型: 核酸
- (C) 链型: 双链
- (D) 拓扑结构: 线状

### (xi) 序列描述: SEQ ID NO: 11

CT AGA GCT TTC GAC	14	
TCT CCG CTG CGC GGC AGC TGG CGG GGG GAG CAG CCA GGT GAG CCC AAG	62	
ATG CTG CTG CGC TCG AAG CCT GCG CTG CGG CCG CCG CTG ATG CTG CTG	110	
Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu		
5	10	15

CTC CTG GGG CCG CTG GGT CCC CTC TCC CCT GGC GCC CTG CCC CGA CCT 158  
 Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro  
     20                 25                 30  
  
 GCG CAA GCA CAG GAC GTC GTG GAC CTG GAC TTC TTC ACC CAG GAG CCG 206  
 Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro  
     35                 40                 45  
  
 CTG CAC CTG GTG AGC CCC TCG TTC CTG TCC GTC ACC ATT GAC GCC AAC 254  
 Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn  
     50                 55                 60  
  
 CTG GCC ACG GAC CCG CGG TTC CTC ATC CTC CTG GGT TCT CCA AAG CTT 302  
 Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu  
     65                 70                 75                 80  
  
 CGT ACC TTG GCC AGA GGC TTG TCT CCT GCG TAC CTG AGG TTT GGT GGC 350  
 Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly  
     85                 90                 95  
  
 ACC AAG ACA GAC TTC CTA ATT TTC GAT CCC AAG AAG GAA TCA ACC TTT 398  
 Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe  
     100                105                110  
  
 GAA GAG AGA AGT TAC TGG CAA TCT CAA GTC AAC CAG GAT ATT TGC AAA 446  
 Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys  
     115                120                125  
  
 TAT GGA TCC ATC CCT CCT GAT GTG GAG GAG AAG TTA CGG TTG GAA TGG 494  
 Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp  
     130                135                140  
  
 CCC TAC CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC 542  
 Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe  
     145                150                155                160  
  
 AAG AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT 590  
 Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe  
     165                170                175  
  
 GCA AAC TGC TCA GGA CTG GAC TTG ATC TTT GGC CTA AAT GCG TTA TTA 638  
 Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu  
     180                185                190  
  
 AGA ACA GCA GAT TTG CAG TGG AAC AGT TCT AAT GCT CAG TTG CTC CTG 686  
 Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu  
     195                200                205  
  
 GAC TAC TGC TCT TCC AAG GGG TAT AAC ATT TCT TGG GAA CTA GGC AAT 734  
 Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn  
     210                215                220  
  
 GAA CCT AAC AGT TTC CTT AAG AAG CCT GAT ATT TTC ATC AAT GGG TCG 782  
 Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser  
     225                230                235                240  
  
 CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA CTT CTA AGA AAG TCC 830  
 Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser  
     245                250                255  
  
 ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT GAT ATT GTT GGT CAG CCT CGA 878  
 Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg  
     260                265                270  
  
 AGA AAG ACG GCT AAG ATG CTG AAG AGC TTC CTG AAG GCT GGT GGA GAA 926  
 Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu  
     275                280                285  
  
 GTG ATT GAT TCA GTT ACA TGG CAT CAC TAC TAT TTG AAT GGA CGG ACT 974  
 Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr  
     290                295                300  
  
 GCT ACC AGG GAA GAT TTT CTA AAC CCT GAT GTA TTG GAC ATT TTT ATT 1019  
 Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile  
     305                310                315                320  
  
 TCA TCT GTG CAA AAA GTT TTC CAG GTG GTT GAG AGC ACC AGG CCT GGC 1067  
 Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly  
     325                330                335

AAG AAG GTC TGG TTA GGA GAA ACA AGC TCT GCA TAT GGA GGC GGA GCG 1115  
 Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Ala  
     340                       345                       350  
  
 CCC TTG CTA TCC GAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA 1163  
 Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys  
     355                       360                       365  
  
 TTG GGC CTG TCA GCC CGA ATG GGA ATA GAA GTG GTG ATG AGG CAA GTA 1211  
 Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val  
     370                       375                       380  
  
 TTC TTT GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT 1259  
 Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro  
     385                       390                       395                   400  
  
 TTA CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC 1307  
 Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
     405                       410                       415  
  
 AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT CGA 1355  
 Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg  
     420                       425                       430  
  
 GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA GAA GGA 1403  
 Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly  
     435                       440                       445  
  
 GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC AAG TAC TTG 1451  
 Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu  
     450                       455                       460  
  
 CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT AAA TAC CTT CTA 1499  
 Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu  
     465                       470                       475                   480  
  
 AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA TCT GTC CAA CTC AAT 1547  
 Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn  
     485                       490                       495  
  
 GGT CTA ACT CTA AAG ATG GTG GAT GAT CAA ACC TTG CCA CCT TTA ATG 1595  
 Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met  
     500                       505                       510  
  
 GAA AAA CCT CTC CGG CCA GGA AGT TCA CTG GGC TTG CCA GCT TTC TCA 1643  
 Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser  
     515                       520                       525  
  
 TAT AGT TTT TTT GTG ATA AGA AAT GCC AAA GTT GCT GCT TGC ATC TGA 1691  
 Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile  
     530                       535                       540                   543  
  
 AAA TAA AAT ATA CTA GTC CTG ACA CTG                                     1718

## (2) SEQ ID NO: 12 的信息:

### (i) 序列特征:

- (A) 长度: 824
- (B) 类型: 核酸
- (C) 链型: 双链

(D) 拓朴结构： 线状

(xi) 序列描述： SEQ ID NO: 12

CTGGCAAGAA GGCTGGTT GGAGAGACGA GCTCAGCTTA CGGTGGCGGT GCACCCCTGC 60  
TGTCCAACAC CTTGCAGCT GGCTTTATGT GGCTGGATAA ATGGGCCCTG TCAGCCCAGA 120  
TGGGCATAGA AGTCGTGATG AGGCAAGGTGT TCTTCGGACG AGGCAACTAC CACTTAGTGG 180  
ATGAAAACCTT TGAGCCTTA CCTGATTACT GGCTCTCTCT TCTGTTCAAG AACTGGTAG 240  
GTCCCGGGGT GTTACTGTCA AGAGTGAAAG GCCCAGACAG GAGCAAACCTC CGAGTGATC 300  
TCCACTGCAC TAAGCTCAT CACCCACGAT ATCAGGAAGG AGATCTAACT CTGTATGTCC 360  
TGAACCTCCA TAATGTACC AGGCACTTGA AGGTACCGCC TCCGTTGTC AGGAAACAG 420  
TGGATACGTA CCTCTGAAG CCTTCGGGGC CGGATGGATT ACTTTCAAA TCTGTCAC 480  
TGAACCGTCA AATTCTGAAG ATGGTGGATG ACCAGACCCCT GCCAGCTTG ACAGAAAAC 540  
CTCTCCCCGC AGGAAGTGCA CTAAAGCCTGC CTGCTTTTC CTATGGTTT TTTGTCAAA 600  
GAATGCCAA AATCGCTGCT TGATATGAA AATAAAAGGC ATACGGTACC CCTGAGACAA 660  
AAGCCGAGGG GGTTGTTATT CATAAAACAA ACCCTAGTT TAGGAGGCCA CCTCCTGCC 720  
GAGTCCAGA GCTTCGGGAG GGTGGGTAC ACTTCAGTAT TACATTCACT GTGGTGTCT 780  
CTCTAAAGAAG AATACTGCAG GTGGTACAG TTAAAGCAC TGTG 824

(2) SEQ ID NO: 13 的信息：

(i) 序列特征：

- (A) 长度： 1899  
(B) 类型： 核酸  
(C) 链型： 双链  
(D) 拓朴结构： 线状

(xi) 序列描述： SEQ ID NO: 13

GGGAAAGCGA GCAAGGAAGT AGGAGAGAGC CGGGCAGGGCG GGGCGGGGTT GGATGGGAG 60  
CAGTGGGAGG GATGAGAAG AGGAGTGGGA GGATGGAGG GCGCAGTGGG AGGGGTTGAGG 120  
AGGCGTAACG GGGCGGAGGA AAGGAGAAAA GGGCGCTGGG GCTCGGCGGG AGGAAGTGCT 180  
AGAGCTCTCG ACTCTCGCT GCGGGGAGC TGGCGGGGG AGCACCCAGG TGAGCCCAA 240  
ATGCTGCTGC GCTCGAAGCC TGGCTGCCG CGCGCGCTGA TGCTGCTGCT CCTGGGGCCG 300  
CTGGGTCCTCC TCTCCCTGG CGCCCTGCCG CGACCTGCCG AAGCACAGGA CGTCGTGGAC 360  
CTGGACTCTCT CAACCCAGGA CGCGCTGCAC CTGGTGGGCC CCTCGTTCT CTCCCTCAC 420  
ATTGACGCCA ACCTGGCCAC TCCCTCATCC TCCCTGGCT TCCAAAGCTT 480  
CGTACCTTGG CCAGAGGCTT GTCTCTCGG TACCTGAGGT TTGGTGGCAC CAAGCAGAC 540  
TTCTTAATT TCGATCCCAA GAAGGAATCA ACCCTTGAG AGAGAAGTTA CTGGCAATCT 600  
CAAGTCACCC AGGATATTG CAAATATGGA TCCATCCCTC CTGATGTGGA GGAGAAGTTA 660  
CGGTTGGAAAT GCCCTACCA GGAGCAATTG CTACTCCGAG AACACTACCA GAAAAGCTC 720  
AAGAACAGCA CCTACTCAAG AGCTCTGTA GATGTGCTAT ACACCTTTGC AAACCTGCTCA 780  
GGACTGGACT TGATCTTGG CCTAAATGGC TTATTAAGAA CAGCAGATTG CGAGTGGAAC 840  
AGTCTTAATG CTCACTGGCT CTGGACTAC TGCTCTCCCA AGGGGTATAA CATTCTTGG 900  
GAACTAGGCCA ATGAACCTAA CAGTTTCCTT AAGAAGGCTG ATATTTCT CAATGGGTG 960  
CAGTTAGGAG AAGATIATAT TCAATTGCTTAA AACCTCTAA GAAAGTCAC CTTCAAAAT 1020  
GCCAAACTCT ATGGCTCTGA TTGTTGTCAG CCTCGAAGAA AGACGGCTAA GATGCTGAAG 1080  
AGCTTCTGA AGGCTGGTGG AGAAGTGATT GATTCACTTA CATGGCATCA CTACTATTG 1140  
AATGGACGGA CTGCTACCAAG GAAAGATTCTT CTAAACCTG ATGATTGGA CATTTTTATT 1200  
TCATCTGTGCA AAAAAGTTT CCAGGGTTT GAGAGCACCAG GGCCTGGCAA GAAGGTCTGG 1260  
TTAGGAGAAA CAAGCTCTGC ATATGGAGC GGAGCGCCCT TGCTATCCGA CACCTTTGCA 1320  
GCTGGCTTTA TTGTTGTCGA TAAATGGGC CTGTCAGCCC GAATGGGAAT AGAAGTGGTG 1380  
ATGAGGCCAG TATTCTTGG AGCAGGAAAC TACCATTTAG TTGGATGAAA CTTCGATECT 1440  
TTACCTGATT ATTGGCTATC TCTCTGTC AAGAAATTGG TGGGCCACCA CGTGTAAATG 1500  
GCAAGCGTGC AAGGTTCAAAG GAGAAGGAAG CTTCGAGTAT ACCTTCATTG CAAACACCT 1560  
GACAATCCAA GGTATAAAGA AGGAGATTTA ACTCTGTATG CCATAAACCT CCATAACGTG 1620  
ACCAAGTACT TGCGGTTACCT ATCTCTTTT TCTAAACAAAGC AAGTGGATAA ATACCTTCTA 1680  
AGACCTTTGG GACCTCATGG ATACTCTCC AAATCTGTCC AACCTCAATGG TCTAACTCTA 1740  
AAGATGGTGG ATGATCAACCTT TGCCACCTT TAAATGGAAA AACCTCTCGG GCCAGGAAGT 1800  
TCACTGGCT TGCAGCTT CTCATATAGT TTTTTGTTGTAAGAAAATGC CAAAGTTGCT 1860  
GCTTGATCT GAAAATAAAA TATACTAGTC CTGACACTG 1899

(2) SEQ ID NO: 14 的信息：

(i) 序列特征:

- (A) 长度: 592
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 14

Met Glu Gly Ala Val Gly Gly Val Arg Arg Arg Asn Gly Ala Glu  
5 10 15  
Glu Arg Arg Lys Gly Arg Trp Gly Ser Ala Gly Gly Ser Ala Arg  
20 25 30  
Ala Leu Asp Ser Pro Leu Arg Gly Ser Trp Arg Gly Glu Gln Pro  
35 40 45  
Gly Glu Pro Lys Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro  
50 55 60  
Pro Leu Met Leu Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro  
65 70 75  
Gly Ala Leu Pro Arg Pro Ala Gln Ala Gln Asp Val Val Asp Leu  
80 85 90  
Asp Phe Phe Thr Gln Glu Pro Leu His Leu Val Ser Pro Ser Phe  
95 100 105  
Leu Ser Val Thr Ile Asp Ala Asn Leu Ala Thr Asp Pro Arg Phe  
110 115 120  
Leu Ile Leu Leu Gly Ser Pro Lys Leu Arg Thr Leu Ala Arg Gly  
125 130 135  
Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys Thr Asp Phe  
140 145 150  
Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe Glu Glu Arg Ser  
155 160 165  
Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys Tyr Gly Ser  
170 175 180  
Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp Pro Tyr  
185 190 195  
Gln Glu Gln Leu Leu Arg Glu His Tyr Gln Lys Lys Phe Lys  
200 205 210  
Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe  
215 220 225  
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu  
230 235 240  
Leu Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu  
245 250 255  
Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu  
260 265 270  
Leu Gly Asn Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe  
275 280 285  
Ile Asn Gly Ser Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys  
290 295 300  
Leu Leu Arg Lys Ser Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro  
305 310 315  
Asp Val Gly Gln Pro Arg Arg Lys Thr Ala Lys Met Leu Lys Ser  
320 325 330  
Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Val Thr Trp His  
335 340 345  
His Tyr Tyr Leu Asn Gly Arg Thr Ala Thr Arg Glu Asp Phe Leu  
350 355 360  
Asn Pro Asp Val Leu Asp Ile Phe Ile Ser Ser Val Gln Lys Val  
365 370 375  
Phe Gln Val Val Glu Ser Thr Arg Pro Gly Lys Lys Val Trp Leu  
380 385 390  
Gly Glu Thr Ser Ser Ala Tyr Gly Gly Ala Pro Leu Leu Ser  
395 400 405  
Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys Leu Gly Leu  
410 415 420  
Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val Phe Phe  
425 430 435  
Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro Leu  
440 445 450  
Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
455 460 465  
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu  
470 475 480

Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys  
 485 490 495  
 Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr  
 500 505 510  
 Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp  
 515 520 525  
 Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys  
 530 535 540  
 Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Gln  
 545 550 555  
 Thr Leu Pro Pro Leu Met Glu Lys Pro Leu Arg Pro Gly Ser Ser  
 560 565 570  
 Leu Gly Leu Pro Ala Phe Ser Tyr Ser Phe Phe Val Ile Arg Asn  
 575 580 585  
 Ala Lys Val Ala Ala Cys Ile  
 590 592

## (2) SEQ ID NO: 15 的信息:

### (i) 序列特征:

- (A) 长度: 1899
- (B) 类型: 核酸
- (C) 链型: 双链
- (D) 拓扑结构: 线状

### (xi) 序列描述: SEQ ID NO: 15

AAA GCG AGC AAG GAA GTA GGA GAG AGC CGG GCA GCA GGG GCG GCG GGG	3
TTG GAT TGG GAG CAG TGG GAG GGA TGC AGA AGA GGA GTG GGA GGG	48
ATG GAG GGC GCA GTG GGA GGG GTG AGG AGG CGT AAC GGG GCG GAG	93
Met Glu Gly Ala Val Gly Gly Val Arg Arg Arg Asn Gly Ala Glu	138
5 10 15	
GAA AGG AGA AAA GGG CGC TGG GGC TCG CGG GGA GGA AGT GCT AGA	183
Glu Arg Arg Lys Gly Arg Trp Gly Ser Ala Gly Gly Ser Ala Arg	20 25 30
GCT CTC GAC TCT CCG CTG CGC GGC AGC TGG CGG GGG GAG CAG CCA	228
Ala Leu Asp Ser Pro Leu Arg Gly Ser Trp Arg Gly Glu Gln Pro	35 40 45
GGT GAG CCC AAG ATG CTG CTG CGC TCG AAG CCT GCG CTG CGG CGG	273
Gly Glu Pro Lys Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro	50 55 60
CCG CTG ATG CTG CTG CTC CTG GGG CGG CTG GGT CCC CTc TCC CCT	318
Pro Leu Met Leu Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro	65 70 75
GGC GCC CTG CCC CGA CCT GCG CAA GCA CAG GAC GTc GTG GAC CTG	363
Gly Ala Leu Pro Arg Pro Ala Gln Ala Gln Asp Val Val Asp Leu	80 85 90
GAC TTc TTC ACC CAG GAG CCG CTG CAC CTG GTG AGC CCC TCG TTc	408
Asp Phe Phe Thr Gln Glu Pro Leu His Leu Val Ser Pro Ser Phe	95 100 105
CTG TCC GTC ACC ATT GAC GCC AAC CTG GCC ACG GAC CGG CGG TTc	453

Leu Ser Val Thr Ile Asp Ala Asn Leu Ala Thr Asp Pro Arg Phe		
110	115	120
CTC ATC CTC CTG GGT TCT CCA AAG CTT CGT ACC TTG GCC AGA GGC	498	
Leu Ile Leu Leu Gly Ser Pro Lys Leu Arg Thr Leu Ala Arg Gly		
125	130	135
TTG TCT CCT GCG TAC CTG AGG TTT GGT GCC ACC AAG ACA GAC TTC	543	
Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys Thr Asp Phe		
140	145	150
CTA ATT TTC GAT CCC AAG AAG GAA TCA ACC TTT GAA GAG AGA AGT	588	
Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe Glu Glu Arg Ser		
155	160	165
TAC TGG CAA TCT CAA GTC AAC CAG GAT ATT TGC AAA TAT GGA TCC	633	
Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys Tyr Gly Ser		
170	175	180
ATC CCT CCT GAT GTG GAG AAG TTA CGG TTG GAA TGG CCC TAC	678	
Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp Pro Tyr		
185	190	195
CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC AAG	723	
Gln Glu Gln Leu Leu Arg Glu His Tyr Gln Lys Lys Phe Lys		
200	205	210
AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT	768	
Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe		
215	220	225
GCA AAC TGC TCA GGA CTG GAC TTG ATC TTT GGC CTA AAT GCG TTA	813	
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu		
230	235	240
TTA AGA ACA GCA GAT TTG CAG TGG AAC AGT TCT AAT GCT CAG TTG	858	
Leu Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu		
245	250	255
CTC CTG GAC TAC TGC TCT TCC AAG GGG TAT AAC ATT TCT TGG GAA	903	
Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu		
260	265	270
CTA GCC AAT GAA CCT AAC AGT TTC CTT AAG AAG GCT GAT ATT TTC	948	
Leu Gly Asn Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe		
275	280	285
ATC AAT GGG TCG CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA	993	
Ile Asn Gly Ser Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys		
290	295	300
CTT CTA AGA AAG TCC ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT	1038	
Leu Leu Arg Lys Ser Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro		
305	310	315
GAT GTT GGT CAG CCT CGA AGA AAG ACG GCT AAG ATG CTG AAG AGC	1083	
Asp Val Gly Gln Pro Arg Arg Lys Thr Ala Lys Met Leu Lys Ser		
320	325	330
TTC CTG AAG GCT GGT GGA GAA GTG ATT GAT TCA GTT ACA TGG CAT	1128	
Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Val Thr Trp His		
335	340	345
CAC TAC TAT TTG AAT GGA CGG ACT GCT ACC AGG GAA GAT TTT CTA	1173	
His Tyr Tyr Leu Asn Gly Arg Thr Ala Thr Arg Glu Asp Phe Leu		
350	355	360
AAC CCT GAT GTA TTG GAC ATT TTT ATT TCA TCT GTG CAA AAA GTT	1218	
Asn Pro Asp Val Leu Asp Ile Phe Ile Ser Ser Val Gln Lys Val		
365	370	375
TTC CAG GTG GTT GAG AGC ACC AGG CCT GGC AAG AAG GTC TGG TTA	1263	
Phe Gln Val Val Glu Ser Thr Arg Pro Gly Lys Lys Val Trp Leu		
380	385	390

GGA GAA ACA AGC TCT GCA TAT GGA GGC GGA GCG CCC TTG CTA TCC Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro Leu Leu Ser 395 400 405	1308
GAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA TTG GGC CTG Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys Leu Gly Leu 410 415 420	1353
TCA GCC CGA ATG GGA ATA gAA GTG GTG ATG AGG CAA GTA TTC TTT Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val Phe Phe 425 430 435	1398
GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT TTA Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro Leu 440 445 450	1443
CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr 455 460 465	1488
AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu 470 475 480	1533
CGA GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA Arg Val Tyr Leu His Cys Thr Asn Thr Asn Pro Arg Tyr Lys 485 490 495	1578
GAA GGA GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr 500 505 510	1623
AAG TAC TTG CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp 515 520 525	1668
AAA TAC CTT CTA AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys 530 535 540	1713
TCT GTC CAA CTC AAT GGT CTA ACT CTA AAG ATG GTG GAT GAT CAA Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Gln 545 550 555	1758
ACC TTG CCA CCT TTA ATG GAA AAA CCT CTC CGG CCA GGA AGT TCA Thr Leu Pro Pro Leu Met Glu Lys Pro Leu Arg Pro Gly Ser Ser 560 565 570	1803
CTG GGC TTG CCA GCT TTC TCA TAT AGT TTT TTT GTG ATA AGA AAT Leu Gly Leu Pro Ala Phe Ser Tyr Ser Phe Phe Val Ile Arg Asn 575 580 585	1848
GCC AAA GTT GCT GCT TGC ATC TGA AAA TAA AAT ATA CTA GTC CTG Ala Lys Val Ala Ala Cys Ile 590 592	1893
ACA CTG	1899

## (2) SEQ ID NO: 16 的信息:

- 序列特征:
  - 长度: 594
  - 类型: 核酸
  - 链型: 双链

(D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 16

```
ATTACTATAG GGCACCGGTG GTGACGGCC CGGGCTGGTA TTGTCTTAAT GAGAAGTTGA 60
TAAAGAATT TGGGTGGTTC ATCTCTTCC AGCTGCAGTT TAGCGTATGC TGAGGCCAGA 120
TTTTTCAGG CAAAGTAAA ATACCTGAGA AACTGCCTGG CCAGAGGACA ATCAGATTTT 180
GGCTGGCTCA ACTGACAAGC AAGTGTATG AAGCTAGATG GGAGAGGAAG GGATGAATAC 240
TCCATTGGAG GCTTTACTCG AGGGTCAGAG GGATACCCGG CGCCATCAGA ATGGGATCTG 300
GGAGTCGGAA ACCTGGGTT CCCACGAGAG CGCGCAGAAC ACCTGCCTGA GGAAGCCTGG 360
TCCGGGATGC CCAGCGCTGC TCCCCGGCC CTCTCCCCG GGCCTCTC CCCAGGCCCTC 420
CCGGCCGCTT GGATCCCGC CATCTCCGCA CCCTTCAAGT GGGGTGGGT GATTTCTGAA 480
GTAAACGTGA CCCCCACCGG GGGGAAAGCG AGCAAGGAAG TAGGAGAGAG CGGGGCAGGC 540
GGGGCGGGGT TGGATTGGGA GCAGTGGGAG GGATGCAGAA GAGGAGTGGG AGGG 594
```

(2) SEQ ID NO: 17 的信息:

(i) 序列特征:

- (A) 长度: 21
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 17

CCCCAGGAGC AGCAGCATCA G 21

(2) SEQ ID NO: 18 的信息:

(i) 序列特征:

- (A) 长度: 21
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 18

AGGCTTCGAG CGCAGCAGCA T 21

(2) SEQ ID NO: 19 的信息:

(i) 序列特征:

- (A) 长度: 22
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 19  
GTAATACGAC TCACTATAGG GC 22

(2) SEQ ID NO: 20 的信息:

- (i) 序列特征:
- (A) 长度: 19
  - (B) 类型: 核酸
  - (C) 链型: 单链
  - (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 20  
ACTATAGGGC ACGCGTGGT 19

(2) SEQ ID NO: 21 的信息:

- (i) 序列特征:
- (A) 长度: 21
  - (B) 类型: 核酸
  - (C) 链型: 单链
  - (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 21  
CTTGGGCTCA CCTGGCTGCT C 21

(2) SEQ ID NO: 22 的信息:

- (i) 序列特征:
- (A) 长度: 23
  - (B) 类型: 核酸
  - (C) 链型: 单链
  - (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 22  
AGCTCTGTAG ATGTGCTATA CAC 23

(2) SEQ ID NO: 23 的信息:

- (i) 序列特征:
- (A) 长度: 22

- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓朴结构: 线状

(xi) 序列描述: SEQ ID NO: 23

GCATCTTAGC CGTCTTCCTT CG 22

# 说 明 书 附 图

图 1

1 CTAGACCTTTCGACTCTCGGTGCGGGCACCTGGGGGGGGAGCAGCCAGGTGAGCCC  
61 AGATGCTGCTGCGCTCGAAGCCYGCCTGCCCGCCGCTGATGCTGCTGCTGCTCTGGGC  
M L D R S K P A L P P P L M L L L G P  
121 CGCTGGGTCCCCCTCTCCCTGGCGCCCTGCCCGACCTGGCGACAGCACAGGACGTCGTCGG  
L G P L S P G A L P R P A Q A Q D V V D  
181 ACCTGGACTCTTCACCCAGGAGCCGCTGCACCTGGTGAGGCCCTCGTCTGCGTC  
L D F F T Q E P L H L V S P S F L S V T  
241 CCATTGACGCCAACCTGGGCCACGGACCCGCGGTCTCTCATCCCTCTGGTTCTCAAAGC  
I D A N L A T D P R F L I L L G S P K L  
301 TTGCTACCTGGCCAGAGGCTTGTCTCTGCGTACCTGAGGGTTGGCTGCCACCAAGACAG  
R T L A R C L S P A Y L R F G G T K T D  
361 ACTTCCTAAATTCGATCCAAAAGGAATCAACCTTTGAAGAGAGAAGTTACTGGCAAT  
F L I F D P K K E S T F E Z R S Y N Q S  
421 CTCAGTCACCAACAGGATAITTCGAAATATGGATCCATCCCTCTGATCTGGAGGAGRAST  
Q V N Q D I C K Y G S I P P D V E E K L  
481 TACGGTTGGAATGGCCCTACCAAGGAGCAATTGCTACTCCGAGAACACTACCAAGAAAAAGT  
R L E W P Y Q E Q L L R E H Y Q K K F  
541 TCAAGAACAGCACCTACTCAAGAAGCTCTGTAGATGCTATACTTTGCAAACGCT  
K N S T Y S R S S V D V L Y T F A N C S  
601 CAGGACTGGACTGATCTTGGCTAAATGGCTTAAAGAACAGCAGATTGCACTGG  
G L D L I F G L N A L L R T A D L Q W N  
661 ACAGTTCTAAATGCTCAAGTCTCTGGACTCTGCTCTCCAAAGGGTATAACATTCTT  
S S N A Q L L L D Y C S S K G Y N I S W  
721 GGGAACTAGGCAATGAACTAAACAGTTCTTAAGAAGGCTGATATTTCAATGGT  
E L G N E P N S F L K K A D I F I N G S  
(\*)  
781 CGCAGTTAGGAAAGATTATTCATTCATTGCTAAACTTCTAAGAAAGCTCACCTCAAA  
Q L G E D Y I Q L H K L L R K S T F K N  
(F)  
841 ATGCAAACCTCTATGTCCTGATGTTGGCTAGCCTCGAAGAAAAGACGGCTAAGATGCTGA  
A K L Y G P D V G Q P R R K T A K M L K  
901 ACAGCTCTGAAAGGCTGGAGAGAASTCATGATTCAAGTACATGGCATCACTACTATT  
S F L K A G G E V I D S V T W H H Y Y L  
961 TGAATGGACGGACTGCTACCAGGGAGATTCTAAACCTGATSTATGGACATTITA  
N G R T A T R E D F L N P D V L D I F I  
1021 TTTCATCTGTGCAAAAGTTCTTCCAGGGGGTGGAGAGCACCAGGCCCTGGCAAGAAGGTCT  
S S V Q K V F Q V V E S T R P G K K V W  
1081 GTTCTGGAGAAACAAGCTCTCCATATGGAGGGGGAGCGCCCTGGCTATCCGACACCTTIG  
L G E T S S A Y G G G A P L L S D T F A  
1141 CAGCTGGCTTATGTGGCTGGATTAATGGCTGTCAGCCGAATGGGAATAGAAGTGG  
A G F M W L D K L G L S A R M G I E V V  
1201 TGATGAGGCAAGTATTCCTGGAGCAGGAACCTACCATTTGTTGATGAAAACCTCGATC  
M R Q V F F G A G N Y X L V D E N F D P  
1261 CTTTACCTGATTATGGCTATCTCTCTGTTCAAGAAATTGGGGGGACCAAGGTGTTAA  
L P D Y W L S L L F K K L V G T K V L M  
1321 TGSCAAGCGTGCAGGTTCTAGAGAGAAGGGAGCTCCAGTATACTCTCATGCAAAACA  
A S V Q G S K R R K L R V Y L H C T N T  
1381 CTGACAATCCAAGGTATAAGAAGGAGATTAACTCTGATGCCATAACCTCCATAACG  
D N P R Y K E G D L T L Y A I N L H N V  
1441 TCACCAAGTACTTGGCTTACCTATCCTTCTAACAAGCAGTGGATAATACCTTC  
T K Y L R L P Y P F S N K Q V D K Y L L  
1501 TAAGACCTTGGGACTCATGGATTACTTCCAAATCTGTCACACTCAATGGCTAACTC  
R P L G P H G L L S K S V Q L N G L T L  
1561 TAAAGATGGGATGATCAPACCTTGGCACCTTAAATGGAAAACCTCTCCGGCCAGGAA  
K M V D D Q T L P P L M E K P L R P G S  
1621 GTTCAGTGGGCTTGGCTAGCTTCTCATATGTTTGTGATAAGAAATGCCAAAGTG  
S L G L P A F S Y S F F V I R N A X V A  
1681 CTGCTTGCACTGAAATAAAATAACTACTCTGACACTG  
A C I

图 2

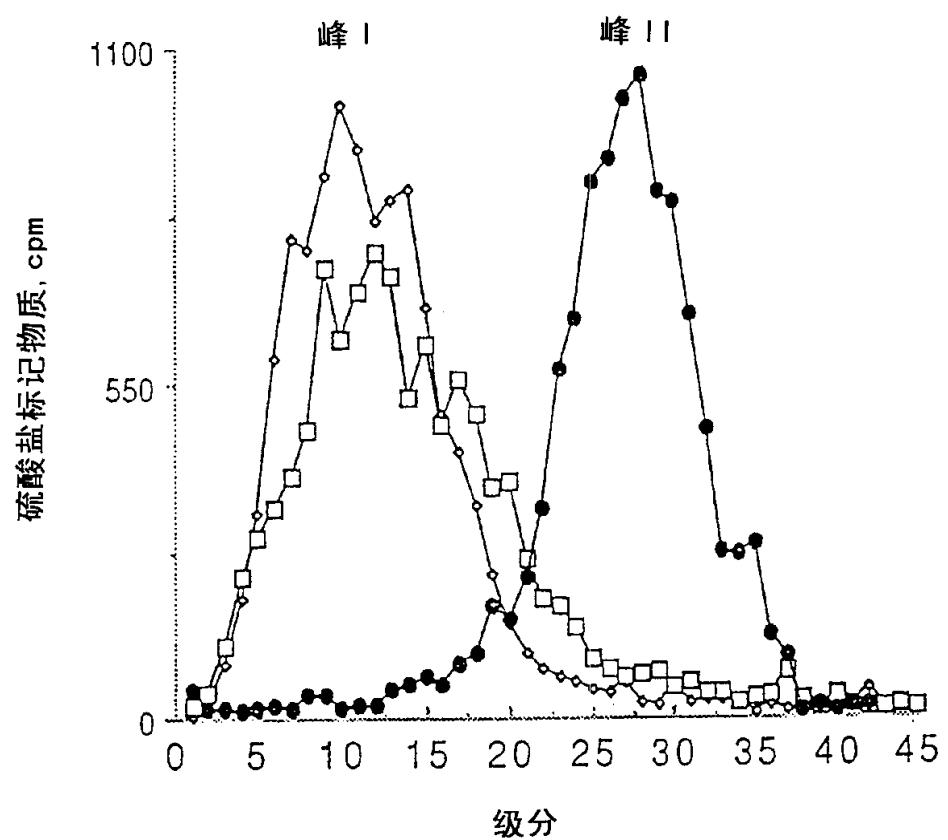


图 3A

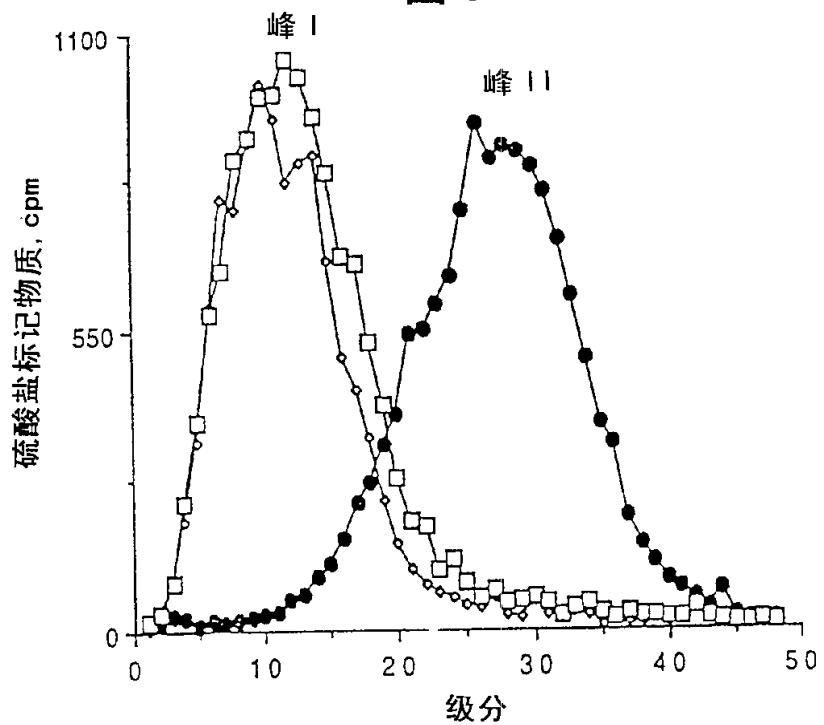


图 3B

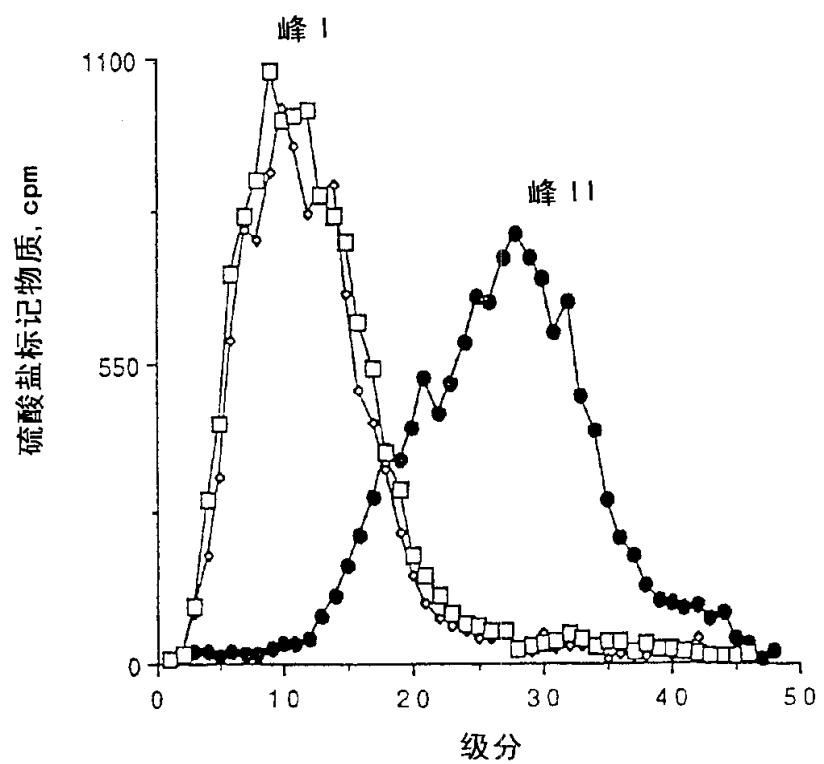


图 4

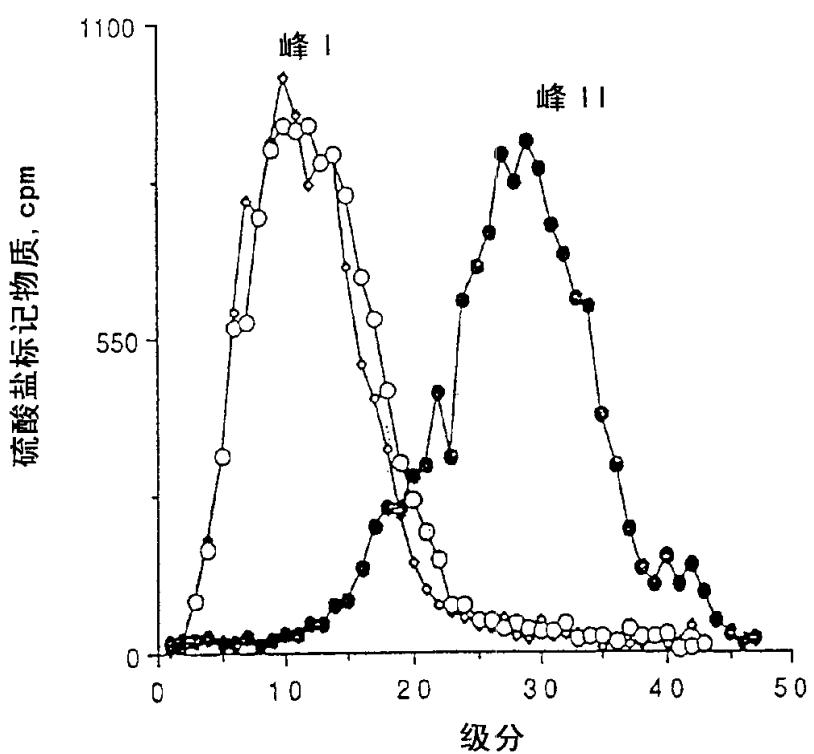


图 5A

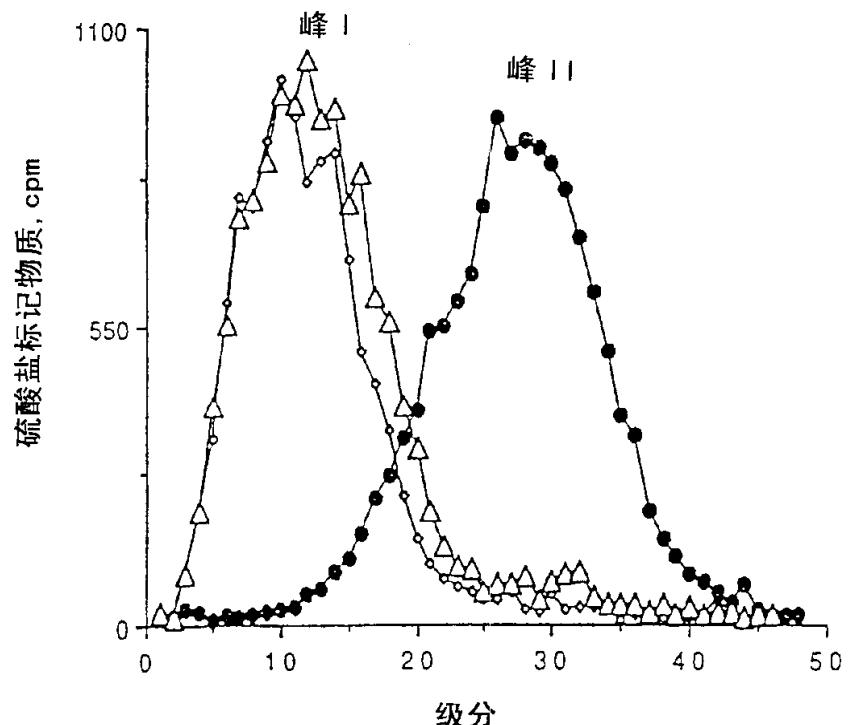


图 5B

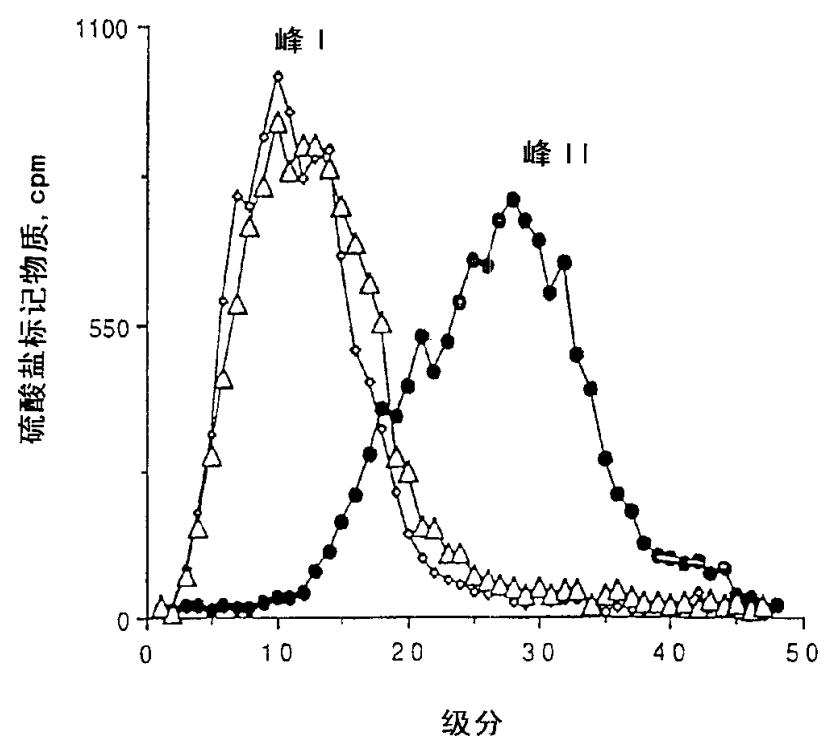


图 6A

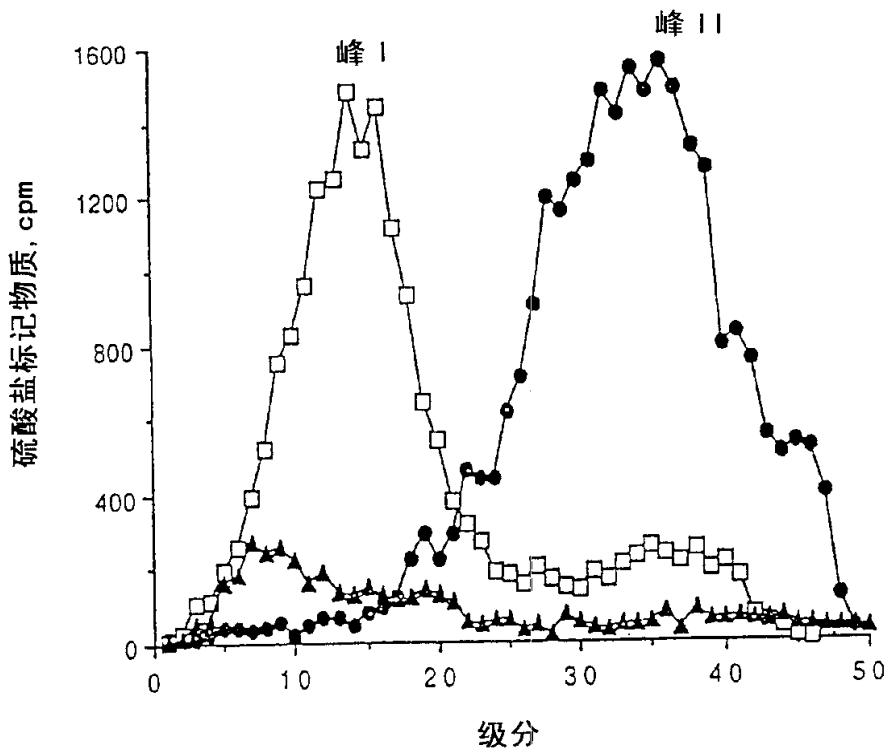


图 6B

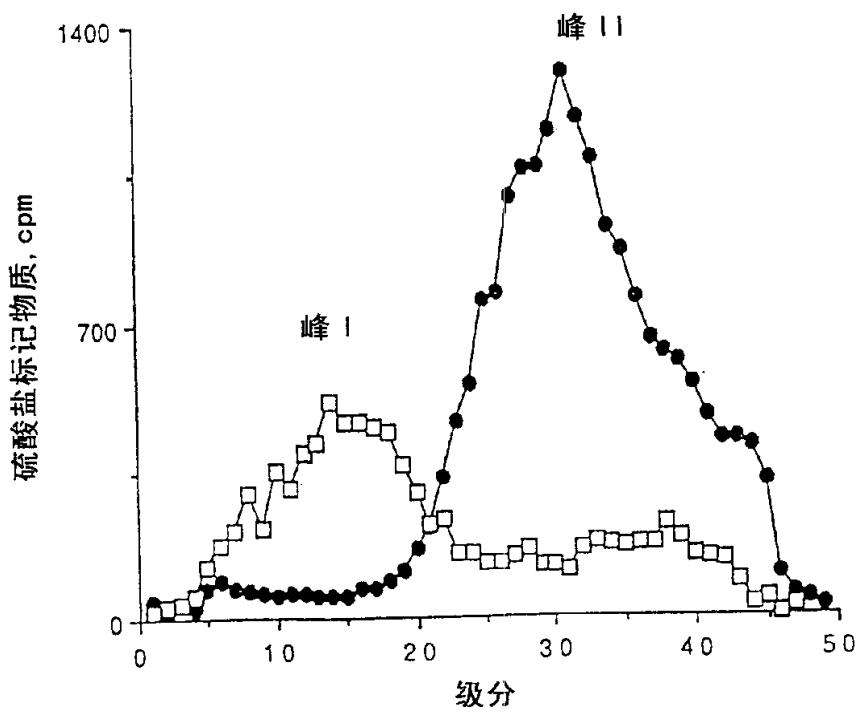


图 7A

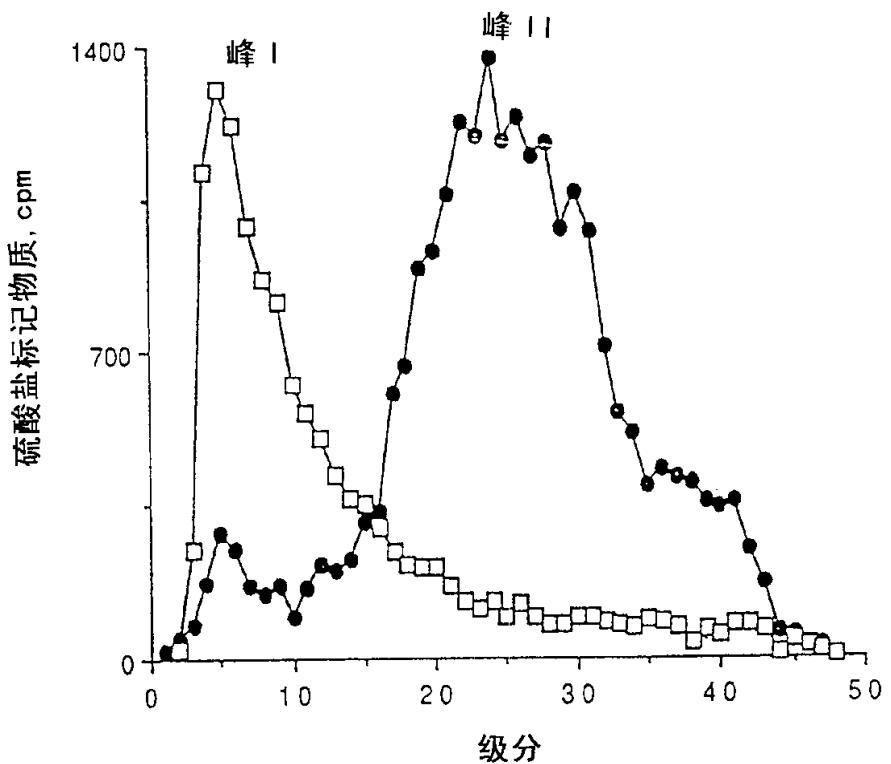


图 7B

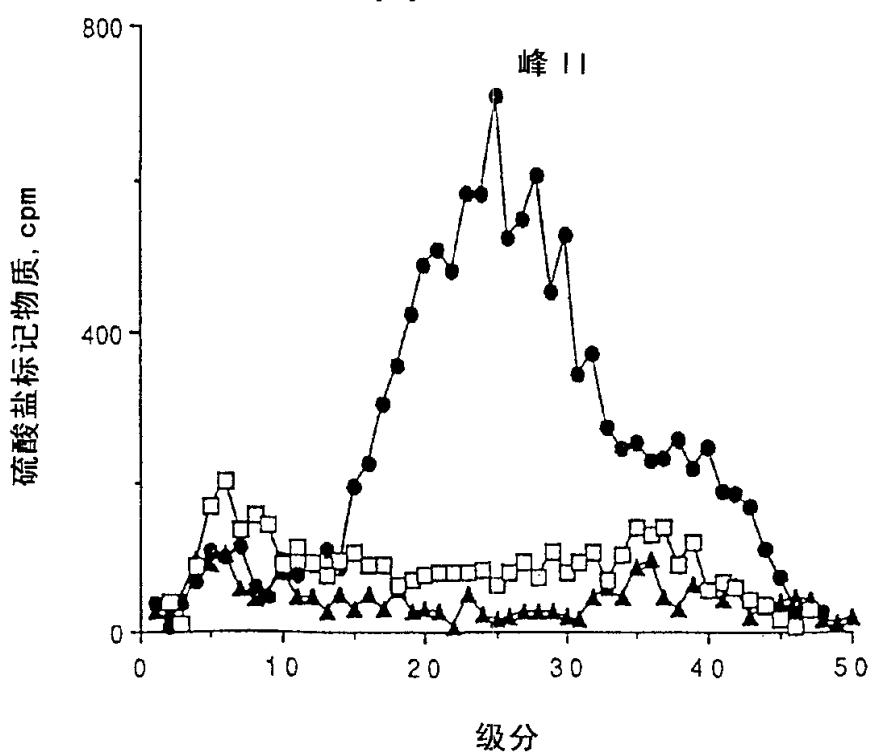


图 8A

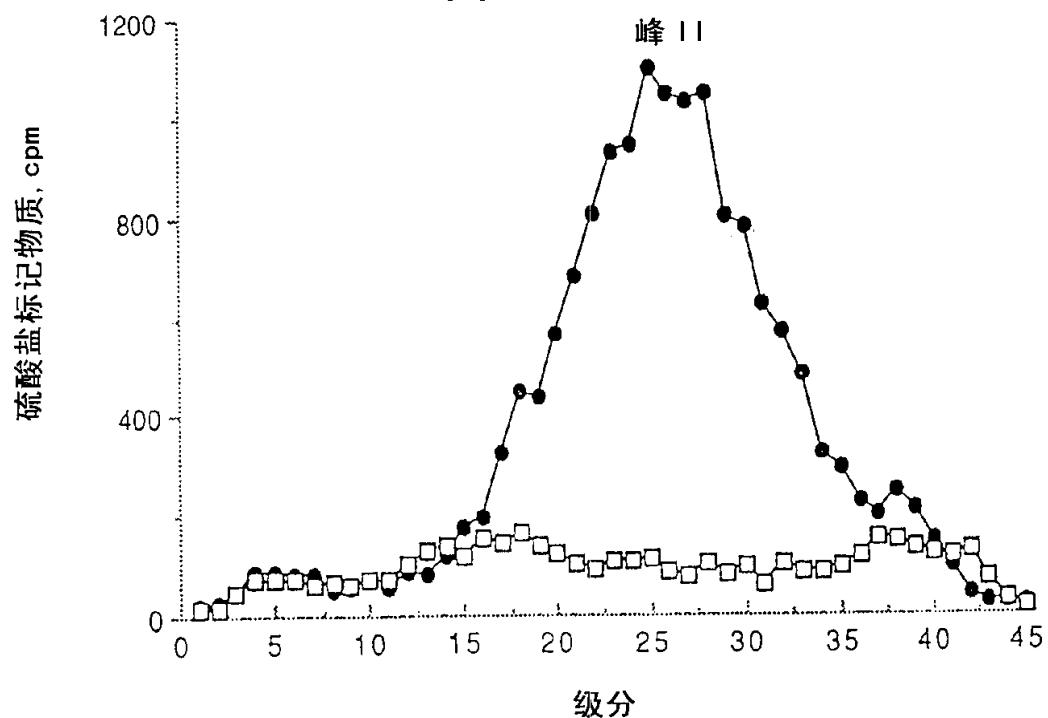


图 8B

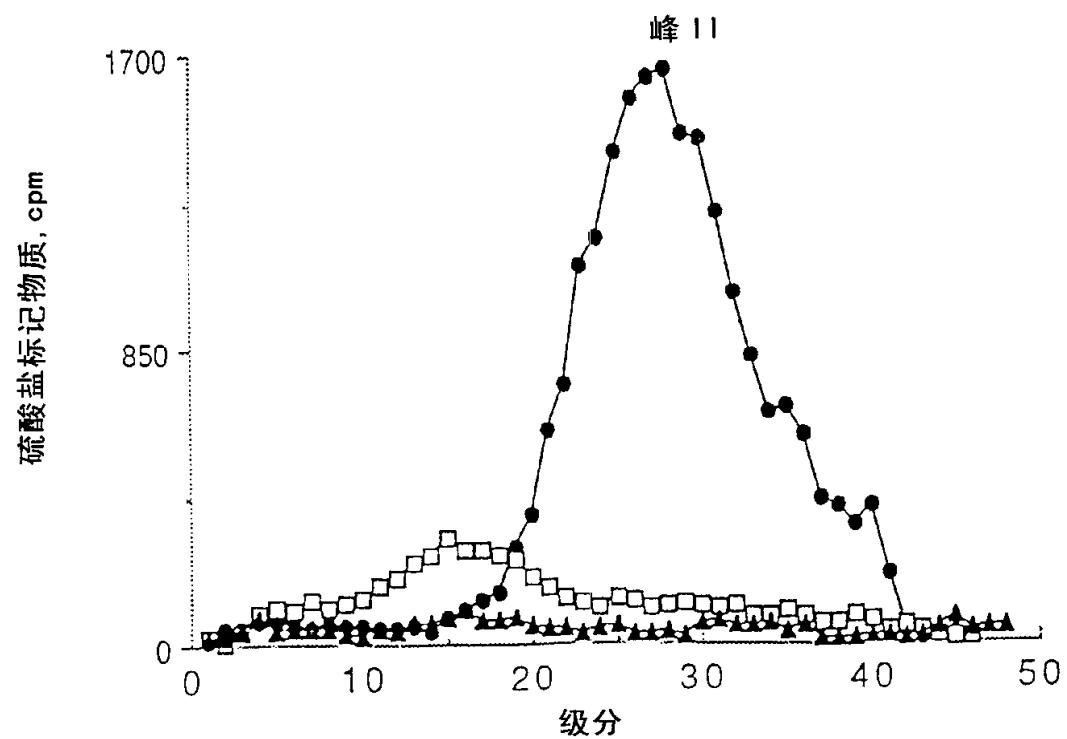


图 9A

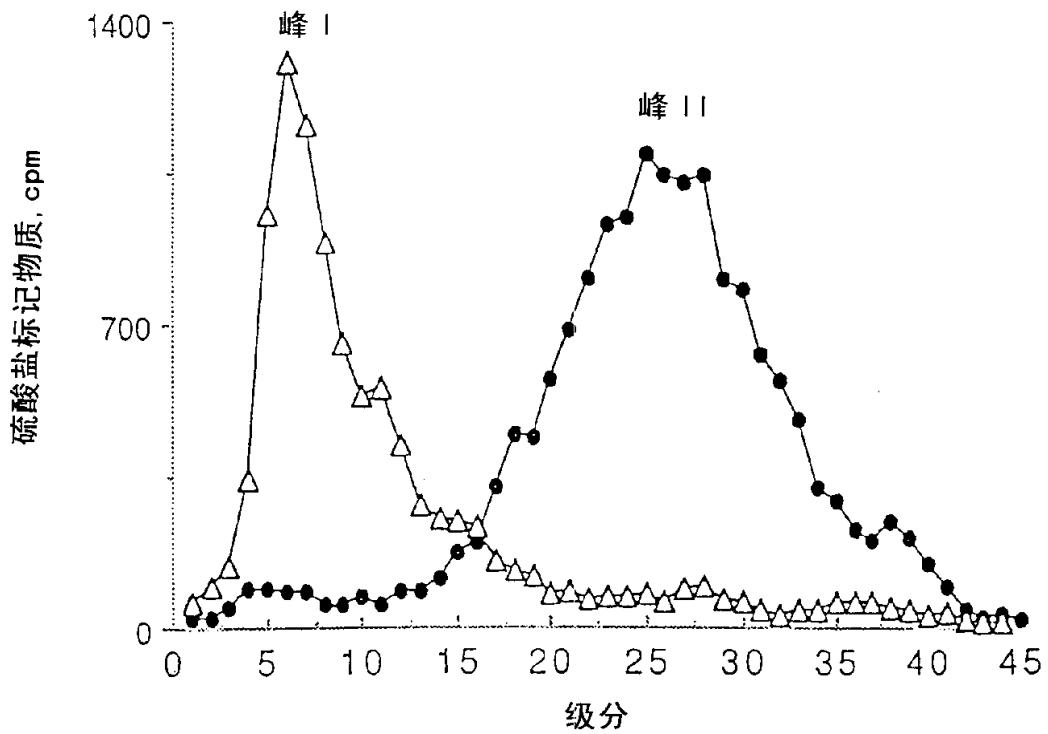


图 9B

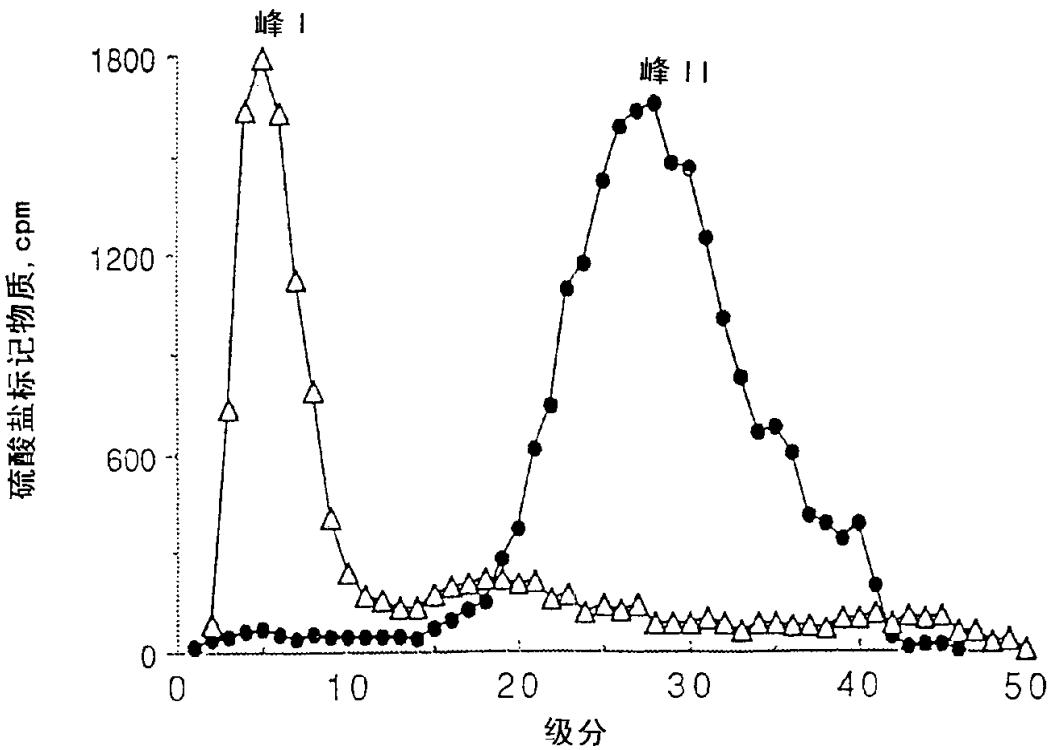


图 10A

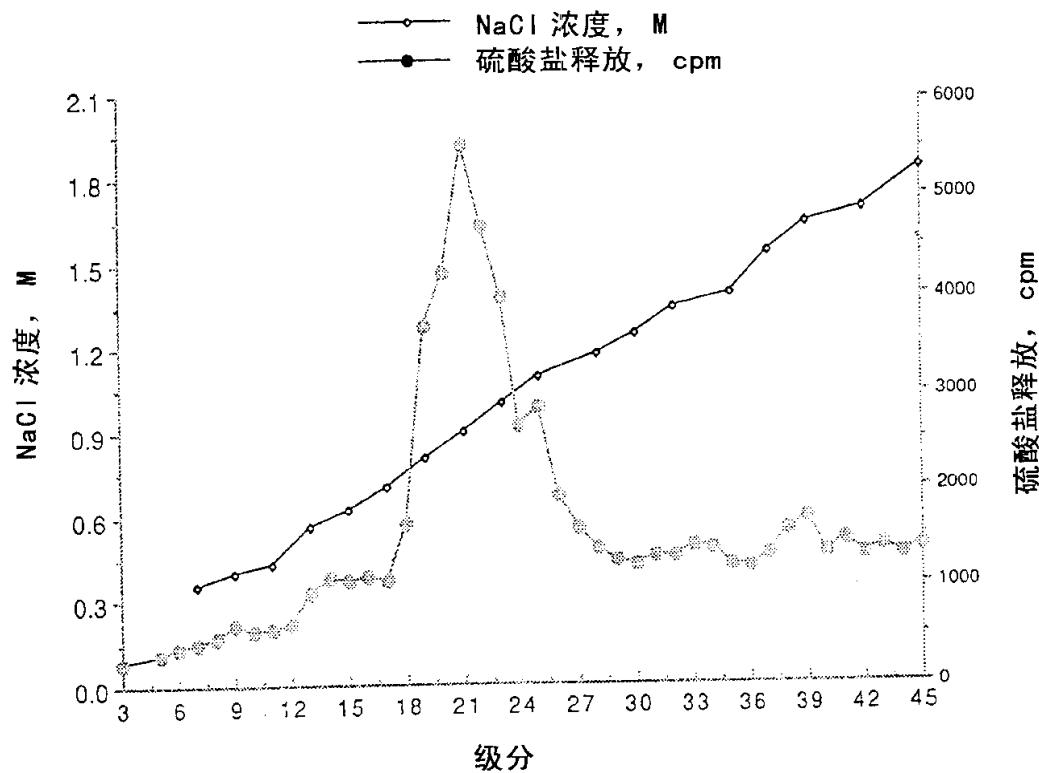


图 10B

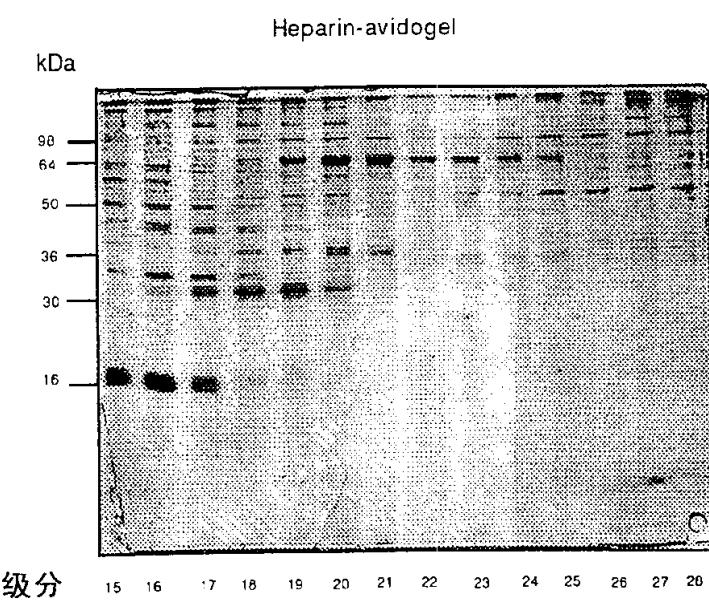


图 11A

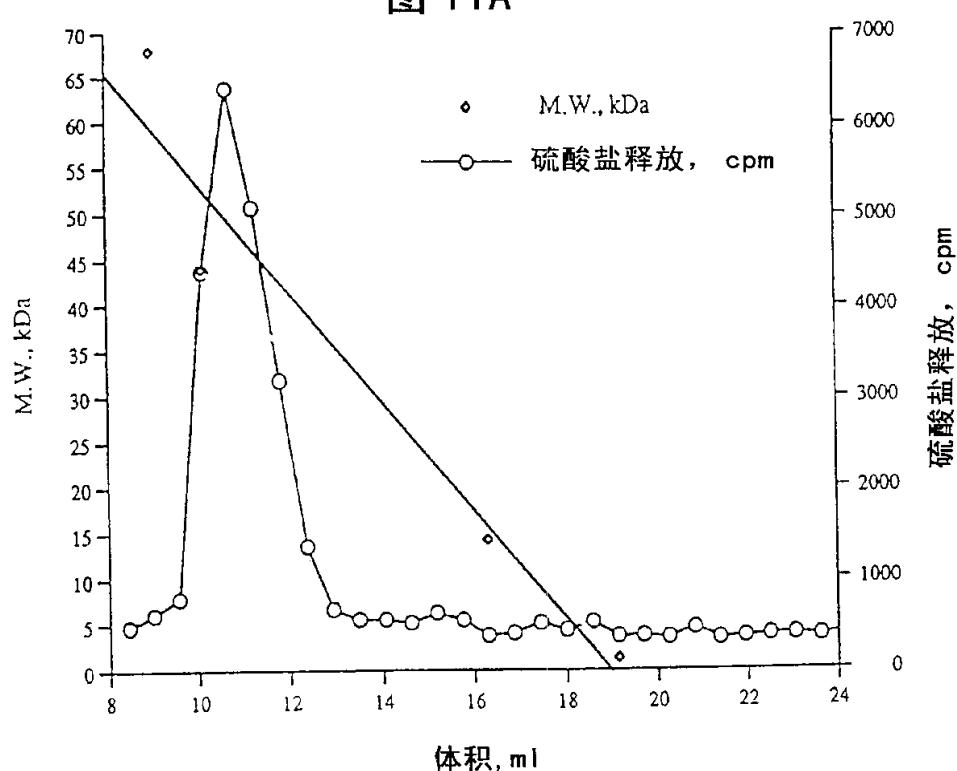


图 11B

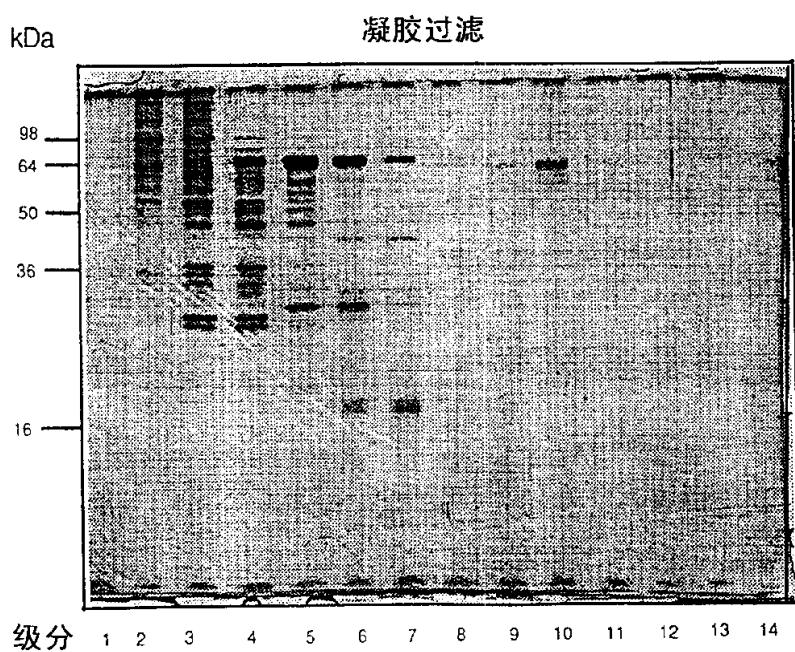


图 12A

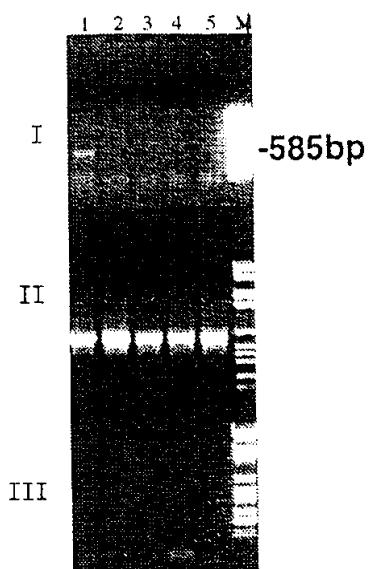


图 12B

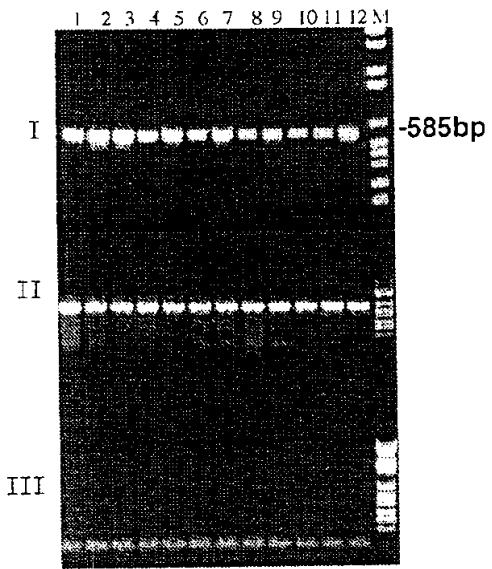


图 12C

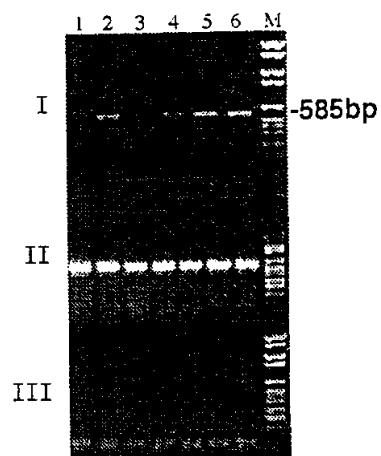


图 12D

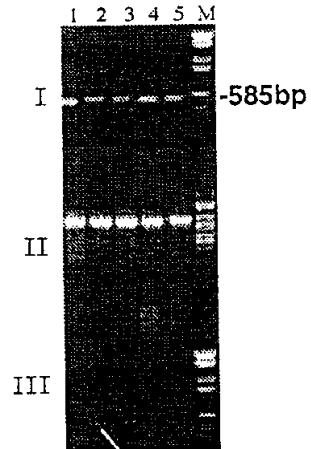
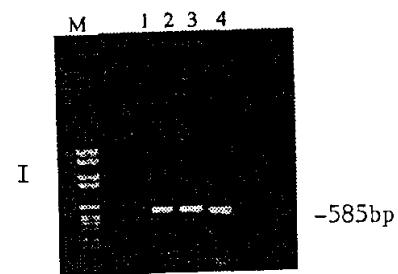


图 12E



### 图 13

小鼠	CTGGCAAGAAGGTCTGGTTGGGAGAGACCGAGCTCAGCTTACGGTGGCGGT	50
人		
小鼠	CTGGCAAGAAGGTCTGGTTAGGAGAAACAAGCTCTGCATATGGAGGCGGA	1115
人		
小鼠	GCACCCCTGCTGCCAACACCTTTGCAGCTGGCTTATGTGGCTGGATAA	100
人		
小鼠	GCGCCCTGCTATCCGACACCTTTGCAGCTGGCTTATGTGGCTGGATAA	1165
人		
小鼠	ATTGGGCCTGTCAGCCCAGATGGGCATAGAAGTCGTATGAGGCAGGTGT	150
人		
小鼠	ATTGGGCCTGTCAGCCCAGATGGGAATAGAAGTGGTATGAGGCAAGTAT	1215
人		
小鼠	TCTTCGGAGCAGGCAACTACCACTTAGTGGATGAAAACTTGAGCCTTA	200
人		
小鼠	TCTTTGGAGCAGGAAACTACCAATTAGTGGATGAAAACCTCGATCCTTA	1265
人		
小鼠	CCTGATTACTGGCTCTCTCTGTTCAAGAAACTGGTAGGTCCCAGGGT	250
人		
小鼠	CCTGATTATTGGCTATCTCTGTTCAAGAAATTGGTGGCACCAAGGT	1315
人		
小鼠	GTTACTGTCAAGAGTGAAAGGCCAGACAGGAGCAAACCTCGAGTGTATC	300
人		
小鼠	GTAAATGGCAAGCGTGCACAGGTTCAAAGAGAAGGAAGCTCGAGTATAACC	1365
人		
小鼠	TCCACTGCACTAACGTCTATCACCCACGATATCAGGAAGGGAGACTAACT	350
人		
小鼠	TTCATTGCACAAACACTGACAACTCAAGGTATAAAGAAGGGAGATTAACT	1415
人		
小鼠	CTGTATGTCCTGAACCTCCATAATGTCACCAAGCACTTGAAGGTACCGCC	400
人		
小鼠	CTGTATGCCATAAACCTCCATAACGTACCAAGTACTTGCAGGTTACCCCTA	1465
人		
小鼠	TCCGTTGTCAGGAAACCAAGTGGATACGTACCTTCTGAAGCCTTCGGGC	450
人		
小鼠	TCCTTTTCTAACAGCAAGTGGATAAAACCTTCTAACACCTTGGGAC	1515
人		
小鼠	CGGATGGATTACTTCCAAATCTGTCCAACGTGACAGGGTCAAATTCTGAAG	500
人		
小鼠	CTCATGGATTACTTCCAAATCTGTCCAACGTCAATGGTCTAACTCTAAAG	1565
人		
小鼠	ATGGTGGATGAGCAGACCCCTGCCAGCTTGACAGAAAAACCTCTCCCCGC	550
人		
小鼠	ATGGTGGATGATCAAACCTTGCACCTTAAATGGAAAAACCTCTCCGGCC	1615
人		
小鼠	AGGAAGTGCACTAAGCCTGCCCTTCCATGGTTTTGTGATCAA	600
人		
小鼠	AGGAAGTTCACTGGGCTTGCACGCTTCTCATATAGTTTTGTGATCAA	1665
人		
小鼠	GAAATGCCAAAATCGCTGCTTGTATATGAAAATAAAA	637
人		
小鼠	GAAATGCCAAAAGTTGCTGCTTGCATCTGAAAATAAAA	1702

图 14

