

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/11

C07K 14/65

A61K 48/00



[12] 发明专利说明书

专利号 ZL 98811833.5

[45] 授权公告日 2005 年 11 月 30 日

[11] 授权公告号 CN 1229496C

[22] 申请日 1998.10.4 [21] 申请号 98811833.5

[30] 优先权

[32] 1997.10.3 [33] US [31] 08/943,608

[86] 国际申请 PCT/IL1998/000486 1998.10.4

[87] 国际公布 WO1999/018195 英 1999.4.15

[85] 进入国家阶段日期 2000.6.5

[71] 专利权人 耶路撒冷希伯来大学伊森姆研究发
展公司

地址 以色列耶路撒冷市

[72] 发明人 A·霍赫贝格 S·阿耶什

审查员 安玉苹

[74] 专利代理机构 中国专利代理(香港)有限公司

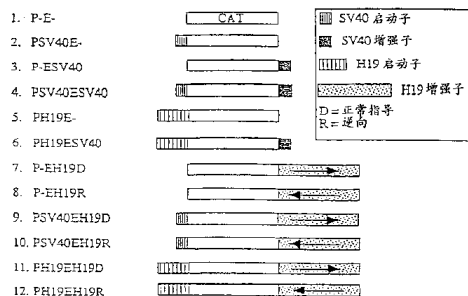
代理人 卢新华 谭明胜

权利要求书 2 页 说明书 31 页 附图 15 页

[54] 发明名称 诱导肿瘤特异性细胞毒性的方法和
组合物

[57] 摘要

本发明涉及在调节转录序列的控制下, 异源序列尤其是编码细胞毒性产物的基因在肿瘤细胞中的特异性表达。特别优选的启动子包括 H19 调节序列、IGF-1 启动子和 IGF-2 P3 和 P4 启动子。本发明提供表达构建体和施用上述表达构建体的方法。本发明的组合物和方法可用于癌症治疗。



A 概括性表示含有不同启动子和增强子的质粒

ISSN 1008-4274

1. 一种包含与调节序列操作性连接的编码细胞毒性基因产物的异源序列的多核苷酸在制备用于治疗患者的癌症的药物中的用途，其中该调节序列来自 H19 调节元件。
- 5 2. 权利要求 1 的用途，其中该 H19 调节序列是 H19 启动子、H19 增强子或 H19 启动子和 H19 增强子两者。
3. 权利要求 2 的用途，其中该 H19 启动子包含 SEQ ID NO: 1 的 1 - 830 位的核苷酸。
4. 权利要求 2 的用途，其中该 H19 启动子包含 SEQ ID NO: 2
10 的序列。
5. 权利要求 2 的用途，其中该 H19 增强子包含克隆在 ATCC 保藏号 209322 的质粒 pH19EH19 中的 H19 增强子的序列。
6. 权利要求 2 的用途，其中该 H19 增强子包含 SEQ ID NO: 3 的序列。
- 15 7. 权利要求 2 的用途，其中该 H19 增强子包含 SEQ ID NO: 4 的序列。
8. 权利要求 2 的用途，其中该 H19 增强子包含 SEQ ID NO: 5 的序列。
9. 权利要求 2 的用途，其中该 H19 增强子被置于异源序列的 3'
20 侧。
10. 权利要求 1 - 9 中任意一项的用途，其中该癌症是膀胱癌。
11. 一种载体，该载体含有一种包含与编码细胞毒性基因产物的异源序列操作性连接的调节序列的多核苷酸，其中该调节序列来自 H19 调节元件。
- 25 12. 权利要求 11 的载体，其中该 H19 调节序列包括 H19 启动子、H19 增强子或 H19 启动子和 H19 增强子两者。
13. 权利要求 12 的载体，其中该 H19 启动子包含 SEQ ID NO: 1 的 1 - 830 位的核苷酸。
14. 权利要求 12 的载体，其中该 H19 启动子包含 SEQ ID NO: 2
30 的序列。
15. 权利要求 12 的载体，其中该 H19 增强子包含克隆在 ATCC 保藏号 209322 的质粒 pH19EH19 中的 H19 增强子的序列。

-
16. 权利要求 12 的载体, 其中该 H19 增强子包含 SEQ ID NO: 3 的序列。
 17. 权利要求 12 的载体, 其中该 H19 增强子包含 SEQ ID NO: 4 的序列。
 - 5 18. 权利要求 12 的载体, 其中该 H19 增强子包含 SEQ ID NO: 5 的序列。
 19. 权利要求 12 的载体, 其中该 H19 增强子被置于异源序列的 3'侧。
 20. 包含权利要求 11 - 19 的任意一项的载体的宿主细胞。

诱导肿瘤特异性细胞毒性的方法和组合物

本申请是 1997 年 10 月 3 日提交的未决申请专利系列号 08/943,608 5 的部分继续申请，其全文在此引入作为参考。

1. 发明领域

本发明属于肿瘤细胞生物学和癌症治疗领域。更具体而言，本发明涉及异源基因尤其是编码细胞毒产物的基因在肿瘤细胞中的特异性表达。

10 2. 发明背景

2.1 H19 基因

H19 基因是少数已知的人类印记基因之一 (Hurst 等, 1996, 自然遗传学 12: 234-237)。在胚胎发生的最早期, H19 在 2 个染色体等位基因上都表达 (DeGroof 等, 1994, 滋养层 8: 285-302)。自此不久, 出现父系等位基因沉默, 而仅母系遗传的等位基因被转录。 15

H19 在胚胎发生期广泛表达, 并首先被鉴定为通过反式作用基因座 *raf* 在肝中和甲胎蛋白协同调节的基因 (Pachnis 等, 1984, 美国国家科学院院报 81: 5523-5527)。另外, H19 已被许多研究组用目的是分离在组织分化期表达的基因的筛选方法独立克隆。例如, Davis 等 20 (1987, 细胞 51: 987-1000) 在筛选 C3H10T1/2 细胞的分化过程中早期活性基因时鉴定了 H19 的小鼠同源物。Pourier 等 (1991, 发育 113: 1105-1114) 发现鼠 H19 在干细胞分化期间和植入期表达。人 H19 基因的转录在人胎盘滋养层细胞分化中也发现 (Rachmilewitz 等 1992, 分子生殖发育 32: 196-202)。

25 尽管 H19 RNA 的转录在整个胎儿期和胎盘发育期许多不同的胚胎组织中发生, H19 表达在出生后下调。然而, 在鼠成年肌肉和肝中报道了相对低水平的 H19 转录 (Brunkow 和 Tilghman, 1991, 基因和发育 5: 1092-1101)。在癌细胞中 H19 也在出生后活化。Ariel 等人 (1997, 分子病理学 50: 34-44) 表明来源于出生前 H19 表达的组织 30 的许多肿瘤中 H19 表达升高。另外, 这些作者发现来源于神经组织的肿瘤尤其是星形细胞瘤和成神经节胶质细胞瘤中有 H19 RNA, 其中这被已知与 H19 的表达无关。鉴于大量癌细胞表达 H19 RNA, 这些作

者推测 H19 是一种癌胚 RNA 并建议研究作为人恶性肿瘤的肿瘤标志的 H19。

已经克隆和测序了人和鼠 H19 基因 (Brannan 等, 1990, 分子细胞生物学 10: 28-36)。比较人和小鼠 H19 基因显示全长 77% 的核苷酸序列一致性。尽管种间核苷酸同源性的保守性, 根据两种基因的开放读框可预测非常低的氨基酸序列一致性 (同上)。而且, 虽然 H19 RNA 被 RNA 聚合酶 II 转录, 然后拼接和聚腺苷化, 但似乎它不被翻译。反之, 发现 H19 RNA 和 28S 细胞质 RNA 有关, 导致人们推测 H19 RNA 可能作为核糖核苷酸蛋白的 RNA 组分起作用 (同上)。

H19 的真正生理作用还没有完全了解。H19 能作为一种显性致死基因起作用; H19 转基因的高异位表达引起小鼠将要出生前的致死性 (Brunkow 等, 上文)。该致死期恰与 H19 的表达被抑制的时间相一致。另一方面, 在携带 H19 敲除等位基因的杂合或纯合的小鼠中, 没有观察到缺陷 (Leighton 等, 1995, 自然 375: 34-39)。敲除母系遗传的等位基因确实干扰遗传连锁和相反印记 IGF-2 基因的印记过程; 由于增加的出生前 IGF-2 的表达, 得到的小鼠出生时比同窝鼠更大 (同上)。因为这 2 个相反印记基因具有顺式作用的调节序列, Leighton 和同事们推测 H19 可能涉及 IGF-2 基因的印记。

提出的 H19 基因产物的另一个功能是肿瘤抑制 RNA。Hao 等人 (1993, 自然 365: 764-767) 报道用 H19 表达构建体转染的 2 种胚胎性肿瘤细胞系, RD 和 G401, 导致在裸鼠中细胞生长阻滞, 形态改变和降低致癌性。已注意到上述肿瘤抑制活性与小鼠中观察到的异位表达的致死性 (Hao 等, 同上) 以及敲除母系 H19 等位基因的小鼠增加体积 (Leighton 等, 同上) 的结果一致。然而, H19 是肿瘤抑制因子的建议存在争议。某些结果被报道不能重复, 可能存在另一种与 H19 紧密连锁的候选肿瘤抑制基因 (Areil 等, 同上)。建议的 H19 作为肿瘤抑制因子的作用还与以下实验数据相抵触, 即 H19 在广泛种类的肿瘤细胞中活化 (如见例如 Lustig-Yariv 等, 1997, 癌基因 23: 169-177)。

2.2 胰岛素样生长因子 (IGF) 基因

IGF-2 是另一种其表达依赖于它们父系来源的印记基因。不过与 H19 相反, 在小鼠和人类中的 IGF-2 为母系印记, 因此从父系遗传的

等位基因表达 (Rainier 等, 1993, 自然 363: 747-749)。人 IGF-2 基因显示复杂的转录类型。有 4 种在组织中和以发育特异的方式活化的 IGF-2 启动子。在胚胎发育过程和癌组织中, 仅 3 种启动子 P2、P3 和 P4 被印记和活化。第 4 种启动子 P1 不被印记, 仅在成年肝和脉络膜丛中活化 (见 Holthuisen 等, 1993, 分子生殖发育 35: 391-393)。IGF-2 基因的 P3 启动子与肝硬变和肝细胞癌的发展有关 (Kim 和 Park, 1998, 韩国医学科学杂志 13: 171-178)。

IGF-2 印记的丢失与 Wilm's 肿瘤有关 (Ogawa 等, 1993, 自然 363: 749-751)。该观察结果导致许多研究者推测印记的丢失和印记基因的双等位表达可能涉及癌的生长失调和发展 (也见 Rainier 等, 1993, 自然 362: 747-749 和 Glassman 等, 1996, 癌症遗传细胞遗传杂志 89: 69-73)。

2.3 肿瘤特异的基因治疗

来自肿瘤有关的基因的调节序列被用于选择地导向表达在衍生于肿瘤细胞中的自杀基因。例如, 在肝细胞癌中诱导甲胎蛋白的表达。Huber 等 (1991, 美国国家科学院院报 88: 8039-8043) 用来自白蛋白基因或胎儿蛋白基因的控制序列将水痘带状疱疹胸腺嘧啶激酶 (VZV TK) 基因编码序列在肝癌细胞中导向表达。用含有这些表达构建体之一的逆转录病毒载体感染的肝癌细胞表达 VZV TK 而变得对正常无毒性的药物原 6-甲基嘌呤阿拉伯核苷 (araM) 敏感。Kaneko 等, (1995, 癌症研究 55: 5283-5287) 构建了在甲胎蛋白控制序列控制下表达 HSV TK 的腺病毒载体。含有该载体的重组腺病毒颗粒被直接注射到在无胸腺裸鼠中产生的衍生于肝细胞癌的肿瘤。接着用 9-(1,3-二羟基-2-丙氧甲基) 鸟嘌呤腹膜内注射引起衍生于肝细胞癌的肿瘤消退。

Osaki 等 (1994, 癌症研究 54: 5258-5261) 把一种表达构建体转染到 A549 肺癌细胞中, 该表达构建体含有连接负责简单疱疹病毒胸腺嘧啶激酶 (HSV TK) 的编码序列的肺癌胚抗原基因的控制序列。所转染的细胞对 9-(1,3-二羟基-2-丙氧甲基) 鸟嘌呤敏感。另外, 在裸鼠中来自皮下注射转染细胞的肿瘤生长被重复的腹膜内注射 9-(1,3-二羟基-2-丙氧甲基) 所抑制。不过, 最近描述该癌胚抗原基因在正常结肠粘膜中表达, 因而限制了这些控制序列作为肿瘤特异调节

区的有用性 (Osaka 等, 同上)。因此, 开发在肿瘤细胞中特异性表达基因产物的基因治疗载体仍有必要。

3. 发明概述

本发明涉及诱导在肿瘤细胞中选择性表达异源基因的方法和组合物。具体而言, 本发明涉及包含操作性与异源基因连接的调节转录序列的多核苷酸, 它引起异源基因的肿瘤特异表达。更具体而言, 该调节转录序列来源于在癌细胞中特异表达的基因组印记基因如 H19, IGF-2 P3 和 P4 启动子, 而异源基因编码细胞毒蛋白或细胞抑制基因产物。在本发明的另一个实施方案中, IGF-1 启动子与异源基因操作性相连接, 以导致肿瘤特异的基因表达。该调节序列将调节在许多不同癌细胞类型中的基因表达。上述方法和组合物可用于广泛种类的癌症和增殖过度病情的治疗。

本发明还涉及一种包含与调节序列操作性连接的编码细胞毒性基因产物的异源序列的多核苷酸在制备用于治疗患者的癌症的药物中的用途, 其中该调节序列来自 H19 启动子调节元件。

本发明的另一方面是含有多核苷酸的表达载体, 该多核苷酸包括上述与异源基因操作性连接的调节区。特别优选的调节区是那些编码 H19 调节区如启动子和增强子序列, IGF-2 P3 或 P4 启动子或 IGF-1 启动子。从那种连接方式, H19 的增强子及其活性部分可用于与 H19 启动子、IGF-1 启动子、IGF-2 P3 启动子或 IGF-2 P4 启动子形成任一种组合。本发明还包括含有上述载体的宿主细胞。关于这方面, 含有由具有或没有 H19 增强子的 H19 启动子控制的异源基因的表达构建体可以和第 2 种构建体共导入到细胞中, 其中该第 2 种构建体含有由与 H19 增强子组合在一起的 IGF-1 启动子或 IGF-2 P3 或 P4 启动子控制的异源基因。在另一方面, 本发明提供用上述载体在肿瘤细胞中表达异源基因的方法。而本发明的另一方面是在基因治疗中利用本发明的载体治疗癌症。

4. 附图简述

图 1A-1C. 人 H19 启动子区的核苷酸序列。显示从核苷酸位 -837 -7 (相对于转录起始点) 的启动子区 (SEQ ID NO: 1)。

图 2. 在 H19 调节序列的控制下用所述的载体表达异源基因的简图。

图 3A-3E. H19 调节序列指导异源基因 (CAT) 在膀胱癌细胞系中的表达。对于 5 种所示的不同细胞系, CAT 的比活 (cpm/ μ g 蛋白) 以所用转染的载体的函数作图。图 3A: HT-1376 细胞、图 3B: EJ28

细胞、图 3C: T24P 细胞、图 3D: 1197 细胞、图 3E: UM-UC-3 细胞。下面的载体将在下面第 6 节更完整地描述: (1) pCAT-基础; (2) pCAT-对照; (3) pH19E; (4) pH19EH19D; 和 (5) pH19EH19R。

5 图 4A-4E. IGF-2 P3 和 P4 启动子指导异源基因在膀胱癌细胞系中的表达。对于 5 种所示的不同细胞系, 萤光素酶的比活 (每 μg 蛋白的计数) 以在转染构建体中所用的指导萤光素酶表达的 IGF-2 启动子的函数作图表示。图 4A: T24P 细胞; 图 4B: 1376 细胞; 图 4C: UM-UC3 细胞; 图 4D: 1197 细胞; 图 4E: EJ28 细胞。所述的载体
10 在下面第 10 节更完整地描述。

图 5. 人 H19 启动子片段的核苷酸序列 (SEQ ID NO: 2)。

图 6. 0.9 kb 的 H19 增强子片段的核苷酸序列 (SEQ ID NO: 3)。

图 7A 和 7B. 2 kb 的 H19 增强子片段的核苷酸序列
(SEQ ID NO: 4)。

15 图 8A-8C. 4 kb 的 H19 增强子片段的核苷酸序列
(SEQ ID NO: 5)。

图 9A-9C. 用含有各种 H19 调节区和 P4 启动子组合的载体的转染指导萤光素酶在肿瘤细胞中表达。图 9A: 5637 细胞; 图 9B: Huh 7 细胞; 图 9C: 293T 细胞。

20 图 10A-10E. 用含有 H19 调节区的载体的转染指导萤光素酶在肿瘤细胞中表达。图 10A: 293T 细胞; 图 10B: T24P 细胞; 图 10C: Huh 7 细胞; 图 10D: 5637 细胞; 图 10E: RT112 细胞。

5. 发明详述

本发明部分基于如下的发现, 来源于在肿瘤细胞中表达的基因组
25 印记基因的调节区能用于导向目的编码序列在癌细胞中表达。具体而言, 已经发现 H19 的表达在各种类型的肿瘤中活化, 这些肿瘤包括但不限于膀胱癌、肝细胞癌、成肝细胞癌、成横纹肌细胞瘤、卵巢癌、颈部癌、肺癌、乳腺癌、头颈部扁平细胞癌、食管癌、类胸腺癌、星形细胞瘤、成神经节胶质神经瘤和成神经细胞瘤。而且发现含有操作
30 性与一种异源基因相连的 H19 启动子区或操作性与一种异源基因相连的 IGF-2 P3 或 P4 启动子的构建体, 或者含有启动子与 H19 增强子下游结合在一起的构建体均在肿瘤细胞中被特异性活化。在本发明

的另一方面，IGF-1 的启动子被用于指导异源基因的表达。

因此，在本发明的一个方面中，本发明提供改变癌细胞表型或选择性杀死癌细胞的方法和组合物。本目的通过把一种多核苷酸输送到所述的细胞来完成，所述的多核苷酸包含来自操作性地与异源基因
5 连结在一起，在癌细胞中表达的基因组印记基因的调节区。例如，该异源基因能编码一种细胞抑制剂或细胞毒剂（如，毒素、反义 RNA 或核酶）。

在癌细胞中表达的来自基因组印记基因的调节区包括但不限于 H19 启动子和增强子，以及 IGF-2 P3 和 P4 启动子。

10 针对在此所描述的本发明的目的，术语“操作性连接”意味着核苷酸序列用如下的方式与调节序列连接在一起，该方式允许所述的核苷酸序列的表达受调节序列指导。

为了本申请的目的，“异源的”基因序列指的是该基因序列一般情况下未与 H19 基因的调节序列操作性地连接在一起。大体来说，异
15 源基因序列包括编码抑制细胞和细胞毒基因产物的序列。

如本文所采用的术语“表达”，意味着目的 DNA 的转录和剪接、加工、稳定性以及可选择的相应的 mRNA 转录体的翻译。根据所传递的 DNA 分子结构，表达可以是瞬时或连续的。

5.1 H19 基因的调节序列，IGF-2 P3 和 P4 启动子和 IGF-1 20 启动子

在此描述的是 H19 调节序列，该调节序列能用于指导异源编码序列在肿瘤细胞中特异性表达。这些 H19 的调节序列包括上游 H19 启动子区和/或下游 H19 增强子区。一种 H19 启动子区的核苷酸序列在图 1A-1C 显示 (SEQ ID NO: 1)。该 830 个核苷酸的序列从 -837
25 延长到帽部位点的 -7（如 Brannan 等的描述，上文）。一个共有的 TATA 序列出现在核苷酸 -27~-35 处。2 个共有的 AP2 结合位点 (8/9 匹配) 出现在转录起始位点上游大约 -500 和 -40 的核苷酸处。当如
在下面更详细讨论的那样被置于异源基因编码区的上游时，大约 830 个碱基对的调节区足以指导操作性连接的异源基因在也表达内源 H19
30 的癌细胞中表达。另外，另一个位于核苷酸 -819-+14 位核苷酸的 H19 启动子区（图 5，SEQ ID NO: 2）也足以指导操作性连接的异源基因在癌细胞中表达。

5 为了提供增强水平的肿瘤细胞特异性的表达，可任选将人 H19 基因的下游增强子区加到 H19 启动子/异源基因的构建体上。如在第 6 节以实施例的形式更详细描述和阐明的，下游增强子区包含在相对于转录起始位点从 +6 kb 延长到 +11 kb 的 SacI 限制性片段上。根据增强子序列预计，当在 H19 启动子控制之下，以反向或正向（相对于内源 H19 基因中的 H19 增强子的方向）置于异源基因编码区的下游时，该下游增强子能发挥其作用。另外，含有在图 6、7A、7B 和 8A-8C（SEQ ID NOS 3-5）所示的序列的该增强子片段也可用于方便基因表达。

10 IGF-1 基因的表达和肺癌及乳腺癌有关。IGF-1 启动子 R（在人 IGF-1 基因序列中位于核苷酸 1-1630 的核苷酸序列（基因库登记号 M12659 M77496），在此引入以供参考；Rotwein 等，1986，生物化学杂志 261：4828-4832）。

15 用 4 种不同的启动子区之一表达 IGF-2 的基因产物。在胚胎组织中这 4 种启动子的 3 种被印记和表达；然而，启动子 P1 仅在成年组织中活化（Sussenbach 等，1992，生长调节 2：1-9）。P3 启动子与肝癌有关。也发现印记的 P4 启动子（IGF-2 基因的核苷酸序列 -546~+102）和 P3 启动子（IGF-2 基因的核苷酸序列 -1229~+140）在人膀胱癌细胞中活化，并且可用于指导操作性连接的异源基因对肿瘤细胞的表达。IGF-2 P3 和 P4 启动子可与 H19 增强子或其活性片段结合在一起使用。

25 这些来自在癌细胞中表达的基因组印记和非印记基因的调节序列能进一步去线性化，以确定获得所希望的肿瘤特异性表达所需的最短调节序列。例如，通过添加、替换或缺失可改变启动子区，然后分析其保留的特异于肿瘤的表达功能。可逐个检测 H19 下游增强子的各个部分增强从 H19 启动子转录的能力。

30 用各种化学和酶方法能引起调节序列中的各种改变，这些方法在本领域的技术人员中众所周知。例如，由限制性位点确定的序列区能被缺失。能用寡核苷酸定点诱变方法以确定的方式改变序列和/或把限制性位点导入到序列特定区中。另外，用诸如 Bal 31 或 ExoIII 和 S1 核酸酶的 DNA 核酸酶能产生缺失突变体。通过把所述的 DNA 与核酸酶温育更长的时间逐步产生调节序列中更大的缺失（见 Ausubel 等，

1989, 分子生物学现代技术, 见诱变技术综述)。

评价改变的序列在合适的宿主细胞, 尤其是表达 H19 衍生于癌, 细胞 (例如膀胱癌细胞, 称为一个实施例) 中, 指导异源编码序列的特异于肿瘤的表达能力。为进一步用途保留其指导肿瘤特异表达能力并结合到重组表达载体中的任何改变的调节序列在本发明的范围之中。在这些调节序列的控制下能表达各种类型的异源基因, 如编码毒性基因产物、潜在的毒性基因产物、如抗增殖或抑制细胞的基因产物的基因。也能表达标志基因, 它们包括酶 (如 CAT、 β -半乳糖苷酶、荧光素酶)、如绿色荧光蛋白的荧光蛋白或抗原标记分子。

10 广义来说, 细胞毒基因产物定义为包括毒素和诱导凋亡剂。此外, 针对本发明的目的, 细胞毒基因产物包括把前药转变成细胞毒产物的药物代谢酶类。用在本发明方法中的细胞毒基因产物的实施例包括白喉毒素、假单胞菌毒素、蓖麻毒蛋白霍乱毒素、PE 40 和肿瘤抑制子基因如视网膜成细胞瘤和 p53。此外, 也可用编码诱导细胞凋亡的凋亡肽的序列。上述凋亡肽包括阿尔海默茨 A β 肽 (见 LaFerla 等, 1995, 自然遗传学 9: 21-30)、前房钠尿肽 (见 Wu 等, 1997, 生物化学杂志 272: 14860-14866)、降钙素基因相关肽 (见 Sakuta 等, 1996, 神经免疫学杂志 67: 103-109)、以及其他已知的或待发现的凋亡肽。

20 把前药转变成细胞毒产物的药物代谢酶包括胸腺嘧啶激酶 (来自单纯疱疹病毒或水痘带状疱疹病毒)、胞嘧啶脱氨酶、硝酸还原酶、细胞色素 p-450 2B1、胸腺嘧啶磷酸化酶、嘌呤核苷磷酸化酶、碱性磷酸酶、羧肽酶 A 和 G2、亚麻苦苷酶 (linamarase)、 λ 内酰胺酶、黄嘌呤氧化酶 (见 Rigg 和 Sikora, 1997 8 月今日分子医学, pp359-366 的有关背景资料)。

25 此外, 用本发明描述的表达构建体可把反义、反基因或 aptameric 寡核苷酸传递到癌细胞中。核酶或单链 RNA 也能在癌细胞中表达以抑制特殊的目的基因的表达。这些反义或核酶分子作用的靶基因应该是那些对细胞维持或维持癌细胞表型所必需的编码基因产物。上述靶基因包括但不局限于 cdk2、cdk8、cdk21、cdc25A、周期蛋白 D1、周期蛋白 E、周期蛋白 A 和 cdk4。

例如, 为了下调内源基因的表达, 把下列载体、反义 RNAS 或核

酶导入到细胞中，其中所述的载体在来源于在癌细胞中表达的印记基因或 IGF-1 启动子的调节序列控制下；而反义 RNAs 或核酶对癌基因形式的 p53、c-fos、c-jun、kr-ras 和/或 Her2/neu 的转录体具有特异性。表达 H19 和能活化 H19 调节序列（或特异性地活化 IGF-1、
5 IGF-2 P3 或 P4 启动子）的肿瘤细胞能特异地被导向用于表达反义 RNA 或核酶 RNA。

反义方法包括设计与靶 mRNA 互补的寡核苷酸（在此情况下，mRNA）。该反义寡核苷酸将结合到互补的靶 mRNA 转录体，然后阻止翻译。绝对的互补性，尽管是优选的，并不需要。在此所指的对部分 RNA 的序列“互补性”意味着一种具有能和 RNA 杂交，形成稳定的双链体的足够互补性的序列。杂交能力将依赖于互补性的程度和反义核酸的长度。一般来说，所杂交的核酸越长，它可能含有的与 RNA 形成的碱基错配就越多，但仍能形成稳定的双链体（或三联体，视具体情况而定）。本领域的技术人员可使用标准方法确定错配的可容许
10 程度，以确定杂交复合体的熔点。

与靶信使 RNA 的 5'末端互补的寡核苷酸，例如直到并包括 AUG 起始密码子的 5'非翻译序列，应最有效地起抑制翻译的作用。然而，最近研究表明，与 mRNAs 的 3'非翻译序列互补的序列也有效地抑制 mRNAs 的翻译。一般性地参见，Wagner R, 1994, 自然 372: 333-
20 335。因此，与靶基因转录体的 5'-或 3'-非翻译、非编码区互补的寡核苷酸可用于反义方法，以抑制内源基因的翻译。与 mRNA 的 5'非翻译区互补的寡核苷酸应包括 AUG 起始密码子的互补部分。与 mRNA 编码区互补的反义寡核苷酸是不太有效地翻译抑制剂，但根据本发明可以使用。不管被设计成与靶 mRNA 的 5'、3'或与编码区杂交，反义
25 核酸至少长为 6 个核苷酸，优选地为长 6-约 50 个核苷酸范围的寡核苷酸。具体来说，该寡核苷酸至少有 10 个核苷酸、至少 17 个核苷酸、至少 25 个核苷酸或至少 50 个核苷酸。

不管如何选择靶序列，首先进行体外研究以定量确定反义寡核苷酸抑制基因表达的能力是优选。这些研究应该使用对照组，该对照组
30 区分寡核苷酸的反义基因抑制和非特异的生物学效应。将所述的靶 RNA 或蛋白水平与内部对照 RNA 或蛋白的水平进行比较的这些研究方法，也是优选的。

设计催化性裂解必需靶基因的核酶分子也能用于阻止靶 mRNA 的翻译。(如见, PCT 国际公开 WO 90/11364, 1990 年 10 月 4 日公开; Sarver 等, 1990, 科学 247: 1222-1255)。当该核酶特异于一种编码对癌细胞生长必需的基因转录体时, 上述核酶能引起癌细胞表型的逆转。虽然在特异识别序列的位点裂解 mRNA 的核酶可被用于摧毁靶 mRNAs 时, 但优选地使用锤头核酶。锤头核酶在与靶 mRNA 形成互补碱基对的侧翼区所示的位置裂解 mRNAs。唯一的要求是靶 mRNA 具有下列 2 个碱基的序列: 5'-UG-3'。锤头核酶的构建和产生在本领域众所周知, 并且在 Haseloff 和 Gerlach, 1988, 自然, 334: 585-591 的方法中更完整地描述。优选工程化核酶, 这样裂解识别位点位于靶 mRNA 的 5'末端附近; 即能增加有效性并使细胞内非功能性 mRNA 转录体的积累最少。

用于本发明的核酶也包括 RNA 内切核糖核酸酶(下文被称为“Cech 型核酶”), 例如在嗜温四膜虫中天然产生的核酶(称之为 IVS, 或 L-19 IVS RNA), 该酶由 Thomas Cech 和合作者进行了深入的描述(Zaug 等, 1984, 科学 224: 574-578; Zaug 和 Cech, 1986, 科学, 231: 470-475; Zaug 等, 1986, 自然, 324: 429-433; 大学专利公司出版的公开的国际专利申请号 WO 88/04300; Been 和 Cech, 1986, 细胞, 47: 207-216)。Cech 型核酶拥有一个 8 碱基对的活性位点, 该位点与靶 RNA 序列杂交, 然后出现靶 RNA 的裂解。本发明设想使用靶定存在于靶基因中的 8 个碱基对活性位点序列的那些 Cech 类型的核酶。

5.2 在肿瘤细胞中的基因活化

重新活化印记基因表达的细胞也能特异地活化表达构建体, 其中该表达构建体含有操作性与异源基因相连的上述印记基因调节区。上述细胞, 尤其是肿瘤细胞, 是本发明基因治疗方法的合适靶物。应用 RNA 分析、原位杂交和报道基因构建体技术, 可以测定 H19 和 IGF-2 P3 和 P4 在肿瘤和细胞系中的特异表达。另外, 用指导异源基因表达的 IGF-1 启动子、可类似地在基因治疗中确定和导向具有活化的 IGF-1 基因表达的肿瘤细胞。

对于大部分 RNA 分析应用, 用本领域熟知的任何技术制备特异与目的基因转录体杂交的标记探针。该标记探针可含有至少 15-30 个

与 H19 核苷酸序列互补的碱基，更优选地含有与 H19 转录体互补的 50 - 150 个碱基。针对 H19 表达的特别优选的杂交探针是一种与 H19 信使 RNA 的 3' 末端互补，从聚 A 位点上游大约 800 个碱基对到聚 A 位点的多核苷酸。

- 5 在以下由实施例方式说明的本发明具体实施方案中，用 T7 或 T3 表达质粒体外产生标记过的反义 RNA 探针。在标记核苷酸的存在下通过随机引物法，例如用 Prime-It 试剂盒 (Stratagene, La Jolla, CA; 商品名号 300392)，也能标记 H19 探针。另外，标记探针也能用 PCR 反应或标准缺刻翻译反应的方法产生，其中在 PCR 反应中，用 H19 编码区的 cDNA 克隆和被设计用来扩增编码区的某一区域的探针。

适于多核苷酸探针的标记物包括掺入了放射性同位素 (如 ^{35}S 和 ^{32}P)、荧光素、发光和显色标签分子和酶组分的核苷酸。

- 用标准技术原位把标记探针与细胞或组织样品杂交，例如用以下实施例描述的方法及本文引作参考的未决美国专利申请系列号 15 08/704,786 的方法。另外，如果能获得足够量的合适细胞，可进行标准 RNA 分析 (如 Northern 分析、RNA 酶保护法或 RNA 酶引物延伸法)，测定目的基因 mRNA 的表达水平。

- 另外，有可能“原位”进行上述基因表达分析，即直接从活检或切除物获得的病人组织的组织切片 (固定和/或固定) 上进行，这样没有 20 有必要纯化核酸。诸如上面描述的那些核酸试剂可在这些原位方法中用作探针和/或引物 (例如见, Nuovo, G.J.1992, “原位杂交的 PCR 法: 方法和应用”, Raven 出版社, 纽约)。

- 确定一种细胞类型或肿瘤能特异地活化含有操作性与异源基因相连的特异调节区的表达构建体的一种替代方法事实上是把上述表达构建体转染到细胞中。为这些目的，异源基因优选是一种标记基因产物。 25 在一次对标记基因产物分析中的阳性结果显示：细胞或细胞系能从该调节区活化表达。

用这些技术，具有活化的 H19 表达的实例性肿瘤类型表示如下：

A. 儿童实体瘤

- 30 1. Wilm's 肿瘤
 2. 肝胚细胞瘤
 3. 胚胎性成横纹肌细胞瘤

B. 胚细胞肿瘤和滋养层肿瘤

1. 睾丸胚细胞瘤
2. 非成熟卵巢畸胎瘤
3. 骶尾瘤
- 5 4. 绒毛膜癌
5. 胎盘点滋养层肿瘤

C. 成人表皮肿瘤

1. 膀胱癌
2. 肝细胞癌
- 10 3. 卵巢癌
4. 颈部癌
5. 肺癌
6. 乳腺癌
7. 头颈部扁平细胞癌
- 15 8. 食管癌
9. 类胸腺癌

D. 起源于神经的肿瘤

1. 星形细胞瘤
2. 成神经节胶质神经瘤
- 20 3. 成神经细胞瘤

相应地，通过本发明的方法可治疗上述癌症。事实上，活化 H19 表达的任何肿瘤可通过本发明的方法治疗。

另外，上面提到的技术可用于确定活化 IGF-1 和 IGF-2 P3 及 P4 启动子的肿瘤。上述肿瘤也可通过本发明的方法治疗。例如，
25 在儿童肿瘤中，诸如 Wilm's 瘤、成横纹肌细胞瘤和成肝细胞瘤中，IGF-2 被活化。

5.3 把在调节序列控制下的多核苷酸导入到宿主细胞的方法

本发明也涉及一种用含有操作性与异源基因连接的调节区的多核苷酸转染的宿主细胞。上述宿主细胞可用培养方法维持，或者是动物
30 的一部分，优选哺乳动物。一般来说，用本发明披露的方法和材料，目的多核苷酸可插入到随后被传递的各种载体的任一种中。用完全可靠的分子生物学技术（一般见，Sambrook 等（1989）分子克隆 I-III

卷,冷泉港实验室出版社,冷泉港,纽约和分子生物学当代方法(1989), John Wiley & Sons, 所有卷和定期的最新资料,在此引入仅供参考), 能产生上述载体。一般来说,如果需要翻译,也可改造目的异源基因以包含一种如有必要的合适的 3'聚腺苷化序列。

5 5.3.1 培养细胞

用含有操作性与异源基因连接的印记基因调节区的多核苷酸转染的宿主细胞可以是任一种原核或真核细胞。把多核苷酸连接到基因构建体如载体中,然后转化或转染到真核(酵母、鸟类、昆虫或哺乳动物)或原核(细菌细胞)的宿主细胞中是在微生物或组织培养技术中广泛使用的标准方法。

适于在诸如大肠杆菌的细菌细胞中培养目的多核苷酸的载体包括如下类型的质粒:来源于 pBR322 的质粒、来源于 pEMBL 的质粒、来源于 pEX 的质粒、来源于 pBTac 的质粒以及来源于 pUC 的质粒。为了在酵母中复制, YEP24、YIP5、YEP51, pYES2 和 YRP17 质粒是用于把遗传构建体导入到啤酒糖酵母中的克隆和表达载体(例如见, Broach 等, 1993, 基因表达的实验操作, 由 M. Inouge 编著, 学术出版社, 83 页)。这些载体由于存在 pBR322 ori 能在大肠杆菌中复制, 并且由于酵母 2 μ m 环形质粒的复制决定子能在酵母中复制。此外,也可用如氨苄霉素的药物抗性标记。

相似地,用于本发明的多核苷酸的优选的哺乳动物载体含有 2 种原核生物的序列,以便于该载体在细菌中繁殖。应用连接的选择性标记基因,设计上述载体当转染到哺乳动物细胞中时,以长期稳定的方式整合到哺乳动物的染色体上。另外,诸如牛乳头瘤病毒(BPV-1)或 Epstein-Barr 病毒等病毒的衍生物能用于瞬时表达。用于制备宿主生物质粒转化的各种方法在本领域众所周知。有关合适的载体系统以及一般重组方法见 Sambrook 等,见上文。

5.3.2 基因治疗

本发明也包括含有操作性连接用于基因治疗的异源基因的基因调节区的多核苷酸在治疗癌症和过度增殖疾病中的用途。为了基因治疗的目的,本发明的表达构建体可用任一种生物有效载体施用,如能体内有效地把重组基因传递到细胞的任一种制剂或组合物。这些方法包括把所述的基因插入到病毒载体或重组的细菌或真核生物的质粒中,

所述的病毒载体包括重组逆转录病毒、腺病毒、腺病毒有关的病毒、单纯疱疹病毒-1。病毒载体直接转染细胞；质粒 DNA 用下列物质辅助下传递，例如阳离子聚合物、阳离子脂质体（如 lipofectin，胆固醇衍生物如 D. D. A. B. 和阳离子磷脂）或其衍生物（如连接抗体）、聚赖氨酸偶联物、短杆菌肽 S、人工病毒包被蛋白或其他上述细胞内载体；也可以体内进行裸基因构建体的直接注射、电穿孔或 CaPO_4 沉淀法。用于癌症基因治疗的基因转移和表达系统的最新综述见 Cooper, 1996, 肿瘤学讨论会 23: 172-187。

应当理解，由于转导合适的靶细胞代表基因治疗的重要的第一步，具体的基因传递系统的选择性将依赖于目标靶物的表型和如局部或全身给药途径等这些因素。而且，应当认为，提供用作体内转导表达构建体的具体基因构建体也用于体外转导细胞，例如用在上面描述的离体组织培养系统中。

优选的体内转导核酸到细胞中的方法是使用含有核酸的病毒载体，例如该核酸为一种在 H19 调节序列控制下的特异细胞毒基因。用病毒载体感染细胞的优点是大部分靶细胞能接受核酸。此外，在病毒载体中编码的分子，如在病毒载体中含有的 cDNA，可在接受病毒载体核酸的细胞中有效表达。可用本发明披露的方法和组合物传递的合适载体，包括但不限于单纯疱疹病毒载体、腺病毒载体、腺病毒有关的病毒载体、逆转录病毒载体、假狂犬病毒载体、阿尔发-疱疹病毒载体等。病毒载体尤其是适用于修饰非复制细胞的病毒载体，以及如何使用上述载体目的多核苷酸表达相连接的方法的完整综述，在“病毒载体：基因治疗和神经科学应用”Caplitl 和 Loewy 编著，学术出版社，圣地亚哥（1995）一书中可找到。

已经表明：通过修饰病毒颗粒表面的病毒包装蛋白（例如见 PCT 公开 WO 93/25234 和 WO 94/06920），有可能限制病毒和接着基于病毒的载体的感染谱。例如，修饰逆转录病毒载体的感染谱的策略包括：将特异于细胞表面抗原的抗体耦联至病毒 env 蛋白中（Roux 等，1989，美国国家科学院院报 86: 9079-9083；Julan 等，1992，普通病毒学杂志，u3: 3251-3255；和 Goud 等，1983，病毒学 163: 251-254）；或将细胞表面受体配体耦联至病毒 env 蛋白中（Neda 等，1991，生物化学杂志；266: 14143-14146）。耦联方法可以是和蛋白或其他种类

的分子化学交联的形式(如把 env 蛋白转化成脱唾液酸糖蛋白的乳糖),以及产生融合蛋白(如单链抗体/env 融合蛋白)。例如,使用该技术,例如通过将抗肿瘤有关分子的抗体或癌细胞表面蛋白的抗体耦联至重组病毒表面可靶定癌细胞。该技术,在用于限制或另外指导对某些组织类型的感染的同时,也能用于把外向载体转变成兼嗜性载体。

优选的用于本发明的病毒基因传递系统使用衍生于腺病毒的载体。可操作处理腺病毒的基因组,使其编码和表达目的基因产物,而在正常裂解性病毒生活周期中的复制能力方面被失活。例如见 Berkner 等,1988,生物技术 6:616;Rosenfeld 等,1991,科学 252:431-434;和 Rosenfeld 等,1992,细胞 68:143-155。合适的来源于腺病毒株 AD 型 5 d1324 或其他腺病毒株(如 Ad2、Ad3、Ad7 等)的腺病毒载体在本领域的技术人员中众所周知。在某些情况下重组腺病毒是有利的,因为它们可用于感染各种类型的细胞类型,这些细胞包括气管表皮(Rosenfeld 等,1992,见上文的引用)、内皮细胞(Lemarchand 等,1992,美国国家科学院院报 89:6482-6486)、肝细胞(Herz 和 Gerard,1993,美国国家科学院院报 90:2812-2816)和肌肉细胞(Quantin 等,1992,美国国家科学院院报 89:2581-2584)。而且,病毒颗粒相对稳定,容易纯化和浓缩,并可被修饰以影响感染谱。此外,导入的腺病毒 DNA(和含在其中的外源 DNA)不整合到宿主细胞的基因组中但以附加体形式存在,因而避免可能出现的插入诱变的潜在问题,在插入诱变的情形中,导入的 DNA 成为整合到宿主基因组中(如逆病毒 DNA)。再者,相对于其他基因传递载体,腺病毒基因组携带外源 DNA 的能力大(高达 8 个碱基对)(Berkner 等,见上文的引用,Haj-Ahmand 和 Graham 1986,病毒学杂志 57:267)。普遍使用并因此受本发明推崇的大部分复制缺陷的腺病毒载体,被缺失所有或部分病毒的 E1 和 E3 基因,但仍保留多达 80%的腺病毒遗传物质(如见, Jones 等,1979,细胞 16:683;Berkner 等,同上;和 Graham 等,分子生物学方法, E. J. Murray 编著(Humana, Clifton NJ, 1991) 7 卷, 109-127 页)。

另一种用于本发明目的表达构建体传递之一的病毒载体系统是腺伴随病毒(AAV)。腺伴随病毒是一种天然产生的缺陷病毒,该病毒需要另一种病毒如腺病毒或疱疹病毒作为辅助病毒,以进行有效的复制和

增殖性生命周期（见 Muzyczka 等的综述，当代微生物学和免疫学专题 158：97-129）。它也是可以把其 DNA 整合到非分裂细胞并表现为高频率的稳定相互作用的少数病毒之一（例如见 Flotte 等，1992，美国呼吸协会细胞分子生物学杂志 7：349-354，Samulski 等，1989，病毒学杂志，63：3822-3828；和 McLaughlin 等，1989，病毒学杂志 63：1963-1973）。含有少达 300 个 AAV 碱基对的载体能被包装和整合。对外源 DNA 的空间限制为大约 4.5 kb。诸如在 Tratschin 等，1985，分子细胞生物学 5：3251-3260 中描述的一种 AAV 载体可用于把 DNA 导入到细胞中。用 AAV 载体，已把各种核酸导入到不同类型的细胞中（例如见 Hermonat 等，1984，美国国家科学院院报 81：6466-6470；Tratschin 等，1985 分子细胞生物学 4：2072-2081；Wondisford 等，1988，分子内分泌学 2：32-39；Tratschin 等，1984，病毒学杂志 51：611-619；和 Flotte 等，1993，生物化学杂志 268：3781-3790）。

除了如上述所描述的病毒输送方法外，也可以使用非病毒方法以引起所需的外源基因在动物组织中的定向表达。大多数基因输送的非病毒法依赖于由哺乳动物细胞用于吸收和胞内运送大分子的普通机理。在优选的实施方案中，本发明的非病毒基因传递系统依赖于靶向细胞吸收目的表达构建体的胞吞通路。该类型的说明性基因传递系统包括来源于脂质体的系统、聚赖氨酸偶联物和人工病毒包被。

在临床应用方面，用于治疗性表达构建体的基因传递系统可由众多方法之任一种导入到病人中，其中每一种方法在本领域众所周知。例如，基因传递系统的药用制剂可通过如静脉注射而全身性导入，并且由于异源基因的调节序列受控表达，或者与基因传递载体靶向具体细胞类型结合的调节序列，该构建体在靶细胞中的特异表达根据由细胞类型或组织类型表达所提供的转染特异性显著地发生。在另一种实施方案中，伴随着导入相当局限的动物中，重组表达构建体的初始传递更受限制。例如，可用导管（见美国专利 5,328,470）或立体注射法（例，Chen 等，1994，美国国家科学院院报 91：3054-3057）导入基因传递载体。可在基因治疗构建体中，用例如由 Dev 等，1994，癌症治疗综述 20：105-115 所描述的技术经电穿孔法传递本发明的表达构建体。

基因治疗构建体的药用制剂本质上由在可接受的稀释剂中的基因传递系统所组成，或者可包含一种缓释基质，其中所述基因传递载体包埋在该基质中。另外，在完整的基因传递系统可完整地从重組细胞产生时，例如逆转录病毒载体，该药用制剂可包括产生基因传递系统的一种或多种细胞。

5.3.3 治疗终点和剂量

普通技术人员应当懂得，从医生或病人的角度来看，任何基本上缓解或预防和癌症病情有关的非所希望的症状（如，疼痛、敏感性和体重减轻等）是合乎要求的。另外，任何肿瘤重量或生长率的降低以及改善肿瘤组织病理图像是所希望的。因此，关于本申请的目的，在此所用的术语“治疗”、“治疗用途”、或“医药用途”指的是所要求保护的组合物的任何及所有的用途，其中这些组合物治疗疾病状态或症状，或者换言之用任何提及的方法预防、阻止、延缓或逆转疾病的发展或其他非所需的症状。

有效剂量和治疗方案可通过常规方法确定，在实验室动物中用低剂量开始，然后增加剂量并同时监测作用效果，并也可以用全身性地改变剂量方案。动物研究，优选地哺乳动物研究通常被用来确定每公斤体重生物活性剂的最大可耐受剂量，或 MTD。本领域的技术人员通常外推有效剂量并避免对其他物种，包括人类的毒性。

进行对人的效能研究前，在正常个体中的 I 期临床研究有助于确立安全剂量。当确定对给定个体的最佳剂量时，临床医生可能考虑许多因素。这些因素中最主要的是所选择的异源基因产物的毒性和半衰期。其他因素包括病人的个体大小、病人的年龄、病人的一般情况、被治疗的具体癌症、疾病的严重度、病人所用的其他药物、基因产物的体内活性等。考虑动物研究的结果和临床文献后，选择实验剂量。

例如，含有操作性与一种异源基因连接的 H19 调节区的腺病毒载体的直接注射到肿瘤块内的一般人用剂量为 1×10^7 pfu - 1×10^{10} pfu/天，其中所述的异源基因编码一种诸如胸腺嘧啶激酶的细胞毒剂。更优选地，把上述腺病毒载体直接注射到肿瘤内的每日剂量为 1×10^8 pfu - 1×10^{10} pfu。对于一种含有操作性与不同水平毒性的细胞毒基因产物连接的 H19 调节区的腺病毒载体来说，这些值当然将作相应的改变。

也能用相似剂量的一种腺病毒载体作为表示性起点，其中该腺病毒载体含有操作性与编码如胸腺嘧啶激酶的细胞毒剂的异源基因的 IGF-2 P4 启动子。

尤其是考虑体内使用时，本发明的各种生化组分优选地具有高纯度并且基本上无潜在的有害污染物（例如，至少国家食品（NF）级、通常至少分析级和优选地至少药用级）。由于所确定的化合物必须使用前合成，上述合成或随后的纯化优选产生这样的产物，其中该产物在合成或纯化方法中基本上无任何潜在的所用的毒性剂。

对于用于治疗个体的癌症病情，本发明在一方面提供以无菌分装的小瓶或安瓿形式的一种试剂盒或试剂包，其中它含有与编码细胞毒剂的异源基因操作性连接在一起的 H19 调节区的多核苷酸载体或载体释放细胞。在一实施方案中，以即可给药的制剂的形式，该试剂盒包含单剂或复剂形式的含有一种操作性与编码细胞毒剂的异源基因连接的 H19 调节区的多核苷酸载体，其中所述的包装内含一种指导其内含物用于治疗癌症的标记物。另外，根据本发明的另一个实施方案，该试剂包提供一种无菌分装的含有上述释放载体的细胞或细胞系的小瓶或安瓿瓶。为了贮存和运输，应该冷冻载体释放细胞或细胞系。另外，该试剂包也可以含有用于培养所述的释放载体的细胞或细胞系的培养基和试剂。

描述本发明后，提供下列实施例用以说明而不是限制本发明。

6. 实施例：H19 调节序列有利于异源基因在肿瘤细胞系中的表达

本节描述含有置于 H19 调节序列的控制下的 CAT 报道基因的各种表达构建体的构建方法，以及将它们转移到几种不同的膀胱癌细胞系。

6.1 材料和方法

6.1.1 细胞系和转染方法

膀胱癌细胞系 HT-1376、EJ28、T24P、1197 和 UM-UC-3 从美国典型培养物保藏中心（ATCC）获得，并按照 ATCC 推荐的方法维持。

瞬时转染用磷酸钙转染方法进行。把沉淀物（7 μ g 质粒）加入到 30mm 平皿 0.3×10^6 个细胞的 1 ml 培养基中。16 小时后，移去转染培养基并加入新鲜的培养基。转染后 24-96 小时收获细胞，然后用丁酰

基-CoA 有机相抽提法 (Sambrook 等, 1989) 测定 CAT 活性。把上层有机相的等分试样 (100 μ l) 转移到含有 3 ml 闪烁液的孔中, 然后计数。

6.1.2 表达载体的构建方法

5 pCAT 基本质粒 (含有多个克隆位点之前的一个 CAT 报道基因)、pCAT 启动子质粒 (含有在 SV40 启动子控制下的 CAT 报道基因)、pCAT 增强子质粒 (含有位于 CAT 报道基因下游的 SV40 增强子) 和一个用于在 CAT 报道基因上游插入启动子的多克隆位点) 以及 pCAT 对照质粒 (含有 SV40 启动子和增强子控制下的 CAT 报道基因), 全部按商业途径从 Promega (Madison, WI) 购得。

为构建含有 H19 启动子控制下的 CAT 报道基因的质粒 pH19E, 首先把 H19 启动子区 (SEQ ID NO: 1) 克隆到 pBluescript II SK⁺ (Promega) 中。从人胎盘 DNA 用下列引物, 扩增含有 H19 启动子序列的多核苷酸。

15

ESPCR21: CGGTTCCCCACTTCCCCAGTTT (SEQ ID NO: 6)
ESPCR22: CGGAAGTCGACAACCCTCACCAAAGGCCAAGGT
(SEQ ID NO: 7)

20 用 Klenow 酶把该 PCR 产物的末端抹平并克隆到 pBluescript II SK⁺ 的 EcoRV 位点上。所插入的 DNA 用内切酶 PvuII、EcoRI 和 ApaI 消化证实。该启动子的方向与载体的 lacZ 编码区的方向相反。然后用 HindIII 和 PstI 切割法切下该启动子区, 所得到的大约 0.9 kb 的片段插入到 pCAT 基本质粒 HindIII-PstI 位点中以产生 pH19E。

25 按下列方法构建表达质粒, 其中该表达质粒含有在 H19 启动子/CAT 报道基因下游 2 个方向插入的 H19 增强子区。一条含有 H19 下游增强子 (相对于 H19 转录起点 + 6.0 kb ~ + 11 kb) 的 5 kb SacI 片段被克隆到 pUC19 的 SacI 位点。然后用 EcoRI 和 HindIII 切下该增强子片段, 并连接到 pBluescript II SK⁺ 的 EcoRI-HindIII 位点上而产生 pBhH19En-Sa。用 BamHI 部分消化 pBhH19En-Sa, 含有 H19 增强子 (和一个内部 BamHI 位点) 的该 5 kb 片段被克隆到在 pH19E 中
30 H19 启动子/CAT 报道基因的下流的 BamHI 位点上。产生了在正方向

(pH19EH19D) 和反方向 (pH19EH19R) 含有 H19 增强子的质粒。

6.2 结果和讨论

用 pCAT 基本质粒 (在图 2 命名为 P-E)、pCAT 对照质粒 (在图 2 命名为 pSV40ESV40)、pH19E、pH19EH19D 和 pH19EH19R 的每一种转染 5 种不同的膀胱癌细胞系 HT-1376、EJ28、T24P、1197 和 UM-UC-3。每种构建体的表达结果在图 3A-3E 表示。在每种细胞系中, 用含有 SV40 增强子和 SV40 启动子的 pCAT 对照质粒观察到最高水平的 CAT 活性。该构建体用作阳性对照, 因为已经确立了 SV40 调节序列作为基因表达的诱导子。然而 SV40 调节序列其诱导基因表达的能力并不是肿瘤细胞特异的。用含有在 H19 启动子的控制下的 CAT 报道基因的 pH19E 转染的细胞系, 也显示相对于背景明显增加的 CAT 的表达。由 H19 启动子诱导 CAT 活性的水平范围在从 HT-1376 细胞系中的 5 倍至在 UM-UC-3 细胞系中的 10 倍。把 H19 增强子加入到 H19 启动子/CAT 报道基因的构建体中进一步增加在某些细胞系中的表达水平。例如, 在细胞系 EJ28、T24P 和 1197 中, H19 增强子明显地增加从 H19 启动子/CAT 报道基因起始的表达。然而, 增强子的方向在不同细胞系中得到不同的结果。在细胞系 HT-1376 和 UM-UC-3 中, 该增强子对表达有微弱影响或没有影响。

这些结果表明人 H19 启动子区指导操作性连接的异源基因在各种类型的来源于膀胱癌的细胞系中表达。在一些来源于膀胱癌的细胞系中, 该 H19 增强子能进一步增加在 H19 控制下的报道基因的表达。

7. 实施例: 一种在 H19 调节序列控制下的毒素基因

7.1 材料和方法

修饰在上述 6 节中描述的表达构建体, 替代 CAT 以表达编码毒性产物或药物原的序列。例如, 用本领域众所周知的标准克隆方法, 移去编码 CAT 基因产物的序列, 并用编码单纯疱疹病毒胸腺嘧啶激酶 (HSV-TK) 的序列取代。

把 H19/药物原表述质粒转染到在 6 节描述的来源于膀胱癌的细胞系中。当转染到膀胱癌细胞系中时, 在 9-(1,3-二羟-2-丙氧甲基)鸟嘌呤的存在下, H19/HSV-TK 表达质粒诱导膀胱癌细胞特异的细胞毒性。

8. 实施例: 在化学诱导的膀胱癌的小鼠模型中的 H19 表达

8.1 材料和方法

以每笼 6 只小鼠饲养 70 只 5 周龄的雌性 C3H/He 小鼠 (Charles River), 并使小鼠适应 12 小时光照/12 小时黑暗周期的空调房间。在 8 周龄时, 开始实验, 把小鼠强制分成对照组 (10 只小鼠) 和实验组 (60 只小鼠)。随意喂给实验组小鼠溶在饮用水中的 0.05% N-丁基-N-(4-羟丁基)亚硝胺 (Tokyo Kasei Kogyo Co. Ltd. 东京, 日本)。对照组小鼠喂给自来水。实验开始后的第 4、8、12、16、20 和 26 周杀死来自两组的动物。用标准方法切下膀胱, 固定, 并包埋在石蜡块中。

8.1.1 探针的制备

含有小鼠 H19 编码区的 2.1 kb 的片段亚克隆到 pBluescript II KS 质粒 (Stratagene, La Jolla, CA) 中 T7 和 T3 RNA 聚合酶结合位点之后。用 T7 聚合酶 (Boehringer Mannheim) 和 Amersham RPN 2006 试剂盒, 从 HindIII 线性化的质粒 DNA 中, 体外产生 [³⁵S] 标记的反义 H19 RNA。体外产生的转录体具有 10⁷ cpm/μg 的比活。用 T3 聚合酶 (Boehringer Mannheim) 和 EcoRI 线性化模板制备的有义 H19 mRNA 被用作对照。

8.1.2 原位杂交

福尔马林固定组织的石蜡切片 (5μM) 被固定在 3-氨基丙基三乙氧基甲硅烷 (3-aminopropyltriethoxysilane) (Tespas, Sigma) 包被的显微镜用载玻片上, 然后 37℃ 干燥过夜。切片用二甲苯脱蜡, 用 4% 多聚甲醛固定, 然后用蛋白酶 K (Sigma) 处理。切片被乙酰化, 以降低探针的非特异结合, 并通过系列乙醇脱水。

按照 Rangini 等, 1991, Mech. Dev. 35: 13-24 描述的方法 (省略 thio-AMP 步骤), 用 [³⁵S] 标记的 RNA 探针 (比活 50,000 CPU/μl) 杂交。切片与显影膜接触 10 天, 然后用苏木精和伊红复染。在明暗视野照明下用 Polyvar (Reichert Jung) 显微镜检查载玻片并照相。对照组包括与正义 RNA 探针杂交及 RNA 酶预杂交处理。另外, 来自成年健康小鼠 (不表达 H19) 和胚胎小鼠膀胱 (表达 H19) 的膀胱切片分别用作阴性和阳性对照。

8.2 结果和讨论

到第 26 周, 所有存活的实验组小鼠出现可触摸的膀胱肿瘤。在化

学诱导的膀胱肿瘤中观察到 H19 的广泛表达。相反，在正常成年膀胱中，未检测到 H19 的表达。因此，该化学诱导膀胱癌的小鼠模型可用作一种动物模型，用来说明含有操作性与毒素基因相连接的 H19 调节区的构建体体内特异于肿瘤的细胞毒性。

5 9. 实施例：用 H19 调节序列以在膀胱癌小鼠模型中表达异源基因的基因治疗

为了体内传递到小鼠膀胱中，把 H19/毒素或前药表达质粒结合到脂质体中（如按照 Takashita 等，1993，临床研究杂志 93：652-651 所描述的方法，在此引入仅供参考）。用于本实验的小鼠是在上面第 8 节描述的患有化学的膀胱肿瘤。

简言之，将溶在 500 μ l 的 Optimen's 无血清培养基（BRL 生命技术公司，Gaithersburg, MD）的 50 μ g 质粒 DNA 加到 250 μ l Lipofectamine 及 250 μ l 水中。混合液在室温下温育 30 分钟，然后稀释在 10ml 的平衡盐溶液中（BSS（-）：140mM NaCl, 5.4mM KCl, 10mM Tris -HCl, pH 7.6）。把溶液在 15,000 rpm 离心 30 分钟沉淀后，该脂质体重新悬浮在 1ml 含有 1mM CaCl₂ 的 BSS（-）中。大约 0.2 ml 的浓缩脂质体经导管施用到载有化学诱导的膀胱肿瘤的小鼠中。带有膀胱肿瘤的对照组小鼠接受没有 DNA 的脂质体，或者含有在 H19 调节序列控制下的无关基因的构建体。在确定的时间点，杀死每组的小鼠，用标准方法切下膀胱，固定并包埋在石蜡块中。在原位杂交中用上面描述的 H19 探针或针对假单胞杆菌毒素基因的编码序列的探针处理另外的切片。此外，比较对照组和实验组之间肿瘤的大小、数量和坏死程度。发现假单胞菌毒素的表达与 H19 的表达共同定位在来自实验组小鼠的膀胱肿瘤内。此外，与对照组小鼠中的膀胱肿瘤相比较，实验组小鼠中的膀胱肿瘤大小减少并且坏死程度降低。

10. 实验例：在肿瘤细胞系中从 IGF-2 P3 和 P4 启动子起始的表达

10.1 材料和方法

在本实验中，构建各种含有置于 4 种不同的 IGF-2 启动子之一控制下的萤光素酶报道基因的表达式构建体，并转移到几种不同膀胱癌的细胞系中。制得了下列人 IGF-2 启动子/萤光脂酶构建体：

质粒构建体	IGF-2 基因核苷酸序列	启动子
Hup1	-980 to +54	启动子 1
Hup2	-379 to +271	启动子 2
Hup3	-1229 to +140	启动子 3
Hup4	-546 to +102	启动子 4

在 Sussenbuch 等, 1992, 生长调节 2: 1-9 中描述 IGF-2 的启动子序列, 在此引入以供参考。萤光素酶报道基因按商业途径从 Promega 公司, Madison, WI (目录号 #E1641) 购到。

5 10.2 结果和讨论

按上面第 6 节所描述的方法, 把得到的表达质粒转移到人膀胱癌细胞系 HT-1376、EJ28、T24P、1197 和 UM-UC-3 中。用试剂盒 (Promega 公司, Medison, WI, 目录号 E1500) 测定萤光素酶活性。在图 4A-4E 中显示的结果说明, IGF-2 P4 启动子指导在每种所测定的膀胱癌细胞系中萤光素酶报道基因的表达。在细胞系 1197 中, IGF-2 P3 启动子也指导萤光素酶报道基因的表达。在后续的实验

10 中, IGF-2 P3 和 P4 启动子被显示指导萤光素酶在其他肿瘤细胞系包括绒毛膜癌细胞和成横纹肌细胞瘤细胞中的表达。

11. 实施例: H19 启动子和 IGF-2 启动子与 H19 增强子一起作用促进异源基因的表达

15

11.1 材料和方法

4 种萤光素酶报道基因载体, pGL3-基本、pGL3-启动子、PGL3-增强子和 pGL3-对照从 Promega 公司获得。用多种不同的转染试剂和磷酸钙方法 (Gorman 等, 1982, 分子细胞生物学 2: 1044-1051), 把这些载体转染到培养细胞系中, 所用的转染试剂包括 lipofetamine (Gibco/BRL)、fugene (Boehringer)、8 种脂质试剂的完美转染试剂盒 (Invitrogen)、TFX-10、TFX-20、快速转染试剂盒 (Promega)。

20

被克隆到 pBluescript II SK 的 EcoRV 位点上的 H19 启动子 (pbh19p#1) 在第 6 节描述, 同上。通过用 SmaI 和 HindIII 切割切下 H19 启动子, 然后将所得到的 0.9 kb 的片段插入到 pGL3-基本载体的 SmaI-HindIII 位点上, 以产生 Luc-pbh19 构建体。

25

用引物 5'-ATATGGTACCGACAACCCTCACCAAAG-3' (上游, SEQ ID NO: 8) 和 5'-ATATAAGCTTCTTCTCCCTCACCTGCTC-3' (下游, SEQ ID NO: 9), 从 pbh19p#1 质粒经 PCR 法扩增来自 nt-819—+14 的 H19 启动子区。用 KpnI 和 HindIII 消化所得到的 PCR 产物, 然后连接到 pGL-3-基本载体的 KpnI-HindIII 位点上, 产生 Luc-PBH19 构建体。该 PCR 产生的 H19 启动子通过自动染料终止子循环测序法 (ABT Prism 377 DNA 测序仪, Perkin Elmer), 在 2 个方向测序。图 5 显示通过 PCR 产生的 H19 启动子的核苷酸序列 (SEQ ID NO: 2)。

按第 6 节描述的 (同上) 5 kb H19 下游增强子用 DamH 消化, 产生 2 条片段 4.1 kb 和 3'末端的 0.9 kb。把 H19 增强子的 0.9 kb 和 4.1 kb 的 BamHI 片段分别插入到 Luc-PBH19 质粒的 BamHI 位点上构建 Luc-PBH19-0.9EH19 和 Luc-PBH19-4EH19 构建体。该增强子序列定位在 H19 启动子/萤光素酶报道基因的下流。

将 0.9 kb BamHI 增强子片段连接至 pGL-基本载体的 BamHI 位点以产生 Luc-0.9EH19 载体。用 KpnI-BamHI 切下 pbh19p#1 质粒的 H19 启动子, 并连接至 Luc-0.9EH19 构建体的 KpnI-BglII 位点, 产生包含如节 6, 见上文所描述的启动子克隆和位于 H19Luc 报道基因下游的 0.9 kb 增强子的 Luc-pbh19-0.9EH19 表达载体。

用 Sussenbach 等, 1992, 生长调节 2: 1-9 中所描述的方法, 构建分别在人 IGF-2 启动子 P1、P2、P3 和 P4 控制下含有萤光素酶基因命名为 Hup-1、Hup-2、Hup-3 和 Hup-4 的表达载体。用引物 5'-ACAGGTACCTCTAGAGTCGACCT-3' (上游, SEQ ID NO: 10) 和 5'-ATATAAGCTTGCTCCCATCCTGCA-3' (下游, SEQ ID NO: 11), 从 Hup-4 构建体通过 PCR 法扩增 512 bp 的 P4 区。所得到的 PCR 产物用 KpnI-HindIII 消化, 然后连接到 pGL3-基本报道基因载体的 KpnI-HindIII 位点, 以产生 Luc-P4 报道基因的载体。

还制备了含有 IGF-2 P4 启动子和 H19 增强子的表达载体。一条来源于前面描述的 4.1 kb 片段的 2 kb BamHI 增强子片段被克隆到 Luc-P4 构建体的 BamHI 位点, 产生 Luc-P4-2EH19 表达载体。

用自动 DNA 测序法对 0.9 kb、2 kb 和 4.1 kb 的 H19 增强子测序。0.9 kb 增强子的核苷酸序列在图 6 显示 (SEQ ID NO: 3)。2 kb 增强

子的核苷酸序列在图 7A 和 7B 显示 (SEQ ID NO: 4)。4.1 kb 增强子的核苷酸序列在图 8A - 8C 显示 (SEQ ID NO: 5)。

11.2 结果和讨论

5 当用几种转染试剂把 4 种含萤光素酶基因的载体导入到培养的细胞系时, 对大部分所测定的细胞系来说, 磷酸钙沉淀法产生最高的转染效率。因此, 钙沉淀法随后被用来转染各种表达载体。此外, 增加质粒 DNA 的浓度不抑制转染效率, 甚至它们在平台期以上的浓度使用时也如此。

10 将膀胱癌细胞系 5637、肝细胞癌 (HCC) 细胞系 Huh7 和肾瘤细胞系 293T 用不同构建体的每一种转染, 所述的构建体含有和 H19 增强子结合在一起在 H19 或 IGF - 2 P4 启动子控制下的萤光素酶报道基因。

15 相对于背景, 用含有报道基因和 H19 启动子的 Luc - ph19 和 Luc - PH19 转染的细胞表现为增加基因表达 (图 9A - 9C)。在所测定的每种细胞系中, 含有 PCR 产生的启动子的构建体 Luc - PH19 比 Luc - ph19 表现为更高的活性。把 H19 0.9kb 的增强子片段添加到 Luc - ph19 报道载体 (Luc - ph19 - 0.9EH19) 上进一步在细胞系 5637 和 293T 中分别增加表达水平, 为 2 - 4 倍。

20 相对于背景, IGF - 2 P4 启动子也增加在所有的细胞系中萤光素酶的表达。把 2 kb H19 增强子片段添加到 Luc - P4 表达载体上增强 P4 启动子活性。由 2 kb 增强子片段诱导的萤光素酶活性水平的范围从 293T 细胞系的 2 倍至 Huh7 细胞系中的 6 倍, 而该增强子仅微弱地增强在 5637 细胞系中的启动子活性。

25 图 10A - 10E 显示含有 PCR 产生的 H19 启动子和 4.1 kb H19 增强子片段的构建体 Luc - ph19 - 4EH19 的表达。除了在 5637 细胞系外, 该增强子在细胞系中将启动子活性增强 3 - 28 倍。

12. 克隆的保藏

30 根据用于专利程序的微生物保藏的国际承认的布达佩斯条约的有关规定, 将下述质粒保藏于美国典型培养物保藏中心 (ATCC)、Manassas, VA。

克隆	ATCC 登记号	保藏日期
pH19EH19	209322	1997 年 10 月 2 日

5 等值方案

前面描述的具体内容足以使本领域的技术人员实施本发明。的确，对分子生物学、医学或相关领域的技术人员来说显而易见的在实施本发明时作出上述方案的各种修改方案被认为在下列权利要求的范围之内。

10 在此引用的所有出版物其全文被引出参考。

序列表

<110> 耶路撒冷希伯莱大学伊森姆研究开发公司

<120> 诱导对肿瘤特异的细胞毒性的方法和组合物

<130> 9457-0014-228

<140> PCT/IL98/00486

<141> 1998-10-04

<150> US 09/165,240

<151> 1998-10-01

<150> US 08/943,608

<151> 1997-10-03

<160> 11

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 830

<212> DNA

<213> 人类

<400> 1

ctgcagggcc	ccaacaaccc	tcaccaaagg	ccaaggtggt	gaccgacgga	cccacagcgg	60
ggctggctggg	ggagtcgaaa	ctcgccagtc	tccactccac	tcccaaccgt	ggtgccccac	120
gcggggcctgg	gagagtcctgt	gaggccgccc	accgcttgct	agtagagtgc	gcccgcgagc	180
cgtaagcaca	gcccggcaac	atgcggtctt	cagacaggaa	agtggccgcg	aatgggaccg	240
gggtgcccag	cggtctgtgg	gactctgtcc	tgcggaaacc	gcggtgacga	gcacaagctc	300
ggtcaactgg	atgggaatcg	gcctgggggg	ctggcaccgc	gcccaccagg	gggtttgcgg	360
cacttccctc	tgccctctcg	caccccaccc	ctactctcca	ggaacgtgag	gtctgagccg	420
tgatgggtgg	aggaaggggc	cctctgtgct	atccgagtc	ccagggaacc	gcagctggcc	480
cccagccatg	tgcaaagtat	gtgcagggcg	ctggcaggca	gggagcagca	ggcatgggtg	540
cccttgaggg	gagacagtgg	tctgggagg	agaggtcctg	gacctgagg	gaggtgatgg	600
ggcaatgctc	agccctgtct	ccggatgcca	aaggaggggt	gcggggaggc	cytctttgga	660
gaattccagg	atgggtgctg	ggtgagagag	acgtgtgctg	gaactgtcca	gggcggagggt	720
gggccctgcg	ggggccctcg	ggaggsgcct	gctctgattg	gccggcaggg	caggggcggg	780
aattctgsgc	ggccacccca	gttagaaaa	gcccgggcta	ggaccgagga		830

<210> 2

<211> 833

<212> DNA

<213> 人类

<400> 2

gacaaccctc	accaaggggc	aaggtggtga	ccgacggacc	cacagcgggg	tggctggggg	60
agtcgaaact	cgccagtcct	cactccactc	ccaaccgtgg	tgccccacgc	gggcctggga	120
gagtcgtgga	ggccgcccac	cgcttgctag	tagagtgcgc	ccgcgagccg	taagcacagc	180
ccggcaacat	gcggtcttca	gacaggaaa	tggccgcgaa	tgggaccggg	gtgcccagcg	240

gctgtgggga	ctctgtccctg	cggaaccg	ggtgacgagc	acaagctcgg	tcaactggat	300
gggaatcggc	ctssggsgct	ggcaccgsgc	ccaccagggg	gtttgcgga	ctccctctg	360
ccccctagca	ccccaccctt	actctccagg	aacgtgagtt	ctgagccgtg	atggaggcag	420
gaaggsgccc	tctgtgccat	ccgagteccc	agggaccg	agctggcccc	cagccatgtg	480
caaagtatgt	gcagsgcgct	ggcaggcagg	gagcagcagg	catgggtgtc	cctgagggga	540
gacagtgggc	tgggagggag	aagtcctggc	cctgagggag	gtgatggggc	aatgctcagc	600
cctgtctccg	gatgccaag	gaggggtg	gggagggcgt	ctttggagaa	ttccaggatg	660
ggtgctgggt	gagagagacg	tgtgctggaa	ctgtccagg	cgaggtggg	ccctgcgggg	720
gccctcggga	ggggccctgct	ctgattggcc	ggcagggcag	gggagggaat	tctggcgggg	780
gccaccaccag	ttagaaaaag	ccggsgctag	gaccgaggag	cagsgtgagg	gag	833

<210> 3

<211> 877

<212> DNA

<213> 人类

<400> 3

caaggacatg	gaatttcgga	ccttctgtcc	ccaccctctc	tgtgagcct	aggaaacctct	60
gagcagcagg	aaggccttgg	gtctagagcc	tagaaatgga	ccccacgtc	cacctgccca	120
gcctagaccc	ccagcattga	aggggtggta	gacttctgtg	gagaggaagc	cactaagcgg	180
gatggacacc	atcgccact	ccaccggcc	ctggccagcc	ctggccagtc	cagcccagtc	240
cagcccagcc	ctgcccctcc	cagcccgtcc	cagcccagct	catccctgcc	ctaccagcc	300
cagccctgtc	ctgcccctgcc	cagcccagcc	cagcccagcc	ctgcccctgcc	ctgcccctgcc	360
cttcccagcc	ctgaccttcc	cagcccgtcc	cagcccagct	catccctgcc	ctaccagct	420
cagcccgtcc	ctgcccctgcc	ctgcccgtcc	cagcccagcc	cagcccagcc	ctgcccgtcc	480
ctgcccagct	cagcccgtcc	caccccagcc	cagcccagcc	cagcatgctg	tctctggatg	540
gtgagcacag	gcttgacctt	agaaagaggc	tggcaacgag	ggctgaggcc	accaggccac	600
tgggtgtctc	cgggtcagac	aagcccagag	cctgtctccc	tggcacgggt	cggggctgtc	660
accggcagca	tgtgtggat	gtgcatggcc	tcagggtctg	tggctccagg	ctgccccgc	720
cctggctccc	gagggcacc	ctcttatg	atgaaccttg	tggcacacc	acctctgagc	780
gtccccctgt	cctgcccgt	gcacccctg	agcagccccc	tgtgtgtttc	atgggagttc	840
tagcaaggaa	ggggagctcg	aattcctgca	gcccggg			877

<210> 4

<211> 1960

<212> DNA

<213> 人类

<400> 4

ccgggtaccg	agctcccagg	aagataaatg	atttcctcct	ctctagagat	gggggtggga	60
tctgagcact	cagagccaag	ggcgagtg	gtccggggcg	gggccctcct	cggccctccc	120
aacatggggg	ccaggaggtc	agccctcaa	cctggacccc	ggctgggtct	cagggaatgg	180
tctccccag	tggccagct	tgttgtgtt	ttcagatggg	tgtgcatggg	tgtgtgtgtg	240
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgatgcct	gacaagcccc	agagagccaa	300
agacctgagt	ggagatcttg	tgaacttctc	aaagggggat	tggaaagtct	gagaaagagc	360
tgtggtcagc	cttgcctctc	cttaaggctg	tggtaaccac	actaggcata	gcataggcct	420
gcgcccctgc	cctccttccc	tctccggc	ctctccttct	tcttctctcc	ccctctaccc	480
cgtccctgg	cctgctcctg	gtgacaccgt	tggccccctt	ccagggtctga	gggaagccag	540
cgggggcccc	ttcctgaaa	cccacctgca	ggccggcttg	ctgggaagg	gctgctctcg	600
cagaggctcc	cggccgccc	gcagccgttt	cctggaagca	gtcgtgtgtg	gtattctgtt	660
cctgttcagc	actgtgcttg	caaagaaa	agacactgtg	ctccttgtcc	ttaggagacc	720
ccgtccatc	acccaacacc	tggctggaca	caggcgggag	gccgggtccg	cggggagcgg	780
cgggggctg	gggcccggacc	attaaaacac	cacggggcgc	aggcactgca	ggctcctcct	840
cctcctcctg	cccagcgct	ctgctcacag	gcacgtgcca	agcccttagg	ccaggagggc	900
agcagtgggt	gcagaacaag	ctcctgggaa	gggggtgag	ggcggaaccc	cggggagaa	960
ggctggcagg	gctgtggggg	acgctgaccg	tggccccac	gttgagaaa	actgntgccc	1020

tggctggaag	atgggggaga	tgccaagcct	ctgaggcagc	acgagcaggg	tgcatggagg	1080
ccggggcgcg	gggagggctgc	actgcagcat	gcaccccaaa	gcccanaagg	agtggagacc	1140
aggccctgga	atcgagaagt	agaaaggcgg	cttggaggcc	tcggaaccgg	ctgacctcca	1200
acagagtggg	tctccagcct	ggctctgccc	tgccgcaggt	ccccccccct	cattaccagg	1260
cctagagcct	ccagtcccgg	tggccccag	cccaggggtg	aacggcctca	ccctgggtcg	1320
tgggacagag	ggcacgttca	tcaagagtgg	ctcccaagg	acacgtggct	gtttgcagtt	1380
cacaggaagc	attcgagata	aggagcttgt	tttcccagtg	ggcacggagc	cagcaggggg	1440
gctgtggggc	agcccagggt	gcaaggccag	gctgtggggc	tgcaagctgc	ttggggccca	1500
ctcccaggcc	tttgcgggag	gtgggaggcg	ggaggcgsga	gctgcacagt	ggccccaggc	1560
gaggctctca	gccccagtcg	ctctccgggt	gggcagccca	agagggtctg	gctgagcctc	1620
ccacatctgg	gactccatca	cccaacaact	taattaaggc	tgaatttcac	gtgtcctgtg	1680
acttgggtag	acaaagcccc	tgtccaaagg	ggcagccagc	ctaaggcagt	ggggacggcg	1740
tgggtggcgg	gcgacggggg	agatggacaa	caggaccgag	ggtgtgcggg	cgatggggga	1800
gatggacaac	aggaccgagg	gtgtgcgggc	gatgggggag	atggacaaca	ggaccgaggg	1860
tgtgcgggac	acgcattgtca	ctcatgcacg	ccaatggggg	gcgtgggagg	ctggggagca	1920
gacagactgg	gctgggctgg	gcgggaagga	cgggcagatg			1960

<210> 5

<211> 4085

<212> DNA

<213> 人类

<400> 5

ccgggtaccg	agctcccagg	aagataaatg	atttctctct	ctctagagat	gggggtggga	60
tctgagcaat	cagagccaaag	ggcgagtggt	gtccggggcg	gggccctcct	cggccctccc	120
aacatggggg	ccaggagggtc	agccccctca	cctggacccc	ggctgggtct	cagggaaatgg	180
tctccccag	tggcccagct	tgtttgtgtt	ttcagatggg	tgtgcatggg	tgtgtgtgtg	240
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgatgcct	gacaagcccc	agagagccaa	300
agacctgagt	ggagatcttg	tgactttctca	aaagggggat	tggaaaggttc	gagaaagagc	360
tgtggtcagc	cttgctctcc	cttaaggctg	tggtaaccac	actaggcata	gcataggcct	420
gcgccccgtc	cctccttccc	tcctccgcgc	ctctcctttc	tctttctccc	ccctctaccc	480
cgctccctgg	cctgctcctg	gtgacaccgt	tggccccctt	ccagggtctga	gggaagccag	540
cggggggccc	ttcctgaaaag	cccacctgca	ggccggcttg	ctgggaaggg	gctgctctcg	600
cagaggctcc	cgcgcgcctt	gcagccgttt	cctggaaagca	gtcgctgtgg	gtattctgtt	660
ccttgcagc	actgtgcttg	caaagaaagc	agacactgtg	ctccttctcc	ttagggagcc	720
ccgctccatc	acccaacacc	tggctggaca	caggcgggag	gccgggtccg	cgggggagcg	780
cgcggggctg	gggcgggacc	attaaacaca	gcagggcgcc	aggcactgca	ggctcctcct	840
cctcctcctg	cccagcgccct	ctgctcacag	gcacgtgcca	agccccctag	ccaggaggcc	900
agcagtggtg	gcagaacaag	ctcctgggaa	gggggtgcag	ggcggaaccc	cgggggagaag	960
ggctggcagg	gctgtggggg	acgctgaccg	tgggccccac	gttgacagaaa	actggntgcc	1020
tggctggaag	atgggggaga	tgccaagcct	ctgaggcagc	acgagcaggg	tgcatggagg	1080
ccggggcgcg	gggagggtgc	actgcagcat	gcaccccaaa	gcccanaagg	agtggagacc	1140
aggccctgga	atcgagaagt	agaaaggcgg	cttggaggcc	tcggaaccgg	ctgacctcca	1200
acagagtggg	gcccggccctg	gaggcaaaag	ggtgcccggg	gtccggccct	gcctggggga	1260
gctatgtgtc	atgggcaagc	cacaggatat	gtagcccgct	ctgagcctat	ggaccagggg	1320
cagggtctca	aggcagggca	ggggagacag	cacgggggag	caaggagcag	agagggggcc	1380
tcaggctctc	ccaggaggaa	cattctcccc	acaggaggaa	gagacggccc	aggggtgact	1440
gtggggagcc	atggtggcag	ctggggctcg	ggcagatggg	agagaggctg	gagagggtga	1500
ggtgcagggg	tcagggtctc	ggggcccaca	tgctgtgggg	agcaggcagg	cccagggtct	1560
tccgccactc	cccactcccc	cttggctcat	aggctggggc	caagggtggg	gtgggatatg	1620
caggagatgg	ggccaggggg	gcaagcaggg	ccccaaagac	atttagaaaa	accgggttat	1680
gcaggcagca	ttcagagcag	gcggcggtgc	tggcgggggc	cctgggagca	cagagaggca	1740
cacgtagggc	ccccgagggg	ctccccattg	gcccgcagtg	acatcacccc	tgtgtcaaca	1800
gtgatgtctg	cagctccggc	cagccagggt	ttatggagcg	agaccagcc	cggcctgggc	1860
cctcactccc	caggcccaca	cactagccca	ctgttcaggg	tccgggggtg	cggcatggcc	1920
tgggggtcct	ggcaccgctg	ctcctctgcc	caccctaact	tcccggtatc	gcggctgccc	1980

```

cctctgagcg tccccaaacca gtaagtgtgg ggcccagcag gcctgccgtc ctccctctct 2040
tccctctctag agagaaacgt ggaggtcctg gggctggggg cgctcatagc cctgtgacac 2100
aggtgcatgg ggtcaggggt cccagaatgg cccctgggaa ggacctcagc tgggcccggcg 2160
gctctaggct tcaggggtct gtctgcacag gggntagccc ctcccagacc tctgtgaagc 2220
cagtacgggc ctccctctcc tgcctcgctg tctgtccggg gcttcctgga ctgcactgcg 2280
ggccactggt gagagsgtgg acaggggaagg gccgcccgtg tgccctgttcc tgcccacctg 2340
gctgtgtggt cccctccaag tagggacaac ccttctgagg gcttgggggc accctggggg 2400
tgccaggggc tcccagagcc ctgtgagccc ctgggggggtc tggcctgatg cccccctcca 2460
cgtccagggc cggctgtgsc ccagaacccc agcttcccag caggccgggtg tgcgggtggtg 2520
acccaggaga ggcttcgccc cactgaggg gccaccgacc tctgtcagac cacagagacc 2580
cccaaggagt ctgaaggctg gagaccggg gctgggacca ggtgggactt tcccacggag 2640
cgtccccag gccagctgg ggacacgtcc ccttctctc cagacacacc ctgcctgcca 2700
ccaggacaca cgggctgtt ggsggtctct ttttaagtgt tgccactctg aggtgactgt 2760
ccctttccaa agaggtttct ggsgccaggg tgggatgctg cggcctgagc agggagatct 2820
gggcccgcag gggctgggga ctgtctcctg gggaaaggaa cgcctgggag cgtgtgtgct 2880
gaccagggac catccaggga gggcctctg tggggcaagc gggaaaggag cggctggaga 2940
ggcttggcgg ccccgccct gcctcccatt ccttagctcc atgcctgtca acctctgtca 3000
cccagtgagt gatgtccagg ggcctggaa aggtcacagc atgtttgagc ggggtgagag 3060
agaggggaaa gggggggcg gggaaaagta cgtggaggaa gctttaggcc caagggaagga 3120
gacaggggtc tgggagggag ggagccactg gggccgcccg gaaggtccct gcttgctgct 3180
gccaccaga accctcgccc cttagctagc ccccgagcc ccagcctttc tggcntgtgg 3240
ccctctcccc catccccagg tgtcctgtgc aaccaggcct tggacccaaa cctcctgcc 3300
ccctcctctc cctcctcacc ctcccaatgc agtggctctc agcctggctc tgccctgccg 3360
caggtccccc cccctcatta ccaggccctag agcctccagt cccggtggcc cccagcccga 3420
gggtgaacgg cctcaccctg ggtcgtggga cagagggcac gttcatcaag agtggctccc 3480
aagggacacg tggctgtttg cagttcacag gaagcattcg agataaggag cttgttttcc 3540
cagtgggcac ggagccagca gggggsgctgt ggggcagccc agggtgcaag gccagsgctgt 3600
ggggctgcag ctgcccggg ccccaactccc aggcctttgc gggaggtggg aggcgggagg 3660
cggcagctgc acagtggccc caggcgaggc tctcagcccc agtcgctctc cgggtgggca 3720
gcccagagg gtctggctga gcctcccaca tctgggactc catcacccaa caacttaatt 3780
aaggctgaat ttacgtgtc ctgtgacttg ggtagacaaa gcccctgtcc aaaggggcag 3840
ccagcctaag gcagtgggga cggcgtgggt ggcgggcgac gggggagatg gacaacagga 3900
ccgagsgtgt scggscgatg ggggagatgg acaacaggac cgagsgtgtg cgggcgatgg 3960
gggagatgga caacaggacc gagsgtgtgc gggacacgca tgtcactcat gcacgccaat 4020
ggggggcggtg sgaggctggg gagcagacag actgggctcg gctgggcggg aaggacgggc 4080
agatg 4085

```

<210> 6

<211> 22

<212> DNA

<213> 人类

<400> 6

cggttcccca cttccccagt tt

22

<210> 7

<211> 33

<212> DNA

<213> 人类

<400> 7

cggaagtcga caaccctcac caaaggccaa ggt

33

<210> 8

<211> 27

<212> DNA

<213> 人类		
<400> 8		
atatggtacc gacaaccctc accaaag		27
<210> 9		
<211> 29		
<212> DNA		
<213> 人类		
<400> 9		
atataagctt cttctccctc accctgctc		29
<210> 10		
<211> 23		
<212> DNA		
<213> 人类		
<400> 10		
acaggtacct ctagagtcga cct		23
<210> 11		
<211> 24		
<212> DNA		
<213> 人类		
<400> 11		
atataagctt gctcccatcc tgca		24

CTGCAGGGCCCCAACAACCTCACCAAAGGCCAAGGTGGTGACCGACGGACCCACAGCGG
1 -----+-----+-----+-----+-----+ 60
GACGTCCCGGGGTGTTGGGAGTGGTTTCCGGTTCACCACTGGCTGCCTGGGTGTCGCC
|
-837

GGTGGCTGGGGGAGTCGAACTCGCCAGTCTCCACTCCACTCCCAACCGTGGTGCCCCAC
61 -----+-----+-----+-----+-----+ 120
CCACCGACCCCTCAGCTTTGAGCGGTGAGAGGTGAGGTGAGGGTTGGCACCACGGGGTG

GCGGGCTGGGAGAGTCTGTGAGGCCGCCACCGCTTGTCAGTAGAGTGCGCCGCGAGC
121 -----+-----+-----+-----+-----+ 180
CGCCCGGACCCTCTCAGACACTCCGGCGGGTGGCGAACAGTCATCTACCGGGGCGCTCG

CGTAAGCACAGCCCGGCAACATGCGGTCTTCAGACAGGAAAGTGGCCGCGAATGGGACCG
181 -----+-----+-----+-----+-----+ 240
GCATTCGTGTCGGGCCGTTGTACGCCAGAAAGTCTGTCTTTCACCGGCGCTTACCCTGGC

图 1A

```

GGGTGCCCAGCGGCTGTGGGACTCTGTCTCGGAAACCGCGGTGACGAGCACAAGCTC
241 -----+-----+-----+-----+-----+-----+-----+ 300
CCCACGGGTGCGCGACACCCCTGAGACAGGACGCCCTTGGCGCCACTGCTCGTGTTCGAG

GGTCAACTGGATGGGAATCGGCCTGGGGGCTGGCACCGCGCCACCAGGGGGTTTGGCG
301 -----+-----+-----+-----+-----+-----+-----+ 360
CCAGTTGACCTACCCCTTAGCCGGACCCCCGACCGTGGCGCGGGTGGTCCCCAAACGCC

CACTTCCCTCTGCCCCCTCAGCACCCACCCCTACTCTCCAGGAACGTGAGGTCTGAGCCG
361 -----+-----+-----+-----+-----+-----+-----+ 420
GTGAAGGGAGACGGGGAGTCGTGGGTGGGGATGAGAGGTCTTGCACTCCAGACTCGGC

TGATGGTGGCAGGAAGGGGCCCTCTGTGCCATCCGAGTCCCCAGGGACCCGAGCTGGCC
421 -----+-----+-----+-----+-----+-----+-----+ 480
ACTACCACCGTCCCTCCCCGGGAGACACGGTAGGCTCAGGGGTCCCTGGGCGTCGACCGG

CCCAGCCATGTGCAAAGTATGTGCAGGGCGCTGGCAGGCAGGGAGCAGCAGGCATGGTGT
481 -----+-----+-----+-----+-----+-----+-----+ 540
GGGTGGGTACACGTTTCATACACGTCCCGGACCGTCCGTCCCTCGTCGTCCGTACCACA

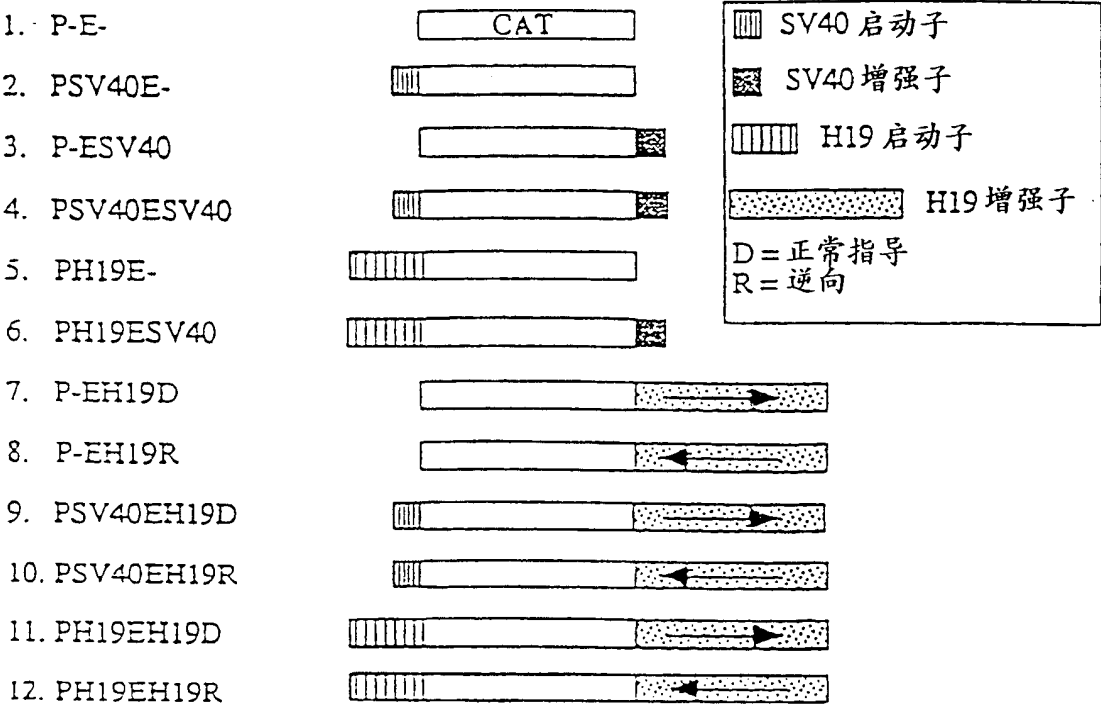
CCCCTGAGGGGAGACAGTGGTCTGGGAGGGAGAGGTCTGGACCCCTGAGGGAGGTGATGG
541 -----+-----+-----+-----+-----+-----+-----+ 600
GGGGACTCCCCCTCTGTCAACAGACCCCTCCCTCTCCAGGACCTGGGACTCCCTCCACTACC

```

图 1B



图 1C



A 概括性表示含有不同启动子和增强子的质粒

图 2

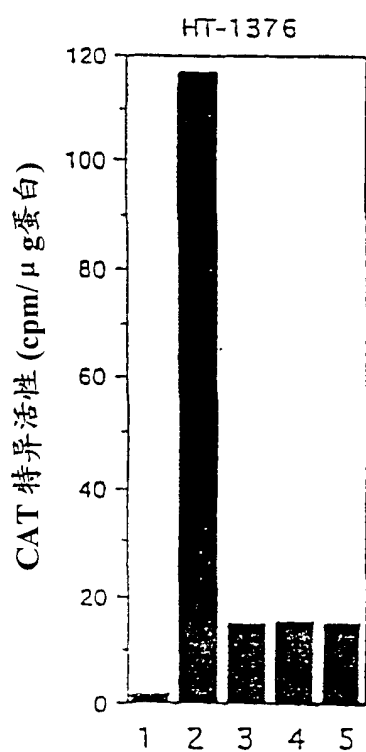


图 3A

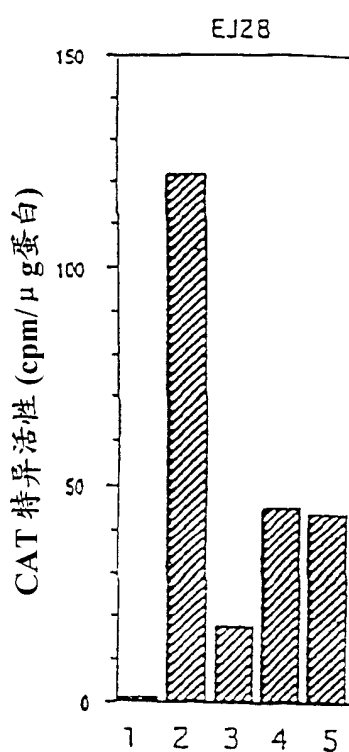


图 3B

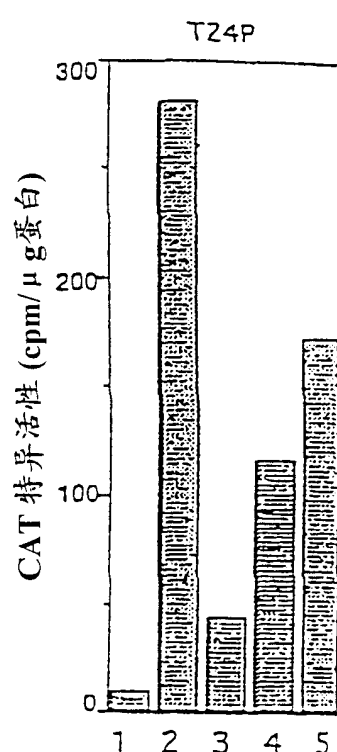


图 3C

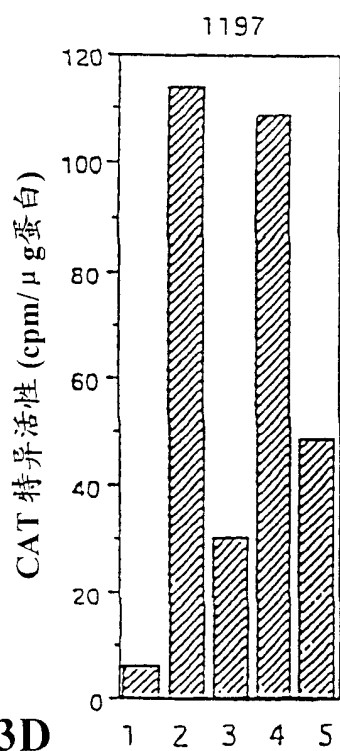


图 3D

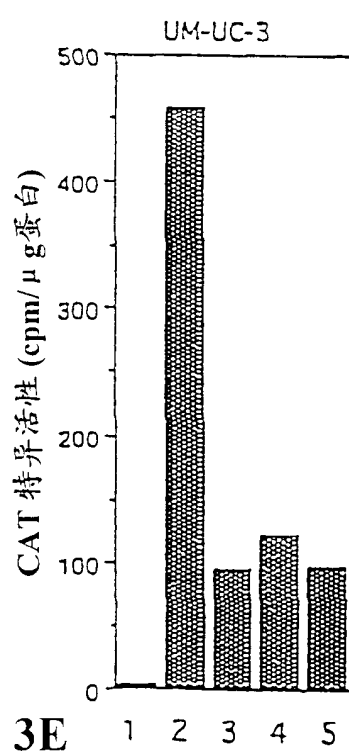
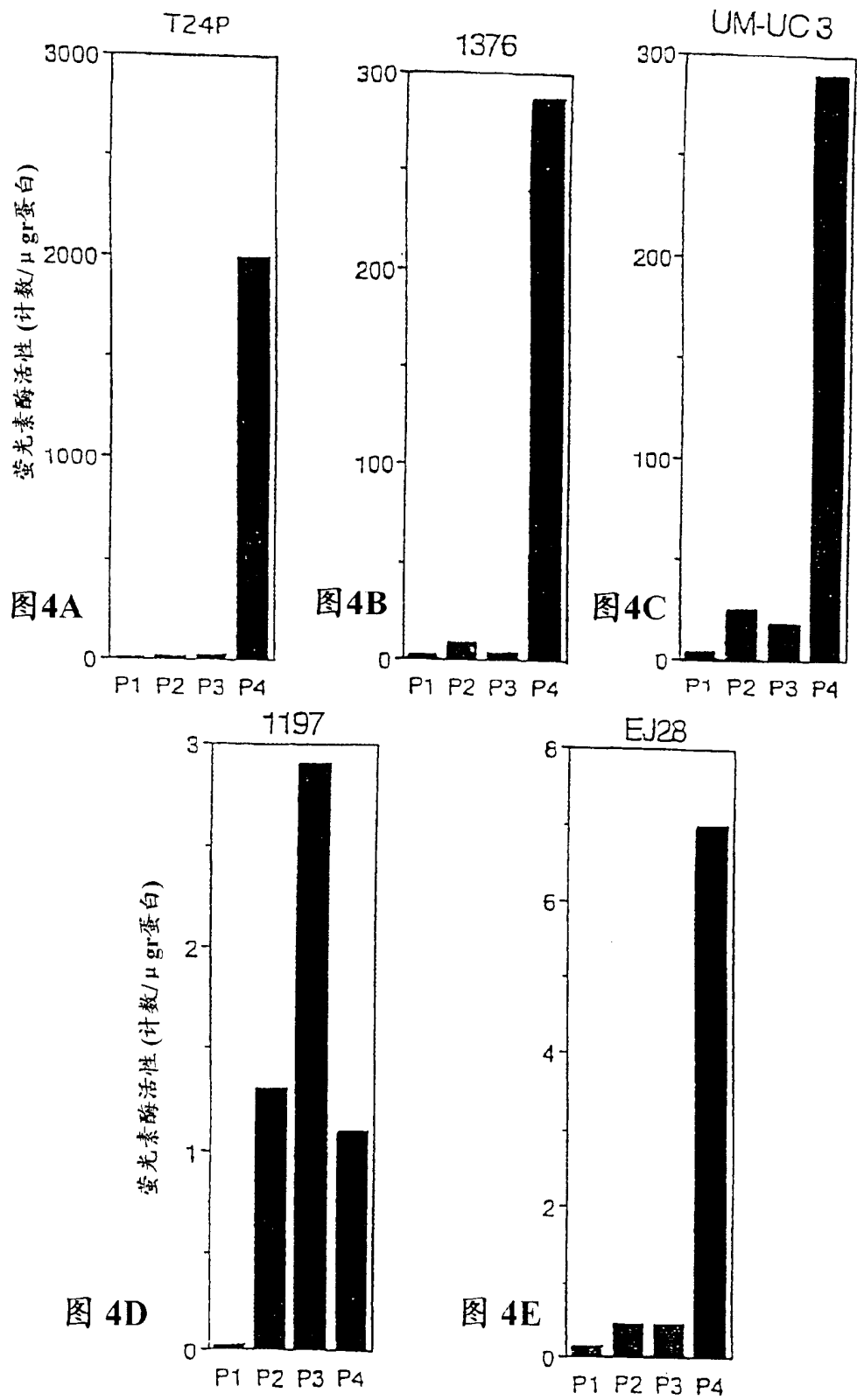


图 3E



1 GACAACCCTC ACCAAGGGCC AAGGTGGTGA CCGACGGACC CACAGCGGGG
51 TGGCTGGGGG AGTCGAAACT CGCCAGTCTC CACTCCACTC CCAACCGTGG
101 TGCCCCACGC GGGCCTGGGA GAGTCTGTGA GGCCGCCAC CGCTTGTCAG
151 TAGAGTGCGC CCGCGAGCCG TAAGCACAGC CCGGCAACAT GCGGTCTTCA
201 GACAGGAAAG TGGCCGCGAA TGGGACCGGG GTGCCAGCG GCTGTGGGA
251 CTCTGTCCTG CGGAAACCGC GGTGACGAGC ACAAGCTCGG TCAACTGGAT
301 GGGAATCGGC CTGGGGGGCT GGCACCGCGC CCACCAGGGG GTTTGCGGCA
351 CTTCCCTCTG CCCCTCAGCA CCCACCCCT ACTCTCCAGG AACGTGAGTT
401 CTGAGCCGTG ATGGTGGCAG GAAGGGGCCC TCTGTGCCAT CCGAGTCCCC
451 AGGGACCCGC AGCTGGCCCC CAGCCATGTG CAAAGTATGT GCAGGGCGCT
501 GGCAGGCAGG GAGCAGCAGG CATGGTGTCC CCTGAGGGGA GACAGTGGTC
551 TGGGAGGGAG AAGTCCTGGC CCTGAGGGAG GTGATGGGGC AATGCTCAGC
601 CCTGTCTCCG GATGCCAAAG GAGGGGTGCG GGGAGGCCGT CTTTGGAGAA
651 TTCCAGGATG GGTGCTGGGT GAGAGAGACG TGTGCTGGAA CTGTCCAGGG
701 CGGAGGTGGG CCCTGCGGGG GCCCTCGGGA GGGCCCTGCT CTGATTGGCC
751 GGCAGGGCAG GGGCGGGAAT TCTGGGCGGG GCCACCCAG TTAGAAAAAG
801 CCCGGGCTAG GACCGAGGAG CAGGGTGAGG GAG

图 5

1 CAAGGACATG GAATTTGGA CTTCTGTCC CCACCCTCTC TGCTGAGCCT
51 AGGAACCTCT GAGCAGCAGG AAGGCCTTGG GTCTAGAGCC TAGAAATGGA
101 CCCCCACGTC CACCTGCCCA GCCTAGACCC CCAGCATTGA AGGGTGGTCA
151 GACTTCCTGT GAGAGGAAGC CACTAAGCGG GATGGACACC ATCGCCCACT
201 CCACCCGGCC CTGCCCAGCC CTGCCCAGTC CAGCCCAGTC CAGCCCAGCC
251 CTGCCCTTCC CAGCCCTGCC CAGCCCAGCT CATCCCTGCC CTACCCAGCC
301 CAGCCCTGTC CTGCCCTGCC CAGCCCAGCC CAGCCCAGCC CTGCCCTGCC
351 CTGCCCTGCC CTTCCCAGCC CTGACCTTCC CAGCCCTGCC CAGCCCAGCT
401 CATCCCTGCC CTACCCAGCT CAGCCCTGCC CTGCCCTGCC CTGCCCTGCC
451 CAGCCCTACC CAGCCCAGCC CTGCCCTGCC CTGCCAGCT CAGCCCTGCC
501 CACCCCAGCC CAGCCCAGCC CAGCATGCGT TCTCTGGATG GTGAGCACAG
551 GCTTGACCTT AGAAAGAGGC TGGCAACGAG GGCTGAGGCC ACCAGGCCAC
601 TGGGTGCTCA CGGGTCAGAC AAGCCCAGAG CCTGCTCCCC TGCCACGGGT
651 CGGGGCTGTC ACCGCCAGCA TGCTGTGGAT GTGCATGGCC TCAGGGCTGC
701 TGGCTCCAGG CTGCCCCCGC CCTGGCTCCC GAGGCCACCC CTCTTATGCC
751 ATGAACCCTG TGCCACACCC ACCTCTGAGC TGTCCCCGCT CCTGCCGCCT
801 GCACCCCTG AGCAGCCCCC TGTGTGTTTC ATGGGAGTCT TAGCAAGGAA
851 GGGGAGCTCG AATTCCTGCA GCCCGG

图 6

1 CCGGGTACCG AGCTCCCAGG AAGATAAATG ATTTCTCCT CTCTAGAGAT
51 GGGGGTGGGA TCTGAGCACT CAGAGCCAAG GGCAGTGG GTCCGGGCGG
101 GGGCCCTCCT CGGCCCTCCC AACATGGGGG CCAGGAGGTC AGCCCTCAA
151 CCTGGACCCC GGCTGGGTCT CAGGGAATGG TCTCCCCAG TGGCCAGCT
201 TGCTTGTGTT TTCAGATGGG TGTGCATGGG TGTGTGTGTG TGTGTGTGTG
251 TGTGTGTGTG TGTGTGTGTG TGTGATGCCT GACAAGCCCC AGAGAGCCAA
301 AGACCTGAGT GGAgATCTTG TGA~~CT~~CTCA AAAGGGGGAT TGAAGGTTC
351 GAGAAAGAGC TGTGGTCAGC CTTGCTCTCC CTTAAGGCTG TGGTAACCAC
401 ACTAGGCATA GCATAGGCCT GCGCCCCGTC CCTCCTTCCC TCCTCCGCGC
451 CTCTCCTTTC TCTTTCTCCC CCCTCTACCC CGCTCCCTGG CCTGCTCCTG
501 GTGACACCGT TGGCCCCCTT CCAGGGCTGA GGGAAGCCAG CGGGGGCCCC
551 TTCCTGAAAG CCCACCTGCA GGCCGGCTTG CTGGGAAGGG GCTGCTCTCG
601 CA~~g~~AGGCTCC CGCCCCCCT GCAGCCGTTT CCTG~~g~~AAGCA GTCGCTGTGG
651 GTATTCTGTT CTTGTCA~~g~~C ACTGTGCTTG CAA~~g~~AAAGC A~~g~~ACACTGTG
701 CTCCTTGTCC TTAGGGAGCC CCGCTCCATC ACCCAACACC TGGCTGGACA
751 CAGGCGGGAG GCCGGGTCCG CGGGGA~~g~~CGG CGCGGGGCTG GGGCCGGACC
801 ATTAAACACA CACGGGCGCC AGGCACTGCA GGCTCCTCCT CTCCTCCTG
851 CCCAGCGCCT CTGCTCACAG GCACGTGCCA A~~g~~CCCCTAGG CCA~~g~~AgGCC
901 A~~g~~CA~~g~~TGGGT GC~~g~~AACAAG CTCCTGGGAA GGGGGTGC~~g~~ GG~~c~~GGACCCC
951 CGGGGA~~g~~AAG GGCTGGCAGG GCTGTGGGGG ACGCTGACCG TGGGCCCCAC

图 7A

1001 GTTGCAgAAA ACTGGNTGCC TGgCTGGAAG ATGGGGGAGA TGCCAAGCCT
1051 CTGAGGCAGC ACGAgCAGGG TGCATGGAGG CCGGGGCGCG GGGAGGCTGC
1101 ACTGCAGCAT GCACCCCAAA GCCCANAGGG AGTGGAgACC AGGCCCTGGA
1151 ATCgAGAAGT AgAAAGGCGG CTTGGAGGCC TCGGAACCGG CTGACCTCCA
1201 ACAGAgTGGG TCTCCAGCCT GgCTCTGCCC TGCCGCAGGT CCCCTCCCcT
1251 CATTACCAGG CCTAGAGCCT CCAGTCCCGG TGGCCCCCAG CCcGAGGGTG
1301 AACGGCCTCA CCCTGGGTcG TGGGACAGAG GGCACGTTCA TCAAGAGTGG
1351 CTCCCAAGGG ACACGTGGCT GTTTGCAGTT CACAGGAAGC ATTcGAGATA
1401 AGGAGcTTGT TTTCCAGTG GGCACGGAGC CAGCAGGGGG GCTGTGGGGC
1451 AGCCCAGGGT GCAAGGcCAG GcTGTGGGGC TGCAGcTGCc TTGGGCCCCA
1501 CTCCCAGGCC TTTGCGGGAG GTGGGAGGCG gGAGGCGGCA GCTGCACAGT
1551 GGCCCCAGGC GAGGCTCTCA GCCCCAGTCG CTCTCCGGGT GGGCAGCCCA
1601 AGAGGGTCTG GCTGAGCCTC CCACATCTGG GACTCCATCA CCCAACAACT
1651 TAATTAAGGC TGAATTTAC GTGTCCTGTG ACTTGGGTAG ACAAAGCCCC
1701 TGTCCAAAGG GGCAGCCAGC CTAAGGCAGT GGGGACGGCG TGGGTGGCGG
1751 GCGACGGGGG AGATGGACAA CAGGACCGAG GGTGTGCGGG CGATGGGGGA
1801 GATGGACAAC AGGACCGAGG GTGTGCGGGC GATGGGGGAG ATGGACAACA
1851 GGACCGAGGG TGTGCGGGAC ACGCATGTCA CTCATGCACG CCAATGGGGG
1901 GCGTGGGAGG CTGGGGAGCA GACAGACTGG GCTGGGCTGG GCGGGAAGGA
1951 CGGGCAGATG

图 7B

```

1  CCGGGTACCG AGCTCCCAGG AAGATAAATG ATTTCTCCT CTCTAGAGAT
51  GGGGGTGGGA TCTGAGCACT CAGAGCCAAG GCGCAGTGG GTCCGGGCGG
101 GGGCCCTCCT CGGCCCTCCC AACATGGGGG CCAGGAGGTC AGCCCTCAA
151 CCTGGACCCC GGCTGGGTCT CAGGGAATGG TCTCCCCAG TGGCCAGCT
201 TGCTTGTGTT TTCAGATGGG TGTGCATGGG TGTGTGTGTG TGTGTGTGTG
251 TGTGTGTGTG TGTGTGTGTG TGTGATGCCT GACAAGCCCC AGAGAGCCAA
301 AGACCTGAGT GGAgATCTTG TGAgCTTCTCA AAAGGGGGAT TGGAAGGTTG
351 GAGAAAGAGC TGTGGTCAGC CTTGCTCTCC CTTAAGGCTG TGGAACCCAC
401 ACTAGGCATA GCATAGGCCT GCGCCCCGTC CCTCCTCCC TCCTCCGCGC
451 CTCTCCTTTC TCTTCTCCC CCTCTACCC CGCTCCCTGG CCTGCTCCTG
501 GTGACACCGT TGGCCCCCTT CCAGGGCTGA GGAAGCCAG CGGGGGCCCC
551 TTCCTGAAAG CCCACCTGCA GGCCGGCTTG CTGGGAAGGG GCTGCTCTCG
601 CAgAGGCTCC CGCCCGCCCT GCAGCCGTTT CgTGgAAGCA GTGCTGTGG
651 GTATTCTGTT CCTTGTGAGC ACTGTGCTTG CAAgAAAGC AgACACTGTG
701 CTCCTTGTC TTAGGGAGCC CCGCTCCATC ACCCAACACC TGGCTGGACA
751 CAGGCGGGAG GCCGGGTCCG CGGGGAgCGG CGCGGGGCTG GGGCCGGACC
801 ATTAAACACA CACGGGCGCC AGGCACTGCA GGCTCCTCCT CCTCCTCCTG
851 CCCAGCGCCT CTGCTCACAG GCACGTGCCA AgCCCTAGG CCAgAgGCC
901 AgCAgTGGGT GCAgAACAAG CTCCTGGGAA GGGGGTGCAg GGcGGACCCC
951 CGGGGAgAAG GGCTGGCAGG GCTGTGGGG ACGCTGACCG TGGCCCCAC
1001 GTTGCAgAAA ACTGGNTGCC TGgCTGG~AG ATGGGGGAGA TGCCAAGCCT
1051 CTGAGGCAGC ACGAgCAGGG TGCATGGAGG CCGGGGCGCG GGGAGGCTGC
1101 ACTGCAGCAT GCACCCCAA GCCANAGGG AGTGGAgACC AGGCCCTGGA
1151 ATCgAGAAGT AgAAAGGCGG CTTGGAGGCC TCGGAACCGG CTGACCTCCA
1201 ACAGAgTGGG GCCGGCCCTG GAGGCAAAGA GGTGCCCGGG GTCCGGCCCT
1251 GCCTGGGGGA GCTATGTGTC ATGGGCAAGC CACAGGATAT GTagCCCGCT
1301 CTGagCCTAT GGACCCagGG CAGGGCTGCA AGGCAGGGCA GGGGAGACAG
1351 CACGGGGGAG CAAGGAGCAG AGAGGGGGCC TCAGGCTCTC CCAGGAGGAA
1401 CATTCTCCCG ACAGGAGGAA GAGACGGCCC AGGGGTGACT GTGGGGAGCC
1451 ATGGTGGCAG CTGGGGTCGT GGCAGATGG AGAGAGGCTG GCGAGGTGAA
1501 GGTGCAGGGG TCAGGGCTCT GGGGCCACA TGCTGTGGG AGCAGGCAGG

```

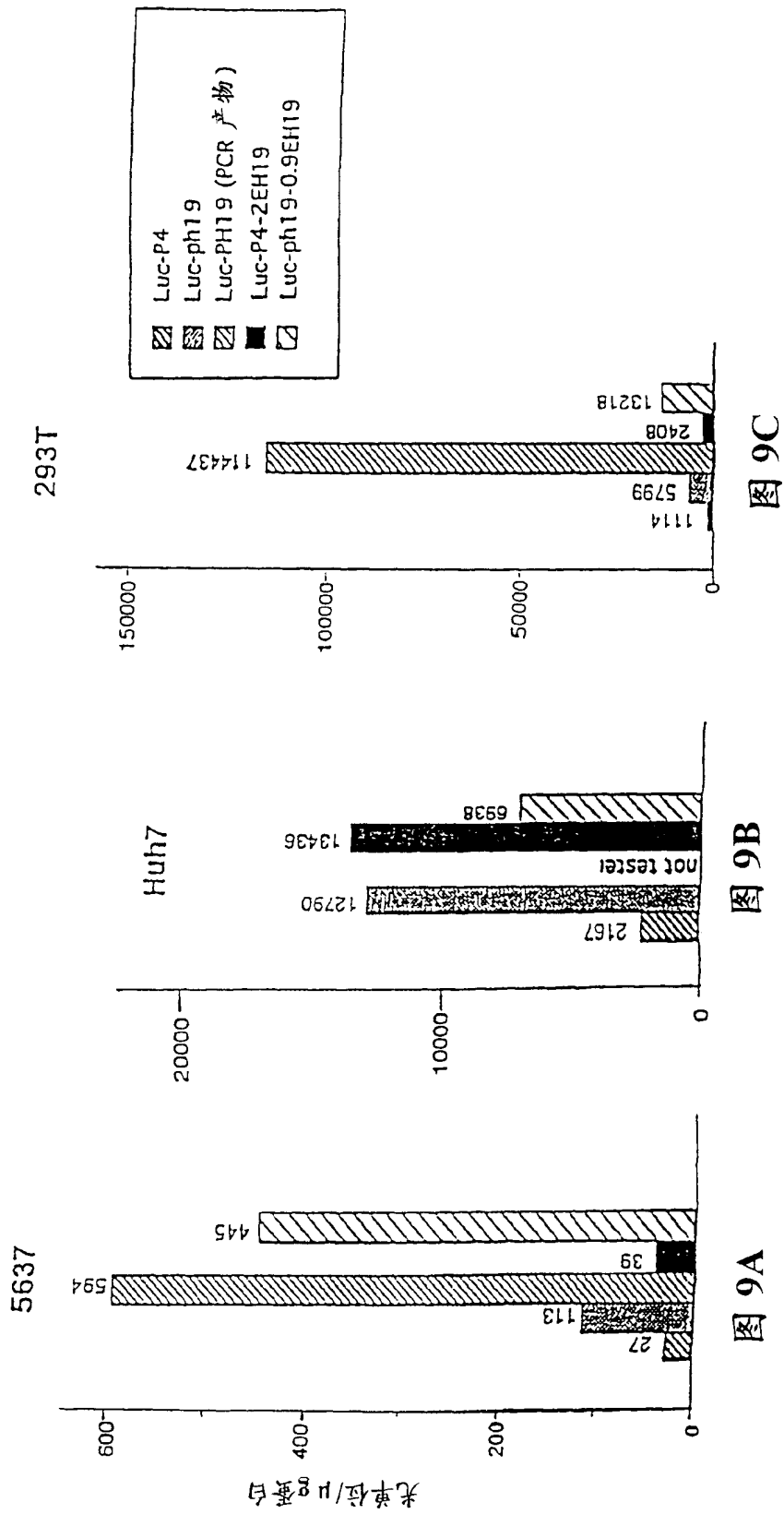
图 8A

1551 CCCAGGGCTC TCCGCCACTC CCCACTCCCC CTTGGGCTCAT AGGCTGGGCC
 1601 CAAGGGTGGG GTGGGATGAG CAGGAGATGG GGCCAGGGG GCAAGCAGGG
 1651 CCCCAAAGAC ATTTAGAAAA ACCGGTTTAT GCAGGCAGCA TTCAGAGCAG
 1701 GCGGCGTGCG TGGCGGGGGC CCTGGGAGCA CAGAGAGGCA CACGTAGGGC
 1751 CCCCAGGGG CTCCCCATTG GCCGGCAGTG ACATCACCCC TGTGTCAACA
 1801 GTGATGTCTG CAGCTCCGGC CAGCCAGGGT TTATGGAGCG AGACCCAGCC
 1851 CGGCCTGGGC CCTCACTCCC CAGGCCACA CACTAGCCCA CTGTTCAGGG
 1901 TCCGGGGTGG CGGCATGGCC TGGGGGTCTT GGCACCGCTG CTCCTCTGCC
 1951 CACCCTAACT TCCCGGCATG GCGGCTGCCC CCTCTGAGCG TCCCCAACCA
 2001 GTAAGTGTGG GGCCAGCAG gCCTGCcGTC CTCcTCcTCT TCCCCTcTAG
 2051 AGAGAAACGT GGAGGTCTTG GGGCTGGGGG CGCTCATAGC CcTGTGACAC
 2101 AGGTGCATGG GGTcAGGGT CCCAGAATGG CCCCTGGGAA GGACCTCAGC
 2151 TGGGCCGGCG GCTcTAGGCT TCAGGGGTCT GTCTGCACAG GGGNTAGCCC
 2201 CTCCAGACC TCTGTGAAGC CAGTACGGC CTCCCCTCCC TGCCCCGTGC
 2251 TCTGTCCGGT GCTTCCTGGA CTGCACTGCG GGCCACTGGT GAGAGGGTGG
 2301 ACAGGGAAGG GCCGCCGTGG TGCCTGTTCC TGCCACCTG GCTGTGTGGT
 2351 CCCCTCCAAG TAGGGACAAC CCTTCTGAGG GCTTGGGGG ACCCTGGGGT
 2401 TGCCAGGGCC TCCAGAGCC CTGTGAGCCC CTGGGGGGTC TGGCCTGATG
 2451 CCCCCCTCCA CGTCCAGGGC CGGCTGTGGC CCAGAACCCC AGCTTCCCAG
 2501 CAGGCCGGTG TCGGGTGGTG ACCCAGGAGA GGCCTCGCCT CCACTGAGGG
 2551 GCCACCGACC TCTGTCAGAC CACAGAGACC CCCAAGGA_gT cTGAAGGCTG
 2601 GAGACCCGGG GCTGGGACCA GGTGGGA_cTT TCCCACGGAG CCGTCCCCAG
 2651 GCCCAGCTGG GGACA_cGTCC CcTTcTcTC CAGACACACC CTGCcTGCCA
 2701 CCAGGACACA CCGGCCTGTT GGGGGTcTcT TTAAAGTGCT TGCCACTCTG
 2751 AGGTGA_cTGT CCCTTTCAA AGAGGTTTcT GGGGCCCAGG TGGGA_iGCCT
 2801 CGGCCTGAGC AGGAGGATcT GGGCCGCCAG GGGCTGGGA CTGTcTCCTG
 2851 GGGAAGGAAG CGCCTGGGAG CGTGTGTGCT GACCCAGGAC CATCCAGGGA
 2901 GGCCCGTcTG TGGGGCAAGC GGAAGGGAG CGGCTGGAGA GGCTTGGCCG
 2951 CCCCCGCCCT GCCTCCATT CCTTAGcTCC ATGCCTGTCA ACCTcTGTCA

图 8B

3001 CCCAGTGAGT GATGTCCAGG GGCCCTGGAA AGGTCACAGC ATGTTTGAGC
 3051 GGGGTGAGAG AGAGGGGAAA GGCGGGGGCG GGGAAAAGTA cGTGGAGGAA
 3101 GCTTTAGGCC CAAGGAAGGA GACAGGGTTc TGGGAGGGAG GGAGCCACTG
 3151 GGGCCGCCGG GAAGGTCCCT GCTTGCTGCT GCCACCCAGA ACCCTCGCCT
 3201 cTTAGcTAGC CCCcGCAGCC CCAGCcTTTc TGGCNTGTGg CCCTCTCCCC
 3251 CATCCCCAGG TGTCcTgTGC AACCAGGCcT TGGACCCAAA CCCTCcTGCC
 3301 CCCTCcTCTC CCTCCTCACC CTCCCAATGC AGTGGTCTCC AGCCTGgcTC
 3351 TGCCCTGCCG CAGGTCCCCT CCCcTCATTA CCAGGCCTAG AGCCTCCAGT
 3401 CCCGGTGGCC CCCAGCCcGA GGGTGAACGG CCTCACCTG GGTcGTGGGA
 3451 CAGAGGGCAC GTTCATCAAG AGTGGCTCCC AAGGGACAcG TGGCTGTTTG
 3501 CAGTTCACAG GAAGCATTcG AGATAAGGAG cTTGTTTTCC CAGTGGGCAC
 3551 GGAGCCAGCA GGGGGGCTGT GGGGCAGCCC AGGGTGCAAG GcCAGGcTGT
 3601 GGGGCTGCAG cTGCcTTGGG CCCCACTCCC AGGCCTTTGC GGGAGGTGGG
 3651 AGGCGgGAGG CGGCAGCTGC ACAGTGGCCC CAGGCGAGGC TCTCAGCCCC
 3701 AGTCGCTCTC CGGGTGGGCA GCCCAAGAGG GTCTGGCTGA GCCTCCCACA
 3751 TCTGGGACTC CATCACCCAA CAACTTAATT AAGGCTGAAT TTCACGTGTC
 3801 CTGTGACTTG GGTAGACAAA GCCCCTGTCC AAAGGGGCAG CCAGCCTAAG
 3851 GCAGTGGGGA CGGCGTGGGT GGCGGGCGAC GGGGGAGATG GACAACAGGA
 3901 CCGAGGGTGT GCGGGCGATG GGGGAGATGG ACAAcAGGAC CGAGGGTGTG
 3951 CGGGCGATGG GGGAGATGGA CAACAGGACC GAGGGTGTGC GGGACACGCA
 4001 TGTACTIONCAT GCACGCCAAT GGGGGGCGTG GGAGGCTGGG GAGCAGACAG
 4051 ACTGGGCTGG GCTGGGCGGG AAGGACGGGC AGATG

图 8C



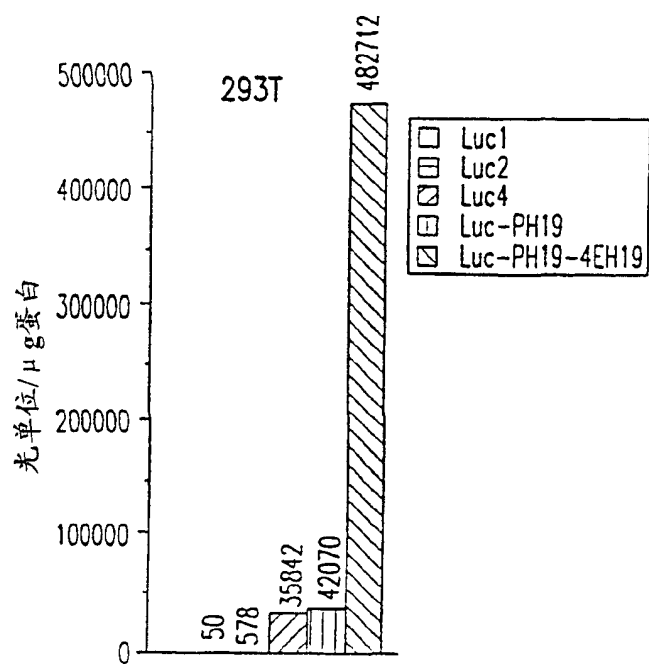


图 10A

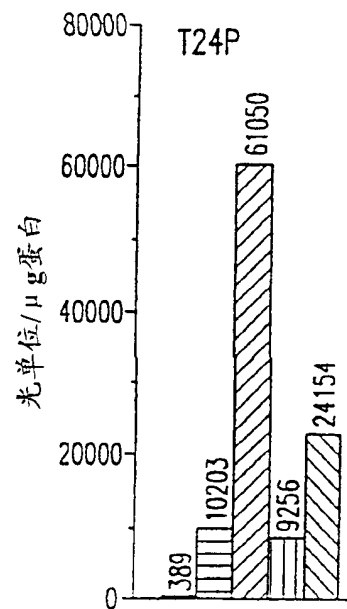


图 10B

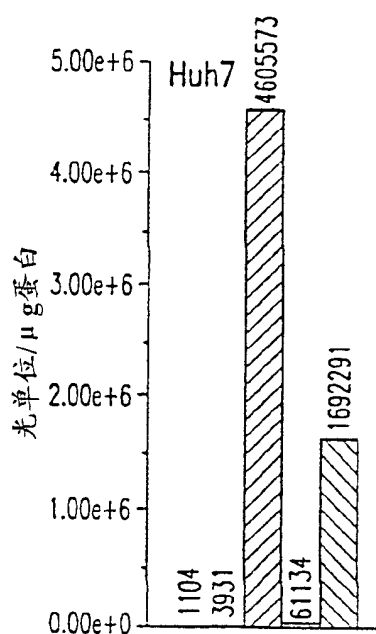


图 10C

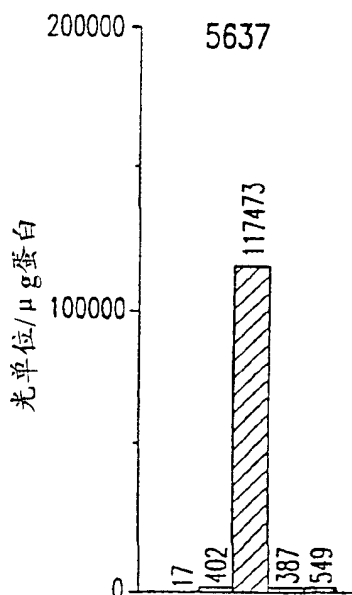


图 10D

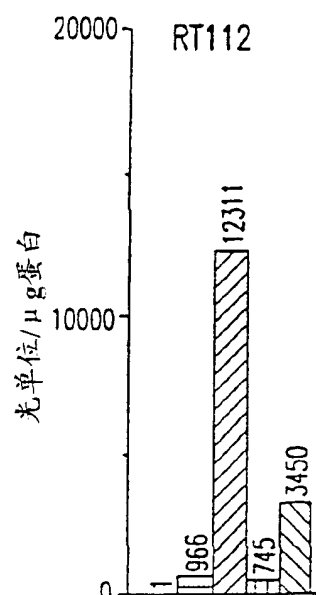


图 10E