

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12Q 1/68

C07K 14/47



[12] 发明专利申请公开说明书

[21] 申请号 02823314. X

[43] 公开日 2005 年 3 月 9 日

[11] 公开号 CN 1592793A

[22] 申请日 2002. 9. 25 [21] 申请号 02823314. X

[30] 优先权

[32] 2001. 9. 25 [33] US [31] 60/324,261

[32] 2002. 8. 23 [33] CA [31] 2,399,569

[86] 国际申请 PCT/JP2002/009873 2002. 9. 25

[87] 国际公布 WO2003/027322 英 2003. 4. 3

[85] 进入国家阶段日期 2004. 5. 24

[71] 申请人 东京大学总长代表日本国

地址 日本东京都

共同申请人 肿瘤疗法科学股份有限公司

[72] 发明人 中村佑辅 古川洋一

[74] 专利代理机构 北京市柳沈律师事务所

代理人 巫肖南 封新琴

权利要求书 1 页 说明书 25 页 序列表 27 页
附图 14 页

[54] 发明名称 肝细胞癌 - 相关基因和多肽, 以及检测肝细胞癌的方法

[57] 摘要

本发明提供在肝细胞癌中上调的基因, 和由这些基因所编码的多肽。还提供载体, 转化子和生产所述重组多肽的方法。还提供这些基因的探针和引物以及针对所述多肽的抗体。所述探针, 引物和抗体可用作检测肝细胞癌的试剂。还提供利用所述试剂检测肝细胞癌的方法。还提供这些基因的反义核苷酸序列, 其可用于抑制肝细胞癌的生长。

I S S N 1 0 0 8 - 4 2 7 4

1. 选自下述的经分离的核酸：
 - (a) 包含 SEQ ID NO : 1 或 NO : 3 所示核苷酸序列的核酸;
 - 5 (b) 编码包含 SEQ ID NO : 2 或 NO : 4 所示氨基酸序列的多肽的核酸;
 - (c) 包含在严紧条件下可与由 SEQ ID NO : 1 或 NO : 3 组成的核苷酸序列或其互补链杂交的链的核酸。
2. 选自下述的经分离的多肽：
 - (a) 由 SEQ ID NO : 1 或 NO : 3 所示核苷酸序列编码的多肽
 - 10 (b) 包含 SEQ ID NO : 2 或 NO : 4 所示氨基酸序列的多肽;
 - (c) 与 SEQ ID NO : 2 或 NO : 4 具有至少 65 % 同一性的多肽。
3. 带有权利要求 1 所述核酸的载体。
4. 带有权利要求 1 所述核酸或权利要求 3 所述载体的转化子。
5. 生产多肽的方法, 该方法包括在培养物中培养权利要求 4 的转化子,
- 15 在所述转化子中表达所述多肽, 和从培养物中回收所述多肽。
6. 可与权利要求 2 所述多肽特异结合的抗体。
7. 检测肝细胞癌的方法, 该方法包括下述步骤：
 - (a) 制备来自受试者的生物样品;
 - (b) 测定至少一种选自下组的多肽的表达水平, 所述多肽是 SEQ ID
 - 20 NO : 1, SEQ ID NO : 3, 和 SEQ ID NO : 5 所示的多肽;
 - (c) 将上述表达水平与在非癌症样品中的测定值相比较; 和
 - (d) 确定受试者是否患癌症。
8. 一种检测肝细胞癌的试剂, 其包含一种核酸, 该核酸含有在严紧条件下可与由 SEQ ID NO : 1, NO : 3 或 NO : 5 组成的核苷酸序列或其互补链
- 25 杂交的链。
9. 检测肝细胞癌的试剂, 其包含权利要求 6 所述抗体。
10. 抑制肝细胞癌生长的方法, 该方法包括向肝细胞癌中引入至少一种可与 SEQ ID NO : 1, NO : 3 或 NO : 5 杂交的反义寡核苷酸。

肝细胞癌-相关基因和多肽, 以及检测肝细胞癌的方法

5 技术领域

本发明涉及在肝细胞癌中上调的基因, 由该基因编码的多肽以及检测肝细胞癌的方法。

背景技术

10 利用 cDNA 微阵列技术可以获得正常细胞与恶性细胞内基因表达的详尽分布图 (Perou, C. M. 等, Nature. 406: 747-752, 2000; Clark, E. A. 等, Nature. 406: 532-535, 2000; Okabe, H. 等, Cancer Res. 61: 2129-2137, 2001)。该方法揭示出癌细胞的复杂的性质, 有助于增进对致癌作用的了解。对在肿瘤中下调的基因的鉴定可对具体癌症进行更加准确无误的诊断, 并开发出新的治疗靶子 (Golub, T. R. 等, Science 286: 531-537, 1999)。

15 肝细胞癌(HCC)是世界范围内癌症死亡的主要原因。尽管近来治疗策略有所改进, 但晚期 HCC 患者的预后仍然很差。尽管分子水平的研究揭示出肝癌的成因与 *TP53*, *CTNBI* 和/或 *AXIN1* 基因的改变相关 (Perou, C. M. 等, Nature. 406: 747-752, 2000; Satoh, S. 等, Nat Genet. 24: 245-250, 2000), 但这些改变似乎只涉及部分 HCC。因此, 仍然需要可以作为诊断标记和/或
20 治疗癌症的药物靶子的主要基因。

本发明人早先通过在 HCC 基因组范围内进行的分析, 鉴定并报道了新基因 *VANGL1* (Yagy, R. 等, 国际肿瘤学杂志(International Journal of Oncology) 20: 1173-1178, 2002)。

发明内容

25 本发明的一个目的在于提供在肝细胞癌中上调的基因, 由该基因编码的多肽以及检测肝细胞癌的方法。

本发明人利用能显示 23,040 个基因的 cDNA 微阵列, 研究了 HCC 的表达分布图。通过这些努力精确定位了 165 个基因, 包括 69 个 EST, 与相应的非癌肝细胞相比它们在癌症组织中似乎高频上调。本发明人从在 HCC
30 中的表达经常提高的转录本中分离了三个基因。这些基因编码与矢车菊根皮素(centaurin)-家族蛋白质具有共同结构特征的产物。

三个基因中的一个相应于单基因簇(UniGene cluster)EST Hs.44579, 该基因是在染色体带 1p36.13 上过表达的新基因。由于该基因的开放阅读框编码的蛋白与发育分化增强因子 2 (*DDEF2*)具有约 60% 同一性, 本发明人将该基因称作发育分化增强因子样 1(*DDEFL1*)。

- 5 另一个在 HCC 中上调的基因相应于单基因簇 EST(hs. 122730)。其推定的氨基酸序列分别与在果蝇中涉及细胞极性和决定细胞命运的斜视(*Strabismus*)(Van Gogh)和 Van Gogh 样 2 (*VANGL2*)具有 40% 和 63% 的同一性。因此, 将该基因称作 Van Gogh 样 1 (*VANGL1*)。

所发现的另一个在 HCC 中上调的基因是 LGN(Genbank 登录号
10 U54999)。LGN 蛋白与抑制性异源三聚体 G 蛋白的 α 亚基($G\alpha_{12}$)相互作用。

DDEFL1 或 LGN 的基因转移促进缺乏任何一个上述基因内源表达的细胞的增殖。此外, *DDEFL1*, *VANGL1* 或 LGN 的表达通过转染它们的特定反义 S-寡核苷酸而降低后, 可抑制肝细胞癌细胞的生长。

上述发现有助于阐明 HCC 的机制, 并研究出新的策略对 HCC 进行诊
15 断和治疗。

本发明具体地提供

(1)选自下述的经分离的核酸:

- (a)包含 SEQ ID NO: 1 或 NO: 3 所示核苷酸序列的核酸;
(b)编码包含 SEQ ID NO: 2 或 NO: 4 所示氨基酸序列的多肽的核酸;
20 (c)包含在严紧条件下可与由 SEQ ID NO: 1 或 NO: 3 或其互补链组成的核苷酸序列杂交的链的核酸,

(2)选自下述的经分离的多肽:

- (a)由 SEQ ID NO: 1 或 NO: 3 所示核苷酸序列编码的多肽;
(b)包含 SEQ ID NO: 2 或 NO: 4 所示氨基酸序列的多肽;
25 (c)与 SEQ ID NO: 2 或 NO: 4 具有至少 65% 同一性的多肽,
(3)带有(1)中的核酸的载体,
(4)带有(1)中的核酸或(3)中的载体的转化子,
(5)生产多肽的方法, 该方法包括在培养基中培养(4)中的转化子, 在所述转化子中表达所述多肽, 和从培养基中回收所述多肽,

30 (6)可与(2)中的多肽特异结合的抗体,

(7)检测肝细胞癌的方法, 该方法包括下述步骤:

- (a) 制备受试者的生物样品;
- (b) 测定至少一种选自如 SEQ ID NO : 1, SEQ ID NO : 3, 和 SEQ ID NO : 5 所示的多肽的表达水平;
- (c) 将上述表达水平与在非癌症样品中的测定值相比较; 和
- 5 (d) 确定受试者是否患癌症,
- (8) 检测肝细胞癌的试剂, 其包含在严紧条件下可与选自 SEQ ID NO : 1, NO : 3 或 NO : 5 的核苷酸序列或其互补序列杂交的链的核酸,
- (9) 检测肝细胞癌的试剂, 其包含(6)中的抗体, 和
- (10) 抑制肝细胞癌生长的方法, 该方法包括向肝细胞癌中引入至少一
- 10 种可与 SEQ ID NO : 1, NO : 3 或 NO : 5 的核苷酸序列杂交的反义寡核苷酸。
- 下面对本发明进行详细描述。

核 酸

本发明提供在肝细胞癌中上调的基因。

- DDEFL1* 的核苷酸序列和氨基酸序列分别如 SEQ ID NO : 1 和 NO : 2
- 15 所示。完整的 *DDEFL1* cDNA 由 4050 个核苷酸组成, 具有由 2712 个核苷酸组成、并编码 903 个氨基酸的蛋白(Genbank 登录号 AB051853)的开放阅读框。*DDFEL1* 的氨基酸序列与 *DDEFL2* 具有 60% 的同一性, 与 *DDEF/ASAP1* 具有 46% 的同一性, 其包含一个 Arf GTP 酶-激活蛋白(ArfGAP)结构域和两个锚蛋白(ankyrin)重复。

- 20 *DDEFL1* 与调节肌动蛋白细胞骨架重新组织相关的矢车菊根皮素家族成员 *DDEF2* 具有 60% 的同一性。这表明 *DDEFL1* 可能也在细胞结构的组建中起作用(Randazzo, P. A.等, Arf GTP 激活蛋白 ASAP1 调节肌动蛋白细胞骨架(The Arf GTPase-activating protein ASAP1 regulates the actin cytoskeleton,) Proc. Natl. Acad. Sci. U S A 97: 4011-4016, 2000)。由于
- 25 *DDEFL1* 也保持 PH 结构域和 ArfGAP 基元, 所以它似乎是通过 GAP 激活的方式调节 Arf 小 GTP 酶的矢车菊根皮素家族新成员。据认为可在 GEFs 的 Dbl 家族中的大多数分子中观察到的 PH 结构域是通过与特定的靶分子相互作用, 和/或通过直接调节催化结构域在蛋白的重新定位中起关键作用
- (Jackson, T. R.等, Trends Biochem Sci. 25: 489-495, 2000; Cerione, R. A.和
- 30 Zheng, Y., Curr. Opin. Cell. Biol. 8: 216-222, 1996; Chardin, P.等, Nature 384:

481-484, 1996)。虽然 *DDEF2* 定位于外周的粘着斑(focal adhesion), 本发明人发现 myc-标记的 *DDEFL1* 蛋白质在细胞质中扩散。

Arf 蛋白质涉及重要的细胞过程, 如液泡膜运输, 内质网(ER)和高尔基体内部空间(compartment)完整性的维持, 以及外周细胞骨架的调整
5 (*Cukierman, E.*等, *Science* 270: 1999-2002, 1995)。迄今为止, 已鉴定出六个 Arf 家族成员(Arf1-Arf6)及其功能 (*Moss, J. and Vaughan, M., J. Biol. Chem.* 270: 12327-12330, 1995)。例如: 研究表明, Arf6 蛋白可以作为细胞骨架的调节剂来改变粘着斑的形态并阻止细胞的扩散, 且 *DDEF2* 表现出对 Arf1 的 GAP 活性。

10 *DDEFL1* 的过表达增进促生长作用和细胞在低血清条件下的存活。这表明 *DDEFL1* 有助于癌细胞在贫营养和含氧量低的条件下生长。*DDEFL1* 在 HCC 中频繁的上调强调了该基因在肝癌发生中的重要性。

VANGL1 的核苷酸序列和氨基酸序列分别如 SEQ ID NO: 3 和 NO: 4 所示。确定的 cDNA 序列由 1879 个核苷酸组成, 包含一个编码 524 个氨基酸
15 的蛋白质(Genbank 登录号 AB057596)的由 1572 个核苷酸组成的开放阅读框。

经鉴定斜视(*stbm*)是决定果蝇突变体具有毛糙眼表型的基因(*Wolff T.*和 *Rubin G.M., Development* 125:1149-1159, 1998)。该基因是维持眼, 腿和刚毛的极性所需的, 并在果蝇中决定 R3 和 R4 光感受器的细胞命运。与 *stbm* 同源的小鼠基因 *Ltap* 在神经管突变体小鼠 Loop-尾中改变, 所述小鼠是神经管缺陷(NTD)的动物模型(*Kibar Z* 等, *Nat Genet.* 28: 251-255, 2001)。因此,
20 *VANGL1* 可能还在细胞极性, 细胞命运决定和/或组织结构中起重要作用。由于 *VANGL1* 在 HCC 中频繁地上调, 并且抑制其表达将显著地减少癌细胞的生长或残留, 所以 *VANGL1* 可以使癌细胞延长存活时间和/或进行去极化
25 生长。

LGN 的核苷酸序列和氨基酸序列分别如 SEQ ID NO: 5 和 NO: 6 所示。*LGN* 的 cDNA 由 2336 个核苷酸组成, 编码 677 个氨基酸的肽。

从前报道 *LGN* 蛋白是与抑制性的异源三聚体 G 蛋白 α 亚基(Gai2)相互作用的蛋白) (*Mochizuki, N.* 等, *Gene* 181: 39-43, 1996)。Gai2 的激活突变曾经
30 在垂体瘤等内分泌肿瘤中有过报道 (*Hermouet, S.* 等, *Proc. Natl. Acad. Sci. USA* 88: 10455-10459, 1991; *Pace, A. M.* 等, *Proc. Natl. Acad. Sci. USA.* 88:

7031-7035, 1991; Lyons, J. et al, Science 249: 655-659, 1990)。然而, 未见到 *LGN* 涉及肿瘤发生或致癌作用的报告。集落形成试验表明 *LGN* 可能具有致癌的活性。*LGN* 增强表达可以活化 *Gai2* 并且介导肝癌发生中的致癌信号。

本发明的核酸包括 cDNA, 基因组 DNA, 化学合成的 DNA 和 RNA。

5 可以是单链或双链。

此处所用“经分离的核酸”是指一种核酸, 其结构不等同于任何天然存在的核酸, 或跨越三个以上独立基因的天然基因组核酸的任何片段。因此该术语包括, 例如 (a) 一种 DNA, 其具有该 DNA 分子天然状态下所在生物体的基因组中天然基因组 DNA 分子的一部分序列; (b) 一种核酸, 其整合到载体或原核或真核生物基因组 DNA 中, 使所得的分子不同于任何天然载体或基因组 DNA; (c) 单独的分子例如 cDNA, 基因组片段, 由聚合酶链式反应(PCR)产生的片段, 或限制性片段; 和 (d) 重组核苷酸序列, 作为杂合基因(即编码融合蛋白的基因)的一部分。具体地说, 从这一定义中排除的是下述的核酸, 即存在于由不同的(i)DNA 分子, (ii)转染细胞, 或(iii)细胞克隆形成的混合物中的 DNA 分子, 例如当它们出现在 DNA 文库, 如 cDNA 或基因组 DNA 文库中。

在一个实施方案中, 本发明的核酸包括, 包含 *DDEFL1* 或 *VANGL1* 的核苷酸序列(具体说来是 SEQ ID NO: 1 或 NO: 3)的核酸。

在另一个实施方案中, 本发明的核酸包括编码包含 *DDEFL1* 或 *VANGL1* 的氨基酸序列的多肽(具体说来是 SEQ ID NO: 2 或 NO: 4)的核酸。因而, 包含基于遗传密码简并性的任意序列的核酸均包括在内。

在又一个实施方案中本发明的核酸包含 SEQ ID NO: 1 或 NO: 3 的变体核酸序列。所述变体包括含有在严紧条件下可与由 SEQ ID NO: 1 或 NO: 3 组成的核苷酸序列或其互补链杂交的链的核酸。

25 此处所用的术语“互补”是指双链核酸的其中一条链的全部碱基均能与另一条链的碱基形成碱基对。另外, “互补”不仅指在至少 15 个毗连核苷酸形成的连续区域内完全匹配的那些, 还包括在该区域内具有至少 65%, 优选 70%, 更优选 80%, 更优选 90%, 最优选 95% 或更高同一性的那些。

30 此处所述两核酸的“百分比同一性”是利用 Karlin 和 Altschul 的算法 (Proc. Natl. Acad. Sci. USA 87: 2264-2268, 1990)经 Karlin 和 Altschul 修改 (Proc. Natl. Acad. Sci. USA 90:5873-5877, 1993)后的算法确定的。该算法已并

入 Altschul 等的 NBLAST 和 XBLAST 程序 (J. Mol. Biol. 215:403-410, 1990)。BLAST 核苷酸检索是利用 NBLAST 程序进行的, 得分(score)= 100, 字长(wordlength)= 12。蛋白质的同源检索可在, 例如在日本 DNA 数据库(DDBJ) 中通过 FASTA 程序, BLAST 程序等进行。BLAST 蛋白质检索是利用
5 XBLAST 程序进行的, 得分(score)= 50, 字长(wordlength) =3。当两序列间存在缺口(gap)时, 利用 Gapped BLAST 进行, 参见 Altsuchl 等 (Nucleic Acids Res. 25: 3389-3402, 1997)。运用 BLAST 和 Gapped BLAST 程序时, 应用各程序(例如 XBLAST 和 NBLAST)的缺省参数。

优选地, 该变体包括与 SEQ ID NO : 1 或 NO : 3 所示的核苷酸序列具有
10 至少 65% 同一性的核苷酸序列。更优选, 该变体与 SEQ ID NO : 1 或 NO : 3 所示的核苷酸序列具有至少 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 或更高的同一性。当所述变体在长度上等同于或长于参照序列, 如 SEQ ID NO : 1 或 NO : 3 时, 所述比较是基于全长的参照序列作出的。当所述变体比参照序列, 例如 SEQ ID NO : 或 NO : 3 短时, 所
15 述比较是基于参照序列中与变体序列相同长度的节段进行的(排除同源性计算时所需的任何环)。

所述严紧条件杂交定义为在下述条件下进行的平衡杂交: 42°C, 2 x SSC, 0.1% SDS (低严紧); 50°C, 2 x SSC, 0.1% SDS (中度严紧); 65°C, 2 x SSC, 0, 1% SDS (高度严紧)。如果需要通过洗涤达到平衡, 所述洗涤利用适合于
20 具体严紧性要求的杂交溶液进行。通常, 温度越高在平衡状态下杂交的两条链间的同源性越高。

对本发明的核酸的长度没有限制, 但其优选包含至少 15, 20, 30, 40 50, 100, 150, 200, 300, 400, 500 1000, 1500, 2000, 2500, 或 3000 个核苷酸。

本发明的核酸包括用作探针或引物的多核苷酸, 所述探针或引物可与
25 SEQ ID NO : 1 或 NO : 3 或其互补链特异杂交。术语“特异杂交”是指在通常的杂交条件下, 优选在严紧条件下与 SEQ ID NO : 1 或 NO : 3 所示的核苷酸序列杂交, 但不与编码其它多肽的 DNA 交叉杂交。

所述引物和探针包含 SEQ ID NO : 1 或 3 所示核苷酸序列或其互补序列
30 中的至少 15 个连续的核苷酸。通常, 所述引物包含 15 到 100 个核苷酸, 优选包含 15 到 35 个核苷酸, 所述探针包含至少 15 个核苷酸, 优选包含至少 30 个核苷酸, 包含 SEQ ID NO : 1 或 NO : 3 所示序列的部分或全部。所述引

物可用于扩增编码本发明多肽的核酸，所述探针可用于分离或检测编码本发明的多肽的核酸。本发明的引物和探针可利用，例如可商购的寡核苷酸合成仪制备。所述探针也可以通过限制酶处理等方式制备成双链DNA片段。

5 本发明的核酸包括可与 SEQ ID NO : 1 或 3 的核苷酸序列中任意位点杂交的反义寡核苷酸。本文术语“反义寡核苷酸”不仅指其全部核苷酸均与相应的 DNA 或 mRNA 特异区域互补的那些寡核苷酸，还包括具有一个或多个错配的核苷酸的寡核苷酸，只要 DNA 或 mRNA 和寡核苷酸能与 SEQ ID NO : 1 或 NO : 3 的核苷酸序列杂交即可。

10 优选所述反义寡核苷酸针对 SEQ ID NO : 1 或 NO : 3 中的至少 15 个连续的核苷酸。更优选上述的反义寡核苷酸包含上述至少 15 连续的核苷酸中的起始密码子。

本发明的反义寡核苷酸包括含低级烷基磷酸酯(例如，甲基磷酸酯或乙基磷酸酯)，硫代磷酸酯(phosphothioate)和磷酰胺酸酯(phosphoramidate)的类似物。

15 本发明的反义寡核苷酸，通过与编码本发明多肽的 DNA 或 mRNA 相结合并抑制其转录或翻译，从而作用于产生本发明的多肽的细胞，促进所述 mRNA 的降解从而抑制本发明多肽的表达。

20 本发明核酸的制备如下。可例如根据编码本发明多肽的 DNA 的核苷酸信息(例如 SEQ ID NO : 1 或 NO : 3)制备引物，通过噬斑 PCR (Affara NA 等 (1994) Genomics 22, 205-210)制备编码本发明多肽的 cDNA。利用可商购的“Qiagen 基因组 DNA 试剂盒(Qiagen, Hilden, Germany)的方法制备基因组 DNA。所得 DNA 的核苷酸序列可以利用本领域已知的方法，例如可商购的“染料终止子测序试剂盒”(Applied Biosystems)确定。本发明的核酸，如下文所述可用于生产重组蛋白并检测肝细胞癌。

25 载体，转化子和重组多肽的生产

本发明还涉及插入了本发明核酸的载体。

本发明的载体包括制备本发明重组多肽的载体。任何能够表达本发明多肽的载体均可使用。

30 表达载体的实例包括细菌(例如大肠杆菌)表达载体，酵母表达载体，昆虫表达载体，和哺乳动物表达载体。在本发明中，哺乳动物表达载体如，

pcDNA3.1-myc/His 或 pcDNA 3.1 载体(Invitrogen)均可使用。可利用本领域公知的方法将本发明的核酸插入到载体中。

5 本发明的载体还包括在体内表达本发明多肽的载体(尤其用于基因治疗)。各种病毒载体和非病毒载体只要其能在体内表达本发明的多肽均可使用。病毒载体的实例包括腺病毒载体, 反转录病毒载体等等。阳离子脂质体可作为非病毒的载体的例子。

10 本发明还提供以可表达的方式带有本发明核酸的转化子。本发明的转化子包括带有插入了本发明核酸的上述表达载体的转化子, 以及带有整合了本发明核酸的基因组的转化子。本发明的核酸可以任何形式保留在所述转化子中, 只要所述转化子可以表达所述核酸即可。

15 本发明对插入了所述载体的细胞没有特殊的限制, 只要所述载体能在所述细胞中表达本发明的核酸即可。例如大肠杆菌, 酵母, 哺乳动物细胞和昆虫细胞均能用作宿主细胞。优选的哺乳动物细胞如 COS7 细胞和 NIH3T3 细胞。可用本领域已知的方法, 如电穿孔法和磷酸钙法将载体引入细胞。

20 可用本领域的常规方法从转化子中分离纯化所述重组多肽。例如, 在收集了转化子并获得了提取物之后可通过离子交换层析, 反向层析法, 凝胶过滤, 或在柱子中固定了针对本发明多肽的抗体的亲和层析法, 或通过这几种柱子的组合, 来纯化并制备目的多肽。

25 当本发明的多肽在宿主细胞(例如动物细胞, 大肠杆菌)中作为与谷胱甘肽-S-转移酶蛋白的融合蛋白表达, 或作为添加了多个组氨酸的重组多肽表达时, 可利用谷胱甘肽柱或镍柱纯化所表达的重组多肽。融合蛋白纯化后还可以根据需要用凝血酶或 Xa-因子切除非目的多肽区域。

多肽

25 本发明提供由 *DDEFL1* 或 *VANGL1* (例如 SEQ ID NO:1 或 NO:3)编码的经分离的多肽。在具体实施方案中, 本发明的多肽包括由 SEQ ID NO:1 或 NO:3 的核苷酸序列编码的多肽, 和包含 SEQ ID NO:2 或 NO:4 所示氨基酸序列的多肽。

30 所述“经分离的多肽”指基本纯且不含其它生物大分子的多肽。基本纯的多肽是干重至少 75%(例如, 至少 80, 85, 95, 或 99%)纯的。纯度可

通过任何适宜的标准方法测定，例如通过柱层析法，聚丙烯酰胺凝胶电泳，或高效液相色谱分析。

本发明的多肽包括 SEQ ID NO :2 或 NO :4 的变体，只要所述变体与 SEQ ID NO :2 或 NO :4 具有至少 65% 的同一性即可。所述变体可以是，包含 SEQ ID NO :2 或 NO :4 所示氨基酸序列、且其中有一个或多个氨基酸被取代，缺
5 失，添加和/或插入的多肽。所述变体也可以是由包含在严紧条件下可与由 SEQ ID NO :1 或 NO :3 所示核苷酸序列杂交的链的核酸所编码的多肽。

已知具有在具体氨基酸序列中缺失，添加和/或替代一个或多个氨基酸残基所得的氨基酸序列的多肽可保持原始的生物活性(Mark, D. F. 等, Proc. Natl. Acad. Sci. USA (1984) 81, 5662-5666, Zoller, M. J. & Smith, M., Nucleic Acids Research (1982) 10, 6487-6500, Wang, A. 等, Science 224, 1431-1433, Dalbadie-McFarland, G. 等, Proc. Natl. Acad. Sci. USA (1982) 79, 6409-6413)。
10

通过取代，缺失，添加，和/或插入而突变的氨基酸的数目没有特别的限定。通常，该数目是氨基酸残基总数的 10% 或更少，优选 5% 或更少，
15 更优选 1% 或更少。

将要突变的氨基酸残基优选经过突变成为侧链性质保守的不同氨基酸。氨基酸侧链性质的实例是：疏水的氨基酸 (A, I, L, M, F, P, W, Y, V)，亲水氨基酸(R, D, N, C, E, Q, G, H, K, S, T)，以及含有下列侧链的氨基酸：脂肪族侧链(G, A, V, L, I, P)；包含羟基的侧链(S, T, Y)；含硫原子的侧链(C, M)；
20 含羧酸和酰胺的侧链(D, N, E, Q)；包含碱基的侧链(R, K, H)；和包含芳族的侧链(H, F, Y, W)(括号中的字母是氨基酸的单字母密码)。“保守氨基酸取代”是用上述一组中氨基酸取代同组中的另一氨基酸。

缺失变体包括 SEQ ID NO :1 或 NO :3 所示氨基酸序列的片段。该片段是具有与本发明的上述多肽部分相同但非全部相同的氨基酸序列的多肽。
25 本发明的多肽片段通常由 8 个或更多氨基酸残基组成，优选由 12 个或更多氨基酸残基组成(例如，15 个或更多氨基酸残基)。优选片段的实例包括下述的截短多肽：具有缺失了一组或两组氨基酸残基的氨基酸序列的多肽，其中所述一组缺失包括氨基末端或羧基末端的缺失，所述两组缺失分别包括氨基末端和羧基末端的缺失。此外，还优选具有结构或功能特点的片段，其
30 包括那些具有 α -螺旋和 α -螺旋形成区， β -折叠和 β -折叠形成区，转角和转角形成区，卷曲和卷曲形成区，亲水区域，疏水区域， α -两亲性区域， β -两亲性

区域, 可变区, 表面成形区域, 底物结合区域, 和高抗原性指数区域。也优选生物学活性片段。生物学活性片段可以介导本发明多肽的活性, 所述片段包括那些具有类似的或改进的活性的片段和减少了不良活性的片段。例如具有通过结合配体将信号转导到细胞内的活性的片段, 以及在动物(包括人)中具有抗原性或免疫原性的片段。这些多肽片段优选保持本发明多肽的抗原性。

所述添加变体包括本发明多肽与另一肽或多肽的融合蛋白。融合蛋白可利用本领域技术人员已知的方法制备, 例如将编码本发明多肽的 DNA 与编码其它肽或多肽的 DNA 相连接, 使框架匹配, 再将其插入表达载体并在宿主中表达。对与本发明的多肽相融合的肽或多肽没有限制。

已知的肽, 例如 FLAG (Hopp, T.P.等, Biotechnology (1988) 6, 1204 1210), 6xHis(包含六个 His(组氨酸)残基), 10xHis, 流感病毒凝集素(HA), 人 C-myc 片段, VSP-GP 片段, p18HIV 片段, T7 标记, HSV-标记, E-标记, SV40 T 抗原片段, lck 标记, α -微管蛋白片段, B-标记, 蛋白 C 片段, 等等均可用作与本发明的多肽相融合的肽。与本发明的多肽相融合的多肽的例子, 如 GST (谷胱甘肽-S-转移酶), 流感病毒凝集素(HA), 免疫球蛋白恒定区, β -半乳糖苷酶, MBP (麦芽糖-结合蛋白)等等。

通过将可商购的编码这些肽或多肽的 DNA 与编码本发明的多肽的 DNA 相融合, 并表达所述融合 DNA, 可以制备融合蛋白。

变体多肽优选与 SEQ ID NO :2 或 NO :4 所示氨基酸序列具有至少 65 % 的同一性。更具体地说, 所述修饰多肽与 SEQ ID NO :2 或 NO :4 所示氨基酸序列具有至少 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 %, 或以上的同一性。当经修饰的多肽长于或等同于参考序列, 例如 SEQ ID NO :2 或 NO :4 时, 所述比较是基于全长的参考序列进行的。当经修饰的多肽比参照序列, 例如 SEQ ID NO :2 或 NO :4 短时, 所述比较是基于参照序列中和变体序列相同长度的节段进行的。

此处所用两个氨基酸序列间的“百分同一性”是应用与上述核酸的描述中同样的方法确定的。

本发明的多肽可应用本领域技术人员已知的方法, 通过上述的遗传工程技术制备成天然的多肽或重组多肽。天然的多肽可通过下述方式获得, 即用重组多肽免疫小动物获得抗体, 将该抗体与柱子偶联, 利用这种柱子

通过亲和层析提取高水平表达本发明多肽的肝组织或细胞。制备重组多肽可以是，将编码本发明多肽的 DNA (例如包含 SEQ ID NO : 或 3 所示核苷酸序列的 DNA) 插入到合适的表达载体中，将所述表达载体引入宿主细胞，使所得的转化子表达所述多肽，再回收所表达的多肽。

- 5 可以通过已知的方法，例如 PCR-介导的定点诱变诱导系统 (GIBCO-BRL, Gaithersburg, Maryland), 寡核苷酸-介导的特定区域(sight - directed)-诱变 (Kramer, W. and Fritz, HJ (1987) 酶学方法 (Methods in Enzymol) 154:350-367) 将突变插入到 SEQ ID NO : 1 或 NO : 3 的氨基酸序列中，来制备变体多肽。

10 抗体

本发明还涉及可与本发明的多肽特异结合的抗体。对本发明的抗体的形式没有特别地限制，其包括多克隆抗体和单克隆抗体。还包括用本发明的多肽免疫动物(如，兔)制备的抗血清，各类多克隆和单克隆抗体，由遗传工程制备的人源化抗体，以及人抗体。

- 15 多克隆抗体可通过下述方式制备，即收获用本发明的多肽免疫小动物(如兔)所获得的血清，通过偶联了本发明多肽的亲和层析柱取得仅仅识别本发明多肽的组分，再利用蛋白 G 或蛋白 A 柱从上述组分中纯化免疫球蛋白 G 或 M。

- 20 制备单克隆抗体是通过下述方式进行的，即用本发明的多肽免疫小动物(如小鼠)，取出所述动物的脾脏，将所述器官匀浆成细胞，再利用试剂，如聚乙二醇将所述细胞与小鼠骨髓瘤细胞融合，从融合细胞(杂交瘤)中选择能产生针对本发明多肽的抗体的克隆，将所得的杂交瘤移植到小鼠的腹腔中，提取腹水。所得的单克隆抗体可通过，例如硫酸铵沉淀，蛋白 A 或蛋白 G 柱，DEAE 离子交换层析，或偶联了本发明多肽的亲和层析柱进行纯
25 化。本发明的抗体可用于纯化和检测本发明的多肽。特别是可用于检测肝细胞癌。

- 人抗体或人源化抗体可通过本领域技术人员已知的方法制备。例如可用本发明的多肽免疫已将其免疫系统转变为小鼠免疫系统的小鼠来制备人抗体。人源化的抗体可通过下述方法制备，例如从单克隆抗体产生细胞克隆
30 出所述抗体的基因，利用 CDR 移植方法将所述基因中的抗原识别部位移植到已知的人抗体中。

检测方法

本发明还提供利用 *DDEFL1 VANGL1* 或 *LGN* 多肽作为标记检测肝细胞癌的方法。

所述检测是通过下述方法进行的，测定来自受试者的生物样品中
5 *DDEFL1 VANGL1*，和 *LGN* 多肽至少其中之一的表达水平，将其与非癌样品中的表达水平相比较，确定受试者是否患癌症。

此处所用的生物样品包括来自需要进行肝细胞癌检测的受试者的任何肝组织或细胞。特别是可以使用肝活组织检查样品。所述生物样品还包括由肝组织或细胞制备的 mRNA, cRNA 或 cDNA 样品。mRNA 和 cDNA 样品
10 可由常规方法制备。cRNA 是指用 RNA 聚合酶从模板 cDNA 转录所得的 RNA。cRNA 可以以连接有 T7 启动子的 cDNA 为模板，利用 T7 RNA 聚合酶来合成。可以利用为获得基于 DNA 芯片的表达分布图的可商购 cRNA 转录试剂盒。

在具体实施方案中，可以通过 RNA, cDNA 或多肽的水平测定 *DDEFL1*,
15 *VANGL1* 或 *LGN* 多肽的表达水平。

mRNA 的表达水平可以通过例如，利用可与 *DDEFL1*, *VANGL1*，或 *LGN* 的核苷酸序列杂交的探针进行 Northern 印迹，或利用可与 *DDEFL1*, *VANGL1*，或 *LGN* 的核苷酸序列杂交的引物进行 RT-PCR 来测定。

用于本发明的检测方法的探针和引物包括可与 SEQ ID NO : 1, NO : 3
20 或 NO : 5 的核苷酸序列或其互补序列特异杂交的核酸。术语“特异杂交”是指在正常杂交条件下，优选在严紧条件下，可与 SEQ ID NO : 1, NO : 3 或 NO : 5 的核苷酸序列杂交，而不与编码其它多肽的 DNA 交叉杂交。

所述引物和探针包含 SEQ ID NO : 1, NO : 3 或 NO : 5 的核苷酸序列或它们的互补序列中至少 15 个连续的核苷酸。通常，所述引物包含 15 到 100
25 个核苷酸，优选 15 到 35 个核苷酸，所述探针包含至少 15 个核苷酸，优选至少 30 个核苷酸，其包含 SEQ ID NO : 1, NO : 3 或 NO : 5 的至少一部分或全部序列。所述引物和探针可通过例如可商购的寡核苷酸合成仪制备。所述探针还可以制备成由限制酶处理所得的双链 DNA 片段。

cDNA 表达水平可通过，例如应用 DNA 阵列的方法测定 (Masami
30 Muramatsu 和 Masashi Yamamoto, 新遗传工程手册(New Genetic Engineering Handbook)280-284 页, YODOSHA Co., LTD.)。具体地说，首先提供来自受试

者的 cDNA 样品和固体支持物,其上固定有与 *DDEFL1*, *VANGL1* 或 *LGN* 的核苷酸序列杂交的多核苷酸探针。可以利用如上所述探针。可以在所述固体支持物上固定两种以上的探针以检测两种以上的靶多核苷酸。根据需要可对 cDNA 样品进行标记以用于检测。对于所用的标记并没有特殊地限定,只要其能被检测即可,例如:荧光标记,放射性标记等。可通过常规的方法进行标记 (L. Luo 等,“激光捕获的邻近神经原亚型基因表达分布图(Gene expression profiles of laser-captured adjacent neuronal subtypes)”, Nat. Med. (1999) 117-122 页)。

使 cDNA 样品与固定于固体支持物上的探针相接触,以便使 cDNA 样品与所述探针杂交。尽管杂交反应溶液和反应条件根据不同的因素,如探针的长度而有所不同,但本领域的技术人员可以通过常规的方法来确定。

cDNA 样品和固体支持物上的探针之间的杂交强度,可根据 cDNA 样品的标记物种类进行确定。例如可通过探测器读出荧光信号来检测荧光标记物。

用不同的荧光标记对实验 cDNA 样品和对照 cDNA 样品(例如来自非癌组织或细胞的 cDNA)进行标记即可同时测定它们的杂交强度。例如上述的一个 cDNA 样品用 Cy5 标记,另一个用 Cy3 标记。Cy5 和 Cy3 的荧光信号强度说明各 cDNA 样品的表达水平 (Duggan 等, Nat. Genet. 21:10-14, 1999)。

在该方法中也可以用 cRNA 而不是 cDNA 进行测定。

此外,可利用针对 *DDEFL1*, *VANGL1*, 或 *LGN* 多肽的抗体,通过下述方法测定多肽的表达水平,所述方法例如 SDS 聚丙烯酰胺电泳法, Western 印迹,点印迹,免疫分析,例如免疫沉淀,荧光免疫分析,放射免疫分析,酶免疫分析(例如酶联免疫吸附试验(ELISA)),和免疫组化染色等等。

在具体实施方案中,将生物样品与固定于固体支持物上的针对 *DDEFL1*, *VANGL1* 或 *LGN* 多肽的抗体相接触,再将所述固体支持物上的抗原-抗体复合物与带有可检测的标记的第二抗体相接触,用适当的方法检测所述标记物。

用于本发明的检测方法的抗体包括任何可与 *DDEFL1*, *VANGL1*, 或 *LGN* 多肽,特别是与具有 SEQ ID NO:2, NO:4 或 NO:6 的氨基酸序列的多肽结合的抗体,包括用 *DDEFL1*, *VANGL1*, 或 *LGN* 多肽免疫动物(如兔)

所得的抗血清,各种类别的多克隆和单克隆抗体,由遗传工程制备的人源化抗体,和人抗体。这些抗体的制备如上所述。

将如上述测定的表达水平与在非癌样品中测定的表达水平相比较,确定受试者是否存在肝细胞癌。当在受试者样品中测得的表达水平高于非癌样品中的表达水平时,则可判断受试者患有癌症或具有患所述癌症的危险性。另一方面,当在受试者样品中测得的表达水平不高于非癌样品中的表达水平时,则可判断受试者未患有癌症。具体地说,受试样品是否比非癌样品的表达水平高可根据相对表达比率(受试样品/非癌样品)确定;当所述相对表达比率高于2.0,可以断定其表达水平较高。

10 检测试剂

本发明提供针对肝细胞癌的检测试剂。

在一个实施方案中,本发明的检测试剂包括具有可与 *DDEFL1*, *VANGL1* 或 *LGN*, 特别是 SEQ ID NO:1, NO:3, 或 NO:5 杂交的至少15个核苷酸的多核苷酸。多核苷酸可作为探针或引物用于上述本发明的检测方法。用作探针时,包含于本发明的检测试剂中的多核苷酸可以是经过标记的。标记的方法包括,例如通过 T4 多核苷酸激酶用 ^{32}P 使所述多核苷酸的 5'-末端磷酸化的标记方法;以随机的六聚物寡核苷酸等作为引物,通过 DNA 聚合酶例如 Klenow 酶引入带有同位素(例如 ^{32}P), 荧光染料, 生物素等标记的底物碱基的方法(随机引导法等等)。

20 在另一个实施方案中,本发明的检测试剂包含可与 *DDEFL1*, *VANGL1* 或 *LGN* 多肽特别是具有 SEQ ID NO:2, NO:4, 或 NO:6 的氨基酸序列的多肽相结合的抗体。本发明的抗体用于在上述本发明的检测方法中检测本发明的多肽。所述抗体可用公知的方法进行标记。此外,所述抗体可被固定到固体支持物上。

25 本发明的检测试剂还可以包含介质或添加剂,包括无菌水,生理盐水,植物油,表面活性剂,脂类,增溶剂,缓冲液,蛋白稳定剂(如牛血清白蛋白和明胶),保存剂等等,只要其不影响本发明检测方法中的反应即可。

抑制肝细胞癌生长的方法

30 本发明还提供抑制肝细胞癌生长的方法。在一个具体实施方案中,该方法可通过将 *DDEFL1*, *VANGL1*, 或 *LGN* 的反义寡核苷酸引入靶细胞来进行。

用于该方法的反义寡核苷酸可与 SEQ ID NO : 1, NO :3, 或 NO :5 的核苷酸序列内的任意位点杂交。所述反义寡核苷酸不仅包括那些全部的核苷酸均与相应的 DNA 或 mRNA 特异区域的核苷酸互补的寡核苷酸, 还包括具有一个或多个错配的核苷酸的寡核苷酸, 只要 DNA 或 mRNA 和寡核苷酸能与 SEQ ID NO : 1 或 NO :3 或 NO :5 的核苷酸序列杂交即可。

所述反义寡核苷酸优选针对 SEQ ID NO : 1, NO :3, 或 NO :5 中至少 15 个连续的核苷酸。更优选上述的反义寡核苷酸包含上述至少 15 连续的核苷酸中的起始密码子。

所述反义寡核苷酸包括含低级烷基磷酸酯(例如, 甲基磷酸酯或乙基磷酸酯), 硫代磷酸酯和磷酰胺酸酯的类似物。

此处, 所述靶细胞可以是哺乳动物细胞, 优选人细胞。

所述引入方法可以是体外, 体内, 或回体(*ex vivo*)的转移方法。在一个实施方案中, 所述反义寡核苷酸可以通过常规的转染方法引入靶细胞。或者, 所述引入可通过常规的基因转移技术, 利用带有反义寡核苷酸的载体, 如腺病毒载体, 反转录病毒载体, 或阳离子的脂质体进行。

在此引用的任何专利, 专利申请, 和出版物均并入本申请作为参考。

附图说明

图 1a-1b 说明 HCC 中称作 B9362 的基因的表达。图 1a 说明由 cDNA 微阵列检测的 B9362 在原始的 20 个 HCC 中的相对表达比率(癌/非-癌)。图 1b 显示对附加的 11 个 HCC 病例通过半定量 RT-PCR 分析的 B9362 表达情况的照片。GAPDH 的表达作为内部对照。

图 2a-2d 说明 DDEFL1 的鉴定结果。图 2a 是 DDEFL1 在各种人组织中的 Northern 印迹分析结果。图 2b 显示 DDEFL1 的结构。图 2c 显示预期的 DDEFL1 蛋白和 ArfGAP 家族成员之间的相似性。图 2d 显示 DDEFL1 和 DDEF2 中 ArfGAP 基元氨基酸序列之间的同一性。箭头指示 CXXCX₁₆CXXC 基元, 表示对 GAP 活性而言必不可少的锌指结构。

图 3a-3b 显示 DDEFL1 的亚细胞定位。图 3a 是 Western 印迹分析结果的照片, 表明 cMyc-标记的 DDEFL1 蛋白在经 pcDNA-DDEFL1-myc 质粒转染的 COS7 细胞中表达。图 3b 是显示所述细胞的免疫细胞化学的照片, 表明 cMyc-标记的 DDEFL1 蛋白定位于细胞质。

图 4a-4d 显示 *DDEFL1* 的生长-促进效果。图 4a 是显示集落形成试验结果的照片, 表明 *DDEFL1* 促进 NIH3T3, SNU423, 和 Alexander 细胞生长。图 4b 是表明 NIH3T3-*DDEFL1* 细胞稳定表达外源 *DDEFL1* 的照片。图 4c 是显示稳定表达外源 *DDEFL1* 的 NIH3T3-*DDEFL1* 细胞在含 10% FBS 的培养基中的生长。图 4d 是显示 NIH3T3-*DDEFL1* 细胞在含 0.1% FBS 的培养基生长的图表($p < 0.01$)。

图 5a-5b 显示经设计能在 SNU475 细胞中抑制 *DDEFL1* 的反义 S-寡核苷酸的生长抑制作用。图 5a 显示反义 S-寡核苷酸的设计和显示通过转染 AS1 或 AS5 反义 S-寡核苷酸使 *DDEFL1* 表达降低的照片。图 5b 是显示 AS1 和 AS5 抑制 SNU475 细胞生长的照片。

图 6a-6b 显示 *VANGL1* 在 HCC 中的表达。图 6a 显示由 cDNA 微阵列检测的 *VANGL1* 在原始的 20 个 HCC 中的相对表达比率(癌/非-癌)。图 6b 是显示用附加的 10 个 HCC 病例由半定量 RT-PCR 分析的 *D3244* 表达情况的照片。T, 肿瘤组织; N, 正常组织。*GAPDH* 的表达作为内部对照。

图 7a 和 7b 说明 *VANGL1* 的鉴定结果。图 7a 是显示 *VANGL1* 在各种人组织中多-组织 Northern 印迹分析结果的照片。图 7b 显示预期的 *VANGL1* 蛋白质结构。

图 8a 和 8b 显示 *VANGL1* 的亚细胞定位。图 8a 是用 pcDNA3.1-myc/His-*VANGL1* 转染, 用小鼠抗-myc 单克隆抗体染色, 用偶联了若丹明的抗小鼠 IgG 第二抗体显现的 SNU475 细胞的照片。核是 DAPI 复染的。图 8b 是经类似染色并观察的模拟细胞的照片。

图 9a-9d 显示经设计能抑制 *VANGL1* 的反义 S-寡核苷酸的生长抑制效果。图 9a 是显示 *VANGL1* 在用对照或反义寡核苷酸处理 12 小时的 SNU475 细胞中表达的照片。图 9b 是显示 S 寡核苷酸抑制 SNU423 细胞生长的照片。图 9c 是由 MTT 试验分析细胞活力的结果图。图 9d 是显示经正义或反义寡核苷酸处理的细胞的荧光活化细胞分选(FACS)结果。

图 10a 和 10b 显示与相应的非癌活组织相比 HCC 的 *LGN* 基因表达。图 10a 显示由 cDNA 微阵列检测的 *LGN* 在原始的 20 个 HCC 中的相对表达比率(癌/非-癌)。图 10b 显示用附加的 10 个 HCC 病例由半定量 RT-PCR 分析的 *LGN* 表达情况的照片。*GAPDH* 的表达作为内部对照。T, 肿瘤组织; N, 正常组织。

图 11 显示 *LGN* 的基因组结构。

图 12a-12c 显示 *LGN* 的亚细胞定位。图 12a 是用 pcDNA3.1-myc/His-*LGN* 转染的 COS7 细胞的图片，其中核是 DAPI 复染的。图 12b 是用 pcDNA3.1 myc/His-*LGN* 转染，用小鼠抗 c-myc 抗体染色，用偶联了若丹明的抗小鼠 IgG 5 第二抗体显示的 COS7 细胞。图 12c 是 a 和 b 的合并。

图 13a 和 13b 显示 *LGN* 的生长-促进效果。图 13a 是显示集落形成试验结果的图片，表明 *LGN* 促进 NIH3T3, SNU423, Alexander 和 SNU475 细胞的生长。图 13b 是显示在含 10% FBS 的培养基中稳定表达外源 *LGN* 的 NIH3T3-*LGN* 细胞的生长优于模拟(NIH3T3-LacZ)细胞。

10 图 14a 和 14b 显示经设计能抑制 *LGN* 表达的反义 S-寡核苷酸在人肝细胞癌 SNU423 细胞中的生长抑制。图 14a 的图片显示，反义 S-寡核苷酸反义 3 的转染降低了 *LGN* 的表达。图 4b 的图片显示，反义 3 抑制了 SNU423 细胞生长。

本发明的最佳实施方式

15 参照下述实施例对本发明进行阐述，但这些实施例不视为对本发明的限制。

实施例 1

1-1. 鉴定通常在人肝细胞癌中上调的 *DDEF1*

20 通过包含 23040 个基因的基因组范围的 cDNA 微阵列方法，将 20 例肝细胞癌(HCC)的表达分布图分别与其相应的非癌肝组织进行了比较。所有的 HCC 组织和相应的非癌组织都是经肝切除术病人同意由其外科手术样品中获得的。发现具有本国(in-house)登录号 B9362 的基因(相应于 EST, UniGene 簇的 Hs.44579)在 1.57 到 5.83 的范围内过表达(图 1a)。在 12 例 HCC 中的 11 例中均观察到其上调的表达(Cy3: Cy5 强度比 > 2.0)，所述 12 例是可以穿
25 过截留滤光片的病例(cy3 和 Cy5 信号都大于 25,000)。由于该基因的开放阅读框编码的蛋白与发育分化增强因子 2(*DDEF2*)具有约 60% 同一性，本发明人将该基因称作发育分化增强因子样 1(*DDEF1*)。为了阐明 cDNA 的微阵列结果，通过半定量 RT-PCR 检测了另外 11 例 HCC 中该转录本的表达。*GAPDH* 的表达作为内部对照。RT-PCR 是如下进行的。用 Qiagen RNeasy
30 试剂盒(Qiagen)或 Trizol 试剂(Life Technologies, Inc.)根据产品说明抽提总 RNA。用聚 dT₁₂₋₁₈ 引物(Amersham Pharmacia Biotech)和 Superscript II 反转录

酶(Life Technologies)将 10 微克小份的总 RNA 反转录成单链 cDNA。稀释单链 cDNA 制剂,以备随后在 20 μ l PCR 缓冲液(TAKARA)中,通过标准 RT-PCR 实验进行 PCR 扩增。扩增反应是在 GeneAmp PCR system 9700 (Perkin-Elmer, Foster City, CA)中如下进行的: 94 $^{\circ}$ C 变性 4 min, 然后以 94 $^{\circ}$ C 30 s, 56 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 45 s 进行 20 (对于 *GAPDH*)或 33 (对于 *DDEFL1*)个循环。*GAPDH* 引物序列为: 正向 5'-ACAACAGCCTCAAGATCATCAG (SEQ ID NO: 7), 反向, 5'-GGTCCACCACTGACACGTTG (SEQ ID NO: 8); *DDEFL1* 引物序列为: 正向 5'-AGCTGAGACATTTGTTCTCTTG (SEQ ID NO: 9), 反向: 5'-TATAAACCAGCTGAGTCCAGAG (SEQ ID NO: 10)。结果证实 *DDEFL1* 的表达在九例肿瘤增加(图 1b)。

1-2. 新基因 *DDEFL1* 的分离和结构

DDEFL1 的表达是利用 *DDEFL1* 的 PCR 产物为探针, 通过多-组织 northern-印迹分析检测的。人多-组织印迹(Clontech, Palo Alto)与 32 P 标记的 *DDEFL1* cDNA 杂交。依照供应商的说明进行预杂交, 杂交和洗涤。所述印迹用增感屏在 -80 $^{\circ}$ C 下放射自显影 72 h。结果表明 4-kb 的转录本在肺, 肝, 小肠胎盘和外周血白细胞中表达(图 2a)。

由于 B9362 比在 Northern 印迹中检测出的小, 通过 5'RACE 实验确定该基因的全部编码序列。5'RACE 实验是利用 Marathon cDNA 扩增试剂盒(Clontech, Palo Alto, CA)依照制造商的说明进行的。利用试剂盒中提供的基因-特异的反向引物(5'-CTCACTTGGCACGTCAGCAGGG (SEQ ID NO: 11)) 和 AP-1 引物扩增 *DDEFL1* cDNA 的 5'部分。cDNA 模板是由人肝 mRNA 合成的。用 TA 克隆试剂盒(Invitrogen)克隆 PCR 产物, 用 ABI PRISM 3700 DNA 测序仪(Applied Biosystems)确定其序列。

完整的 cDNA 由 4050 个核苷酸组成, 具有一个 2712 个核苷酸的开放阅读框, 它编码具有 903 个氨基酸的蛋白(Genbank 登录号 AB051853)。第一个 ATG 的侧翼序列(CCCGCCAATGC (SEQ ID NO: 12))与在真核生物中起始翻译的共有序列相吻合, 并在上游带有一个框内终止密码子。在 NCBI (国家生物信息中心)数据库中用 BLAST 程序检索, 鉴定出一个已分配到染色体带 1p36.12 上的基因组序列(Genbank 登录号 AL357134)。该 cDNA 与基因组序列的比较显示 *DDEFL1* 由 25 个外显子组成(图 2b)。

用简单模块结构检索工具(Simple Modular Architecture Research Tool, SMART)检索蛋白基元, 结果显示预期的蛋白包含两个卷曲螺旋区(密码子 141-172 和 241-278), 一个 PH (Pleckstrin homology)基元(密码子 303-396), 一个 ArfGAP 基元(Arf 的 GTP 酶激活蛋白)(密码子 426-551), 和两个锚蛋白重
5 复(密码子 585-617 和 621-653)。这一结构与矢车菊根皮素 β 1 和矢车菊根皮素 β 2(图 2c)类似。特别是 *DDEFL1* 具有矢车菊根皮素-家族蛋白的特征, 如 PH 结构域, 磷脂酰肌醇 3, 4, 5-三磷酸的靶子, 和 ArfGAP 基元。 *DDEFL1* 的 ArfGAP 基元的氨基酸序列与 *DDEF2* 中相应部分具有 67.8% 的同源性。显然, 代表对 GAP 活性必不可少的锌指结构的 CXXCX₁₆CXXC 基元是完全
10 保守的。

1-3. *DDEFL1* 的亚细胞定位

将 *DDEFL1* 的编码序列克隆到 pcDNA3.1-myc/His 载体(invitrogen)中。将所得的表达 myc-标记型 *DDEFL1* 蛋白的质粒(pDNA-myc/His-*DDEFL1*) 瞬时转染 COS7 细胞(美国典型培养物保藏中心(ATCC))。通过免疫印迹
15 (Western 印迹)检测预期的 myc-标记蛋白如下。将用 pcDNA3.1-myc/His-*DDEFL1* 转染的细胞用 PBS 洗两次, 再用裂解缓冲液 (150 mM NaCl, 1% Triton X-100, 50 mM Tris-HCl pH7.4, 1mM DTT 和 1X 完全蛋白酶抑制剂混合物(Boehringer))收集。在细胞匀浆后, 10,000xg 离心 30 min, 使上清液标准化以利用 Bradford 试验(Bio-Rad)测定蛋白浓度。用 10%
20 SDS-PAGE 分离蛋白, 用小鼠抗 myc 抗体进行免疫印迹。HRP-偶联的山羊抗小鼠 IgG (Amersham)作为第二抗体用于 ECL 检测系统(Amersham)。结果, 用抗 myc 抗体在 western 印迹中检测到了 *DDEFL1* 蛋白(图 3a)。

此外, 免疫细胞化学染色是如下进行的。首先, 将所述细胞用含 4% 多聚甲醛的 PBS 固定 15 min, 然后在 RT 中, 用含 0.1% Triton X-100 的 PBS
25 透析 2.5 min。然后在 4°C 将所述细胞用含 2% BSA 的 PBS 覆盖 24 h 以封闭非特异性杂交。将小鼠抗 myc 单克隆抗体(Sigma)以 1:1000 稀释, 或将小鼠抗 FLAG 抗体(Sigma)以 1:2000 稀释用作第一抗体, 与若丹明偶联型抗-小鼠第二抗体(Leinco 和 ICN)一起孵化, 观察所述反应。核用 4', 6'-联脒-2'-苯基吲哚二氢氯化物(DAPI)复染。在 ECLIPSE E800 显微镜下获得荧光图像。显
30 微分析表明该蛋白主要存在于细胞质中(图 3b)。 *DDEFL1* 也定位在人胚肾 (HEK293)细胞(ATCC)的细胞质中。

1-4. *DDEFL1* 对细胞生长的影响

将 *DDEFL1* 的编码序列克隆到 pcDNA 3.1 载体(invitrogen)中。将涂布于 10-cm 平皿(2×10^5 细胞/皿)的 NIH3T3 细胞(ATCC), 用表达 *DDEFL1* 的质粒(pcDNA *DDEFL1*)和对照质粒(pcDNA-LacZ)转染, 在添加了 10% 胎牛血清和 1% 抗生素/抗霉菌溶液(Sigma)的 Dulbecco 改良 Eagle's 培养基中培养, 并进一步用适当浓度的遗传霉素培养两周。细胞用 100% 甲醇固定, 再用 Giemsa 溶液染色。用 pcDNA-*DDEFL1* 转染的细胞比对照细胞明显产生了更多克隆。用人肝细胞癌 SNU423 (韩国细胞系库)和 Alexander (ATCC)细胞转染也出现类似的集落形成增多现象, 其中 *DDEFL1* 内源表达极低(图 4a)。

为了进一步研究这一生长-促进效应, 建立了稳定表达外源 *DDEFL1* 的 NIH3T3 细胞。利用 FuGENE6 试剂(Boehringer)按照供应商的说明将 pDNA-myc/His-*DDEFL1* 转染到 NIH3T3 细胞中。转染 24 小时后, 向培养基中添加遗传霉素, 转染两周后选出单个集落。用半定量 RT-PCR 确定 *DDEFL1* 的表达(图 4b)。在含 10% FBS 的培养基中, NIH3T3-*DDEFL1* 细胞的生长速率在统计学意义上高于模拟细胞(NIH3T3-*LacZ*) ($p < 0.05$) (图 4c)。在仅含 0.1% FBS 的培养基中, NIH3T3-*DDEFL1* 细胞存活 6 天, 而在相同条件下, 对照 NIH3T3 细胞在 6 天内死亡。在该情况下, NIH3T3-*DDEFL1* 细胞在含 0.1% FBS 的培养基中的生长速率在统计学意义上高于模拟细胞(图 4d) ($P < 0.01$)。

1-5. 在人细胞肝癌 SNU475 细胞中反义 S-寡核苷酸对 *DDEFL1* 表达的抑制

合成下列六对相应于 *DDEFL1* 基因的对照(正义)和反义 S-寡核苷酸。

反义:

- | | | |
|----|------------|---|
| 25 | DDEFL1-AS1 | 5'-TGCTCCGGCATGGCGG-3' (SEQ ID NO: 13); |
| | DDEFL1-AS2 | 5'-GCTGAACTGCTCCGGC-3' (SEQ ID NO: 14); |
| | DDEFL1-AS3 | 5'-TCCAAGATCTCCTCCC-3' (SEQ ID NO: 15); |
| | DDEFL1-AS4 | 5'-TCTCCTTCCAAGATCT-3' (SEQ ID NO: 16); |
| | DDEFL1-AS5 | 5'-GCGCTGAGCCGGCCTC-3' (SEQ ID NO: 17); 和 |
| 30 | DDEFL1-AS6 | 5'-CCTCACCTCCTCCCGC-3' (SEQ ID NO: 18). |

对照:

- DDEFL1-S1 5'-CCGCCATGCCGGAGCA-3' (SEQ ID NO: 19);
 DDEFL1-S2 5'-GCCGGAGCAGTTCAGC-3' (SEQ ID NO: 20);
 DDEFL1-S3 5'-GGGAGGAGATCTTGGA-3' (SEQ ID NO: 21);
 DDEFL1-S4 5'-AGATCTTGGAAGGAGA-3' (SEQ ID NO: 22);
 5 DDEFL1-S5 5'-GAGGCCGGCTCAGCGC-3' (SEQ ID NO: 23); 和
 DDEFL1-S6 5'-GCGGGAGGAGGTGAGG-3' (SEQ ID NO: 24)。

利用 LIPOFECTIN 试剂(GIBCO BRL)将合成的 S-寡核苷酸转染到 SNU475 细胞(韩国细胞系库), 在我们检测的六例肝细胞癌细胞系中, 上述细胞具有最高的 *DDEFL1* 表达水平(未给出数据)。转染 12 和 24 小时后, 与相应的对照 S-寡核苷酸 S1 和 S5 相比, 反义 S-寡核苷酸 AS1 和 AS5 显著地抑制 *DDEFL1* 的表达(图 5a)。转染 6 天后, 用反义 S-寡核苷酸 AS1 或 AS5 转染的存活细胞明显少于用各自的对照 S-寡核苷酸 S1 或 S5 转染的细胞(图 5b)。在三次独立实验均得到一致的结果。

实施例 2

15 2-1. 鉴定通常在人肝细胞癌中上调的 *VANGL1*

实施例 1 的基因组范围内 cDNA 微阵列分析还揭示出另一基因, 它相应于 UniGene 簇 EST (Hs.122730)的本国登录号 D3244。该基因在 12 例临床 HCC 的 10 例中与相应的非癌肝组织对照相比明显上调。这 12 例肿瘤与相应非癌组织相比的相对表达率在 1.5 到 16.0 的范围(图 6a)。在 12 例 HCC 的 20 10 例中均观察到上调的表达(Cy3: Cy5 强度比 > 2.0), 这 12 例病例是可以穿过截留滤光片的病例(cy3 和 Cy5 信号都大于 25,000)。在另外 10 例 HCC 中, 通过用类似于实施例 1-1 的半定量 RT-PCR 证实 D3244 表达的提高, 所用引物对为, 正向 5'- GAGTTGTATTATGAAGAGGCCGA (SEQ ID NO: 25); 反向: 5'- ATGTCTCAGACTGTAAGCGAAGG (SEQ ID NO: 26) (图 6b)。

25 2-2. *VANGL1* 在成人组织中的表达

按照实施例 1-2 的方式, 以 D3244 cDNA 为探针进行多组织 northern 印迹分析, 结果表明 1.9-kb 的转化子以组织-特异性方式在睾丸和卵巢中大量表达(图 7a)。在 NCBI 数据库检索相应于 D3244 的基因组序列, 发现了被分配到染色体带 1p22 的两个序列(Genbank 登录号: AL450389 和 AL592436)。
 30 利用 GENSCAN 和 Gene Recognition 以及 Assembly Internet Link 程序, 预测出候选-外显子序列并进行外显子连接。此外, 按照实施例 1-2 的方式, 用

反向引物((5'-TGTCAGCTCTCCGCTTGCGGAAAAAAG (SEQ ID NO: 27))进行 5'RACE, 以便确定转录本 5'区域的序列。结果获得了 1879 个核苷酸的组装人 cDNA 序列, 其包含一个 1572 个核苷酸的开放阅读框(Genbank 登录号: AB057596)。预期的氨基酸序列与斜视(Van Gogh)和 VANGL2 分别具有 40% 和 63% 的同一性。因此, 将相应于 D3244 的基因称作 Van Gogh 样 1 (VANGL1)。简单模块结构研究工具提示, 预期的蛋白包含公认四个跨膜结构域(密码子 111-133, 148-170, 182-204, 219-241)(图 7b)。

2-3. VANGL1 的亚细胞定位

将表达 c-myc-标记型 VANGL1 蛋白的 pcDNA3.1-myc/His-VANGL1 质粒瞬间转染到 SNU475 细胞(韩国细胞系库)中。按照实施例 1-3 的方式进行免疫细胞化学染色。结果表明所述经标记的 VANGL1 蛋白存在于细胞质中(图 8a 和 8b)。

2-4. 用经设计能降低 VANGL1 表达的反义 S-寡核苷酸抑制肝细胞癌细胞的生长

为测试 VANGL1 的抑制是否可以导致 HCC 细胞的细胞周期阻滞和/或死亡, 合成了下述四对反义和对照(正义) S-寡核苷酸, 并将其转染到 SNU475 细胞中。

反义:

- 反义 1 5'-GTATCCATAGCAATGG-3' (SEQ ID NO: 28);
 反义 2 5'-TGGATTGGGTATCCAT-3' (SEQ ID NO: 29);
 反义 3 5'-TAAGTGGATTGGGTAT-3' (SEQ ID NO: 30); 和
 反义 4 5'-ACTCCTACCTGCCTGT-3' (SEQ ID NO: 31).

对照:

- 正义 1 5'-CCATTGCTATGGATAC-3' (SEQ ID NO: 32);
 正义 2 5'-ATGGATACCCAATCCA-3' (SEQ ID NO: 33);
 正义 3 5'-ATACCCAATCCACTTA-3' (SEQ ID NO: 34); 和
 正义 4 5'-ACAGGCAGGTAGGAGT-3' (SEQ ID NO: 35).

含起始密码子的反义 S-寡核苷酸(反义 3)在 SNU475 细胞中显著地降低 VANGL1 内源表达(图 9a)。

用 3-(4,5-二甲基-噻唑-2-基)-2,5-二苯基四唑溴化物(MTT)试验评价细胞活力如下。将细胞以 5×10^5 细胞/100 mm 皿的浓度涂平皿。接种 24 小时

后将细胞用经设计能抑制 *VANGL1* 的正义或反义 S 寡核苷酸进行一式三份地转染。转染 72 小时后, 将所用培养基替换为含 500 $\mu\text{g/ml}$ 3-(4,5-二甲基-噻唑-2-基)-2,5-二苯基四唑溴化物(MTT) (Sigma)的新鲜培养基, 在 37°C 继续培养 4 小时。然后加入 1 ml 0.01N HCl/10% SDS 以便裂解细胞, 并用 ELISA 5 读板仪在 570 nm 测试波长(参考波长, 630 nm)测定裂解物的吸光度。用与对照细胞相比所得的相对吸光度表示所述细胞的活力。

与对照的正义 S 寡核苷酸正义 3 相比, 反义 S-寡核苷酸反义 3 的转染显著地减少存活细胞的数量(图 9b 和 9c)。这一结果在三次独立实验中得到确认。

10 此外进行了如下的流式细胞术分析。将细胞以 1×10^5 细胞/100 mm 平皿的浓度涂皿, 在给定的时间内用胰蛋白酶处理, 然后用 70% 冷乙醇固定。经 RNase 处理后, 将细胞用碘化丙啶(propidium iodide)(50 $\mu\text{g/ml}$)的 PBS 溶液染色。在 Becton Dickinson FACScan 上进行流式细胞术分析, 然后用 CellQuest 和 ModFit 软件(Verity Software House)进行分析。在至少 20,000 个 15 未经选通(ungated)的细胞中, 确定了核在细胞周期的 G0/G1, S 和 G2/M 期和任意的亚 G1 群中的百分比。

FACS 分析表明, *VANGL1* 抑制显著增加了亚 G1 期细胞的数目(图 9d)。这些结果提示, *VANGL1* 可以对肝细胞癌的细胞生长和/或存活起重要作用。

实施例 3

20 3-1. 人肝细胞癌中 *LGN* 通常增加

通过实施例 1-1 的微阵列分析还选出另一个通常上调的基因 D3636, 它相应于 *LGN*(Genbank 登录号: U54999), 且与相应的非癌活组织相比在 12 例临床 HCC 的 10 例中显著上调。这 12 例肿瘤相对于非癌组织的相对表达式比率在 0.7 到 16.0 的范围。在 12 例 HCC 的 10 例中均观察到 *LGN* 的已上调的表达(Cy3:Cy5 强度比 >2.0), 这些病例是可以穿过截留滤光片的病例(cy3 和 Cy5 信号都大于 25,000) (图 10a)。在另外 10 例 HCC 中, 通过用类似于 25 实施例 1-1 的半定量 RT-PCR 证实 *LGN* 表达也存在上调, 所用引物对为, 正向 5'-ATCTGAAGCACTTAGCAATTGC (SEQ ID NO: 36), 反向: 5'-CTGTAGCTCAGACCAAGAACC (SEQ ID NO: 37) (图 10b)。

30 3-2. *LGN* 的基因组结构

LGN 的 cDNA 由 2,336 个核苷酸组成, 编码 677 个氨基酸的肽。将该 cDNA 序列与基因组 cDNA 序列比较, 表明 *LGN* 基因由 14 个外显子组成(图 11)。

3-3. *LGN* 的亚细胞定位

- 5 用表达 c-myc-标记型 *LGN* 蛋白的 pcDNA3.1-myc/His-*LGN* 质粒瞬间转染 COS7 细胞。按照实施例 1-3 的方式进行免疫印迹分析, 检测到相应于 myc-标记型 *LGN* 蛋白的 72 kDa 带(图 12)。类似地, 按照实施例 1-3 进行免疫细胞化学染色, 结果显示标记的 *LGN* 蛋白存在于细胞质和细胞核中。

3-4. *LGN* 基因转移可以促进细胞生长

- 10 为分析 *LGN* 对细胞生长的作用, 如实施例 1-4 一样用表达 *LGN* 的质粒 (pcDNA3.1-myc/His-*LGN*) 转染 NIH3T3, SNU423, Alexander 和 SNU475 细胞进行集落形成试验。与对照质粒 (pcDNA3.1-myc/His-*LacZ*) 相比, pcDNA3.1-myc/His-*LGN* 在这些细胞中产生明显更多的集落(图 13a)。这一结果在三次独立实验中得到确认。

- 15 为进一步研究 *LGN* 对细胞生长的作用, 建立了稳定表达外源 *LGN* 的 NIH3T3 细胞(NIH3T3-*LGN* 细胞)。与对照 NIH3T3 *LacZ* 细胞相比, NIH3T3-*LGN* 细胞显示较高的生长速率(图 13b)。

3-5. *LGN* 的反义 S-寡核苷酸抑制人肝细胞癌 SNU475 细胞的生长

- 20 合成下述五对相应于 *LGN* 的对照(正义)和反义 S-寡核苷酸, 然后转染到 SNU423 细胞中。

反义:

- 反义 1 5'-CCATCGAGTCATATTA-3'(SEQ ID NO: 38);
反义 2 5'-TTCCTCCATCGAGTCA-3'(SEQ ID NO: 39);
反义 3 5'-AAATTTTCCTCCATCG-3'(SEQ ID NO: 40);
25 反义 4 5'-AGTCTTACCTGTAACG-3'(SEQ ID NO: 41); 和
反义 5 5'-GCTTCCATTCTACAAA-3'(SEQ ID NO: 42).

正义:

- 正义 1 5'-TAATATGACTCGATGG-3'(SEQ ID NO: 43);
正义 2 5'-TGACTCGATGGAGGAA-3'(SEQ ID NO: 44);
30 正义 3 5'-CGATGGAGGAAAATTT-3'(SEQ ID NO: 45);
正义 4 5'-CGTTACAGGTAAGACT-3'(SEQ ID NO: 46); 和

正义 5 5'-TTTGTAGAATGGAAGC-3'(SEQ ID NO: 47).

转染 12 小时后, 与对照 S 寡核苷酸(正义 3)相比, 跨起始密码子的反义 S-寡核苷酸(反义 3)显著地抑制 *LGN* 的表达(图 14a)。转染 6 天后, 经反义 3 转染的存活细胞的数目显著地少于经对照正义 3 转染的存活细胞的数目(图 14b)。在三次独立实验中均得到一致的结果。

工业实用性

本发明提供 *DDEFL1*, *VANGL1* 或 *LGN* 的 cDNA 核苷酸序列和多肽氨基酸序列, 已发现这些基因在肝细胞癌中通常上调。因此这些多肽可以作为标记用于确定是否存在肝癌。这些核苷酸序列信息使得能设计探针和引物来检测或扩增 *DDEFL1*, *VANGL1* 或 *LGN* 基因。还使得能合成抑制 *DDEFL1*, *VANGL1* 或 *LGN* 多肽表达的反义核苷酸序列。氨基酸序列信息使得能制备与 *DDEFL1*, *VANGL1* 或 *LGN* 多肽结合的抗体。所述探针引物以及抗体可用作检测肝细胞癌的试剂。此外, 本发明人证实, 用反义寡核苷酸抑制 *DDEFL1*, *VANGL1* 或 *LGN* 的表达, 能显著地降低 HCC 细胞的生长。由此, 所述反义寡核苷酸可用于抑制 HCC 细胞的生长。本发明还有助于进一步阐明肝细胞的致癌作用机制, 并发现可用于开发肝癌治疗的有效药物的分子靶。

<110> 东京大学总长代表日本国 (JAPAN AS REPRESENTED BY THE PRESIDENT OF THE UNIVERSITY OF TOKYO ONCOTHERAPY SCIENCE, INC.)

<120> 肝细胞癌-相关基因和多肽, 以及检测肝细胞癌的方法

<130> SEN-A0121P

<140>

<141>

<150> US 60/324,261

<151> 2001-09-25

<150> CA

<151> 2002-08-23

<160> 47

<170> PatentIn version 3.1

<210> 1

<211> 4050

<212> DNA

<213> 人 (Homo sapiens)

<400> 1

```

gtgcccccg cgtccgctc cggcagctcc acgctcgcgc ccgcatgcc ggagcagttc      60
agcgtcgccg agttcctggc cgtcaccgcg gaggacctca gctccccggc tggggccgcc      120
gccttcgccg ccaagatgcc ccggtaccga ggggcggcgc tggcgggga ggagatcttg      180
gaaggagacc aagccatcct gcagagaata aagaaggctg tgcgggcaat ccatagctcc      240
ggccttggcc atgiggagaa tgaagagcag taccgagagg ccgtggaatc ctaggcaac      300
agccacctgt ccagaacag ccatgagctg tccacaggct tcctaaactt ggccgtgttc      360
acccgcgagg ttgctgcgct cttcaagaac ctgattcaga acttgaacaa cattgtctct      420
ttccccctgg acagtctgat gaaggggcag ctgagggacg gtcgacagga ttccaaaaaa      480
cagctggaga aggcatggaa ggactatgaa gccaaaatgg ccaagctgga gaaggagcgc      540

```

gatcgggcca	gggtgacagg	agggatccct	ggggaggtgg	cccaggacat	gcagagagag	600
cggcgcatct	tccagctgca	catgtgtgag	tatctgctca	aagccgggga	gagccagatg	660
aagcaaggtc	ctgacttccct	tcagagcctc	atcaagttct	tccacgcccc	gcacaacttt	720
ttccaagatg	gctggaaggc	tgcccagagc	ctgttcccct	tcatcgagaa	gctggcggcc	780
tcagtacatg	cactccatca	ggcccaggag	gacgagctac	agaagctgac	ccagctccgg	840
gactcccicc	gagggacact	gcagcttgag	agcagagagg	aacacctgag	ccggaagaac	900
tcaggatgtg	gctatagcat	ccaccagcac	caaggcaaca	agcagtttgg	gacggagaaa	960
gtgggctttc	tatacaagaa	aagtgacgga	attcgaagag	tctggcagaa	aaggaagtgt	1020
ggagtcaagt	atggctgcct	gaccatctca	cacagcacga	taaaccggcc	cccggatgaag	1080
ctgacctgic	tgacgtgcca	agtgaggcca	aacctgagg	agaaaaagtg	cttcgacctg	1140
gtgaccacaca	accggacgta	ccactttcag	gcagaggacg	agcacgagtg	tgaggcgtgg	1200
gtgtcagtgt	tgagaacag	caaggacgaa	gccctgagca	gcgcccttct	cggggagccc	1260
agcgcitggcc	cggggtcctg	ggggtcggcc	ggccatgatg	gggagccgca	cgacctcaca	1320
aagctgctca	tcgcggaggt	gaagagcagg	cctgggaata	gccagtgtctg	cgactgcggg	1380
gctgcagacc	ccacgtggct	cagcaccaac	ctgggcgtgc	tcacctgcat	ccagtgtctg	1440
ggcgtccacc	gcgaactggg	cgctgccttt	tcgcgatgc	agtcactcac	cttggacctg	1500
ctgggcccct	ccgagttggt	gctggccttg	aacatgggaa	acacgagctt	caatgaggtc	1560
atggaggccc	agctaccctc	acacggcggc	cctaaaccct	cagctgagag	tgacatgggc	1620
acccgcaggg	actacattat	ggccaagtat	gtggagcata	ggtttgcacg	ccggtgcaca	1680
cctgagcctc	agcgactctg	gacagccatt	tgcaacaggg	acctcctgtc	ggtactggag	1740
gcctttgcca	atgggcagga	ctttggacag	ccgctgccag	ggcctgatgc	acaggcacct	1800
gaagaactcg	tcttgcatth	ggctgtcaaa	gtcgccaacc	aggcttccct	gcctctggtg	1860
gatttcaica	tccagaacgg	tggtcacctg	gatgccaagg	ctgctgacgg	gaacacggct	1920

ctgcactacg cagcactcta caaccagccc gactgcctca agctgctgct gaaggggaga 1980
 gctttgggtg gcacagtaaa tgaagcaggc gagacagctc tggacatagc caggaagaag 2040
 caccacaagg agtgtgagga gctgctggag caggcccagg cggggacctt tgccttcctt 2100
 ctacatgtgg actactcctg ggtaatttcc acagagcctg gctctgacag tgaggaggat 2160
 gaggaagaga agcgcgtgctt gctgaagctc ccggcccagg ctactgggc cagtgggagg 2220
 ctggacatca gcaacaagac ctatgagact gtcgccagcc tgggagcagc caccctcag 2280
 ggcgagagtg aggactgtcc cccgcccttg ccagtcaaaa actcttctcg gactttggtc 2340
 caaggggtg caagacatgc cagtggagat cgttctgaag tctccagcct gagttcagag 2400
 gcccctgaga cccttgagag cctgggcagt ccagcctcct cctccagtct gatgagcccc 2460
 ttggaacctg gggatcccag ccaagcccca cccaactctg aagagggcct ccgagagccc 2520
 ccaggcacct ccagaccag cctgacatcc gggaccacc cttcggagat gtacctcccc 2580
 gtcagattca gctccgagag cactcgcctc tatcggcggg gggcgcggag ccctgaagat 2640
 ggtccctcag ccaggcagcc tctgccaga aggaacgtgc cggttggcat cactgaagga 2700
 gatggctcaa ggactgggag tctcccagca agttctgtgc aacttttgca agactagctc 2760
 cttgctggcc cccacatgcc ccatgctagg ccccaatgtt cagagctggg acttgagctc 2820
 aaaaaactgg ggagctgaga catttgttct ctggatctc actctctctg tcccttgtgc 2880
 ctctgtagct ggccttcttc ctgccacagg ccatgcctct accaaggaca catggccttt 2940
 ccctgtagg gctgatggcg gttctttcct atctcattac ccgctagggg cctgggagcc 3000
 ctgtggctgg atctgagtgc tcctgagctg gcttcagctg cagaactctc agtcctcat 3060
 cagatcgaga ctctatttcc cccgtcagtc tggggcttc acaaggcag gagagccctc 3120
 catcactgac ttccagatca gggaccctgc caagtaggga ctgtctctc agccagccat 3180
 ttattagtct aatattcctt cactaaattc caactctatg tctggacctg tgtaggcac 3240
 ttcagatacc acacgagtaa gacaagggcc ctgcagggtt ggtcctttgg tggaaagctg 3300

```

gtcctaaggg ttgggcttgg gaataggcag ggtcagattc cagggcatgg ctctggactc 3360
agcigggtta tacctatatg accattacag ttgtctacag atcacatcca ttctggctgg 3420
tcaacatgca tgctgtactg gctgttaaataaaaaatattc tgaatgtcac tccttttgag 3480
ggacagcaca gccttccta ggcattctcc tatattcccc agccaaattg tagagtcaga 3540
tgcaccacaca ttgacctgtg tccttgattt agcaggaagg aaaggaatag tcgggggtga 3600
tggatgcccc cttctcttct cttctcttgg gtcaactcag gagcctttta gtctgaggga 3660
atggagagggc aaagaaagaa gggagagtaa tagaattggg agggcagaga cttaagggtt 3720
ctgcttcccc gccctagaaa ttctatcatt gctcagcccc aatgagaaag cagatacacc 3780
taagccatca tcaaccacta acatctcaac ttgccagtig ctgggtgctg ggccctggca 3840
ggaatgggcc aagccaagca ggggagacta gagagcacca atggccaaca cagctgcctg 3900
gctggggagg ctgtgctgtt tcccctggag acctgactgg tctgtggttc ccacaggaac 3960
agggttgtct tttagcccc cagtgtctgg ttccattcat ctcagacttg ttatttcact 4020
catctctaataaagattgg ggggtcagtt 4050

```

<210> 2

<211> 903

<212> PRT

<213> 人 (Homo sapiens)

<400> 2

```

Met Pro Glu Gln Phe Ser Val Ala Glu Phe Leu Ala Val Thr Ala Glu
1           5           10           15
Asp Leu Ser Ser Pro Ala Gly Ala Ala Ala Phe Ala Ala Lys Met Pro
           20           25           30
Arg Tyr Arg Gly Ala Ala Leu Ala Arg Glu Glu Ile Leu Glu Gly Asp
           35           40           45
Gln Ala Ile Leu Gln Arg Ile Lys Lys Ala Val Arg Ala Ile His Ser
50           55           60

```

Ser Gly Leu Gly His Val Glu Asn Glu Glu Gln Tyr Arg Glu Ala Val
 65 70 75 80
 Glu Ser Leu Gly Asn Ser His Leu Ser Gln Asn Ser His Glu Leu Ser
 85 90 95
 Thr Gly Phe Leu Asn Leu Ala Val Phe Thr Arg Glu Val Ala Ala Leu
 100 105 110
 Phe Lys Asn Leu Ile Gln Asn Leu Asn Asn Ile Val Ser Phe Pro Leu
 115 120 125
 Asp Ser Leu Met Lys Gly Gln Leu Arg Asp Gly Arg Gln Asp Ser Lys
 130 135 140
 Lys Gln Leu Glu Lys Ala Trp Lys Asp Tyr Glu Ala Lys Met Ala Lys
 145 150 155 160
 Leu Glu Lys Glu Arg Asp Arg Ala Arg Val Thr Gly Gly Ile Pro Gly
 165 170 175
 Glu Val Ala Gln Asp Met Gln Arg Glu Arg Arg Ile Phe Gln Leu His
 180 185 190
 Met Cys Glu Tyr Leu Leu Lys Ala Gly Glu Ser Gln Met Lys Gln Gly
 195 200 205
 Pro Asp Phe Leu Gln Ser Leu Ile Lys Phe Phe His Ala Gln His Asn
 210 215 220
 Phe Phe Gln Asp Gly Trp Lys Ala Ala Gln Ser Leu Phe Pro Phe Ile
 225 230 235 240
 Glu Lys Leu Ala Ala Ser Val His Ala Leu His Gln Ala Gln Glu Asp
 245 250 255
 Glu Leu Gln Lys Leu Thr Gln Leu Arg Asp Ser Leu Arg Gly Thr Leu
 260 265 270
 Gln Leu Glu Ser Arg Glu Glu His Leu Ser Arg Lys Asn Ser Gly Cys
 275 280 285
 Gly Tyr Ser Ile His Gln His Gln Gly Asn Lys Gln Phe Gly Thr Glu
 290 295 300
 Lys Val Gly Phe Leu Tyr Lys Lys Ser Asp Gly Ile Arg Arg Val Trp

Leu Ser Val Leu Glu Ala Phe Ala Asn Gly Gln Asp Phe Gly Gln Pro
 565 570 575

Leu Pro Gly Pro Asp Ala Gln Ala Pro Glu Glu Leu Val Leu His Leu
 580 585 590

Ala Val Lys Val Ala Asn Gln Ala Ser Leu Pro Leu Val Asp Phe Ile
 595 600 605

Ile Gln Asn Gly Gly His Leu Asp Ala Lys Ala Ala Asp Gly Asn Thr
 610 615 620

Ala Leu His Tyr Ala Ala Leu Tyr Asn Gln Pro Asp Cys Leu Lys Leu
 625 630 635 640

Leu Leu Lys Gly Arg Ala Leu Val Gly Thr Val Asn Glu Ala Gly Glu
 645 650 655

Thr Ala Leu Asp Ile Ala Arg Lys Lys His His Lys Glu Cys Glu Glu
 660 665 670

Leu Leu Glu Gln Ala Gln Ala Gly Thr Phe Ala Phe Pro Leu His Val
 675 680 685

Asp Tyr Ser Trp Val Ile Ser Thr Glu Pro Gly Ser Asp Ser Glu Glu
 690 695 700

Asp Glu Glu Glu Lys Arg Cys Leu Leu Lys Leu Pro Ala Gln Ala His
 705 710 715 720

Trp Ala Ser Gly Arg Leu Asp Ile Ser Asn Lys Thr Tyr Glu Thr Val
 725 730 735

Ala Ser Leu Gly Ala Ala Thr Pro Gln Gly Glu Ser Glu Asp Cys Pro
 740 745 750

Pro Pro Leu Pro Val Lys Asn Ser Ser Arg Thr Leu Val Gln Gly Cys
 755 760 765

Ala Arg His Ala Ser Gly Asp Arg Ser Glu Val Ser Ser Leu Ser Ser
 770 775 780

Glu Ala Pro Glu Thr Pro Glu Ser Leu Gly Ser Pro Ala Ser Ser Ser
 785 790 795 800

```

gctacctggg cctcacccgc gcctcttttc ttggacttct agttttcctc acccctattg    600
ccttcacctt ttacctccg atcctgtgga gggatgagct ggagccttgt ggcacaattt    660
gtgaggggct ctttatctcc atggcattca aactcctcat tctgctcata gggacctggg    720
cacttttttt ccgcaagcgg agagctgaca tgccacgggt gtttggtttt cgtgcccttt    780
tgttggctct catctttctc tttgtggttt cctattggct tttttacggg gtccgcattt    840
tggactctcg ggaccggaat taccagggca ttgtgcaata tgcagtctcc ctigtggatg    900
ccctcctctt catccattac ctggccatcg tcttgcctgga gctcaggcag ctgcagccca    960
tgttcacgct gcaggtggtc cgctccaccg atggcgagtc ccgcttctac agcctgggac   1020
acctgagtat ccagcgagca gcattgggtg tccatgaaaa ttactacaaa gatttcacca   1080
tctataacct aaacctccta acagcctcca aattccgagc agccaagcat atggccgggc   1140
tgaaagtcta caatgtagat ggccccagta acaatgccac tggccagtcc cgggccatga   1200
ttgctgcagc tgctcggcgc agggactcaa gccacaacga gttgtattat gaagaggccg   1260
aacatgaacg gcgagtaaag aagcggaaag caaggctggt ggttgcagtg gaagaggcct   1320
tcatccacat tcagcgtctc caggctgagg agcagcagaa agccccaggg gaggtgatgg   1380
accctaggga ggccgccag gccatcttcc cctccatggc cagggtcttc cagaagtacc   1440
tgcgcatcac ccggcagcag aactaccaca gcatggagag catcctgcag cacctggcct   1500
tctgcatcac caacggcatg accccaagg ccttcctaga acggtacctc agtgcgggcc   1560
ccaccctgca atatgacaag gaccgctggc tctctacaca gtggaggctt gtcagtgatg   1620
aggctgtgac taatggatta cgggatggaa ttgtgttcgt ccttaagtgc ttggacttca   1680
gcctcgtagt caatgtgaag aaaattccat tcatcactct ctctgaagag ttcataagacc   1740
ccaaatctca caaatttgc cttcgcttac agtctgagac atccgtttaa aagttctata   1800
tttgtggctt tattaaaaaa aaaagaaaaa tatatagaga gatatgcaaa aaaaataaaa   1860
gacaaaaaca aaaaaaaaaa                                     1879

```

<210> 4
 <211> 524
 <212> PRT
 <213> 人 (Homo sapiens)

<400> 4
 Met Asp Thr Glu Ser Thr Tyr Ser Gly Tyr Ser Tyr Tyr Ser Ser His
 1 5 10 15

 Ser Lys Lys Ser His Arg Gln Gly Glu Arg Thr Arg Glu Arg His Lys
 20 25 30

 Ser Pro Arg Asn Lys Asp Gly Arg Gly Ser Glu Lys Ser Val Thr Ile
 35 40 45

 Gln Pro Pro Thr Gly Glu Pro Leu Leu Gly Asn Asp Ser Thr Arg Thr
 50 55 60

 Glu Glu Val Gln Asp Asp Asn Trp Gly Glu Thr Thr Thr Ala Ile Thr
 65 70 75 80

 Gly Thr Ser Glu His Ser Ile Ser Gln Glu Asp Ile Ala Arg Ile Ser
 85 90 95

 Lys Asp Met Glu Asp Ser Val Gly Leu Asp Cys Lys Arg Tyr Leu Gly
 100 105 110

 Leu Thr Val Ala Ser Phe Leu Gly Leu Leu Val Phe Leu Thr Pro Ile
 115 120 125

 Ala Phe Ile Leu Leu Pro Pro Ile Leu Trp Arg Asp Glu Leu Glu Pro
 130 135 140

 Cys Gly Thr Ile Cys Glu Gly Leu Phe Ile Ser Met Ala Phe Lys Leu
 145 150 155 160

 Leu Ile Leu Leu Ile Gly Thr Trp Ala Leu Phe Phe Arg Lys Arg Arg
 165 170 175

 Ala Asp Met Pro Arg Val Phe Val Phe Arg Ala Leu Leu Leu Val Leu
 180 185 190

 Ile Phe Leu Phe Val Val Ser Tyr Trp Leu Phe Tyr Gly Val Arg Ile
 195 200 205

Leu Asp Ser Arg Asp Arg Asn Tyr Gln Gly Ile Val Gln Tyr Ala Val
 210 215 220
 Ser Leu Val Asp Ala Leu Leu Phe Ile His Tyr Leu Ala Ile Val Leu
 225 230 235 240
 Leu Glu Leu Arg Gln Leu Gln Pro Met Phe Thr Leu Gln Val Val Arg
 245 250 255
 Ser Thr Asp Gly Glu Ser Arg Phe Tyr Ser Leu Gly His Leu Ser Ile
 260 265 270
 Gln Arg Ala Ala Leu Val Val Leu Glu Asn Tyr Tyr Lys Asp Phe Thr
 275 280 285
 Ile Tyr Asn Pro Asn Leu Leu Thr Ala Ser Lys Phe Arg Ala Ala Lys
 290 295 300
 His Met Ala Gly Leu Lys Val Tyr Asn Val Asp Gly Pro Ser Asn Asn
 305 310 315 320
 Ala Thr Gly Gln Ser Arg Ala Met Ile Ala Ala Ala Ala Arg Arg Arg
 325 330 335
 Asp Ser Ser His Asn Glu Leu Tyr Tyr Glu Glu Ala Glu His Glu Arg
 340 345 350
 Arg Val Lys Lys Arg Lys Ala Arg Leu Val Val Ala Val Glu Glu Ala
 355 360 365
 Phe Ile His Ile Gln Arg Leu Gln Ala Glu Glu Gln Gln Lys Ala Pro
 370 375 380
 Gly Glu Val Met Asp Pro Arg Glu Ala Ala Gln Ala Ile Phe Pro Ser
 385 390 395 400
 Met Ala Arg Ala Leu Gln Lys Tyr Leu Arg Ile Thr Arg Gln Gln Asn
 405 410 415
 Tyr His Ser Met Glu Ser Ile Leu Gln His Leu Ala Phe Cys Ile Thr
 420 425 430
 Asn Gly Met Thr Pro Lys Ala Phe Leu Glu Arg Tyr Leu Ser Ala Gly
 435 440 445

ttccagaaga agtgagagat gctctgcagg cagccgtgga ttttatgag gaaaacctat	720
cattagtac tgctttgggt gaccgagcgg cacaaggacg tgcctttgga aatcttggaa	780
acacacatta cctccttggc aacttcaggg atgcagttat agtcatgag cagcgtctcc	840
ttattgcaa agaatttggg gataaagcag ctgaaagaag agcatatagc aaccttggaa	900
atgcatatat atttcttggg gaatttgaaa ctgcctcgga atactacaag aagacactac	960
tgttgcccg acagcttaaa gaccgagctg tagaagcaca gtcttgttac agtcttggaa	1020
atacatatac ttacttcaa gactatgaaa aggccattga ttatcatctg aagcacttag	1080
caattgctca agagctgaat gatagaattg gtgaaggaag agcatgttgg agcttaggaa	1140
atgcatacac agcactagga aatcatgac aagcaatgca ttttgctgaa aagcacttgg	1200
aaatttcaag agaggttggg gataaaagtg gtgaactaac agcacgactt aatctctcag	1260
accttcaaat ggttcttggg ctgagctaca gcacaaataa ctccataatg tctgaaaata	1320
ctgaaattga tagcagtttg aatgggttac tccccagtt gggacgccgg catagtatgg	1380
aaaatatgga acttatgaag ttaacaccag aaaaggtaca gaactggaac agtgaaattc	1440
ttgctaagca aaaacctctt attgccaaac ctctgcaa gctactcttt gtcaacagac	1500
tgaaggggaa aaaatacaaa acgaattcct ccactaaagt tctccaagat gccagtaatt	1560
ctattgacca ccgaattcca aattctcaga ggaaaatcag tgcagatact attggagatg	1620
aagggttctt tgacttatta agccgatttc aaagcaatag gatggatgat cagagatggt	1680
gcttacaaga aaagaactgc catacagctt caacaacaac ttcttccact cccctaaaa	1740
tgatgctaaa aacatcatct gttcctgtgg tatccccaa cacggatgag ttttagatc	1800
ttcttgccag ctacagagt cgccgtctgg atgaccagag ggctagtctc agtaatttgc	1860
cagggtctcg tctaacacaa aacagccagt cggtacttag ccacctgatg actaatgaca	1920
acaaagaggc tgatgaagat ttctttgaca tccttgtaa atgtcaagga tccagattag	1980
atgatcaaag atgtgctcca ccacctgcta ccacaaaggg tccgacagta ccagatgaag	2040

actttttcag ccttatttta cgggccagg gaaagagaat ggatgaacag agagttcttt 2100
 taaaaagaga taaaaacaga gacactgact ttgggctaaa ggactttttg caaaataatg 2160
 ctttgttgga gtttaaaaat tcagggaaaa aatcggcaga ccattagtta ctatggattt 2220
 attttttttc ctttcaaaca cggtaaggaa acaatctatt acttttttcc ttaaaaggag 2280
 aatttatagc actgtaatac agcttaaaat attttagaa tgatgtaaat agttaa 2336

<210> 6

<211> 677

<212> PRT

<213> 人 (Homo sapiens)

<400> 6

Met Arg Glu Asp His Ser Phe His Val Arg Tyr Arg Met Glu Ala Ser
 1 5 10 15

Cys Leu Glu Leu Ala Leu Glu Gly Glu Arg Leu Cys Lys Ser Gly Asp
 20 25 30

Cys Arg Ala Gly Val Ser Phe Phe Glu Ala Ala Val Gln Val Gly Thr
 35 40 45

Glu Asp Leu Lys Thr Leu Ser Ala Ile Tyr Ser Gln Leu Gly Asn Ala
 50 55 60

Tyr Phe Tyr Leu His Asp Tyr Ala Lys Ala Leu Glu Tyr His His His
 65 70 75 80

Asp Leu Thr Leu Ala Arg Thr Ile Gly Asp Gln Leu Gly Glu Ala Lys
 85 90 95

Ala Ser Gly Asn Leu Gly Asn Thr Leu Lys Val Leu Gly Asn Phe Asp
 100 105 110

Glu Ala Ile Val Cys Cys Gln Arg His Leu Asp Ile Ser Arg Glu Leu
 115 120 125

Asn Asp Lys Val Gly Glu Ala Arg Ala Leu Tyr Asn Leu Gly Asn Val
 130 135 140

Tyr His Ala Lys Gly Lys Ser Phe Gly Cys Pro Gly Pro Gln Asp Val

Ser Met Glu Asn Met Glu Leu Met Lys Leu Thr Pro Glu Lys Val Gln
 405 410 415

Asn Trp Asn Ser Glu Ile Leu Ala Lys Gln Lys Pro Leu Ile Ala Lys
 420 425 430

Pro Ser Ala Lys Leu Leu Phe Val Asn Arg Leu Lys Gly Lys Lys Tyr
 435 440 445

Lys Thr Asn Ser Ser Thr Lys Val Leu Gln Asp Ala Ser Asn Ser Ile
 450 455 460

Asp His Arg Ile Pro Asn Ser Gln Arg Lys Ile Ser Ala Asp Thr Ile
 465 470 475 480

Gly Asp Glu Gly Phe Phe Asp Leu Leu Ser Arg Phe Gln Ser Asn Arg
 485 490 495

Met Asp Asp Gln Arg Cys Cys Leu Gln Glu Lys Asn Cys His Thr Ala
 500 505 510

Ser Thr Thr Thr Ser Ser Thr Pro Pro Lys Met Met Leu Lys Thr Ser
 515 520 525

Ser Val Pro Val Val Ser Pro Asn Thr Asp Glu Phe Leu Asp Leu Leu
 530 535 540

Ala Ser Ser Gln Ser Arg Arg Leu Asp Asp Gln Arg Ala Ser Phe Ser
 545 550 555 560

Asn Leu Pro Gly Leu Arg Leu Thr Gln Asn Ser Gln Ser Val Leu Ser
 565 570 575

His Leu Met Thr Asn Asp Asn Lys Glu Ala Asp Glu Asp Phe Phe Asp
 580 585 590

Ile Leu Val Lys Cys Gln Gly Ser Arg Leu Asp Asp Gln Arg Cys Ala
 595 600 605

Pro Pro Pro Ala Thr Thr Lys Gly Pro Thr Val Pro Asp Glu Asp Phe
 610 615 620

Phe Ser Leu Ile Leu Arg Ser Gln Gly Lys Arg Met Asp Glu Gln Arg
 625 630 635 640

Val Leu Leu Gln Arg Asp Gln Asn Arg Asp Thr Asp Phe Gly Leu Lys
 645 650 655

Asp Phe Leu Gln Asn Asn Ala Leu Leu Glu Phe Lys Asn Ser Gly Lys
 660 665 670

Lys Ser Ala Asp His
 675

<210> 7

<211> 22

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的引物序列

<400> 7

acaacagcct caagatcatc ag

22

<210> 8

<211> 20

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的引物序列

<400> 8

gggccaccac tgacacgttg

20

<210> 9

<211> 22

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的引物序列

<400> 9

agctgagaca tttgttctct tg

22

<210> 10
 <211> 22
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的引物序列

<400> 10
 tataaaccag ctgagtccag ag 22

<210> 11
 <211> 22
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的引物序列

<400> 11
 ctcacttggc acgtcagcag gg 22

<210> 12
 <211> 10
 <212> DNA
 <213> 人 (Homo sapiens)

<400> 12
 cccgccatgc 10

<210> 13
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 13
 tgctccggca tggcgg 16

<210> 14

<211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 14
 gctgaactgc tccggc 16

<210> 15
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 15
 tccaagatct cctccc 16

<210> 16
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 16
 tctcctcca agatct 16

<210> 17
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 17
 gcgctgagcc ggcctc 16

<210> 18
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 18
 cctcacctcc tcccgc 16

<210> 19
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 19
 ccgcatgcc ggagca 16

<210> 20
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 20
 gccggagcag ttcagc 16

<210> 21
 <211> 16
 <212> DNA
 <213> 人工的

<220>
 <223> 一个人工合成的寡核苷酸序列

<400> 21

gggaggagat cttgga

16

<210> 22

<211> 16

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的寡核苷酸序列

<400> 22

agatcttgga aggaga

16

<210> 23

<211> 16

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的寡核苷酸序列

<400> 23

gaggccggct cagcgc

16

<210> 24

<211> 16

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的寡核苷酸序列

<400> 24

gcgggaggag gtgagg

16

<210> 25

<211> 23

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的引物序列

<400> 25	
gagttgtatt atgaagaggc cga	23
<210> 26	
<211> 23	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的引物序列	
<400> 26	
atgtctcaga ctgtaagcga agg	23
<210> 27	
<211> 28	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的引物序列	
<400> 27	
tgtcagctct ccgcttgCgg aaaaaaag	28
<210> 28	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 28	
gtatccatag caatgg	16
<210> 29	
<211> 16	
<212> DNA	
<213> 人工的	

<220>

<223> 一个人工合成的寡核苷酸序列

<400> 29

tggattgggt atccat

16

<210> 30

<211> 16

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的寡核苷酸序列

<400> 30

taagtggatt ggtat

16

<210> 31

<211> 16

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的寡核苷酸序列

<400> 31

actcctacct gcctgt

16

<210> 32

<211> 16

<212> DNA

<213> 人工的

<220>

<223> 一个人工合成的寡核苷酸序列

<400> 32

ccattgctat ggatac

16

<210> 33

<211> 16

<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 33	
atggataccc aatcca	16
<210> 34	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 34	
atacccaatc cactta	16
<210> 35	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 35	
acaggcaggt aggagt	16
<210> 36	
<211> 22	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的引物序列	
<400> 36	
atctgaagca cttagcaatt gc	22

<210> 37	
<211> 21	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的引物序列	
<400> 37	
ctgtagctca gaccaagaac c	21
<210> 38	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 38	
ccatcgagtc atatta	16
<210> 39	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 39	
ttcctccatc gagtca	16
<210> 40	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 40	
aaattttcct ccatcg	16

<210> 41	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 41	
agtcttacct gtaacg	16
<210> 42	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 42	
gcttccattc taca aa	16
<210> 43	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 43	
taatatgact cgatgg	16
<210> 44	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	

<400> 44	
tgactcgatg gagga	16
<210> 45	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 45	
cgatggagga aaattt	16
<210> 46	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 46	
cgttacaggt aagact	16
<210> 47	
<211> 16	
<212> DNA	
<213> 人工的	
<220>	
<223> 一个人工合成的寡核苷酸序列	
<400> 47	
ttttagaat ggaagc	16

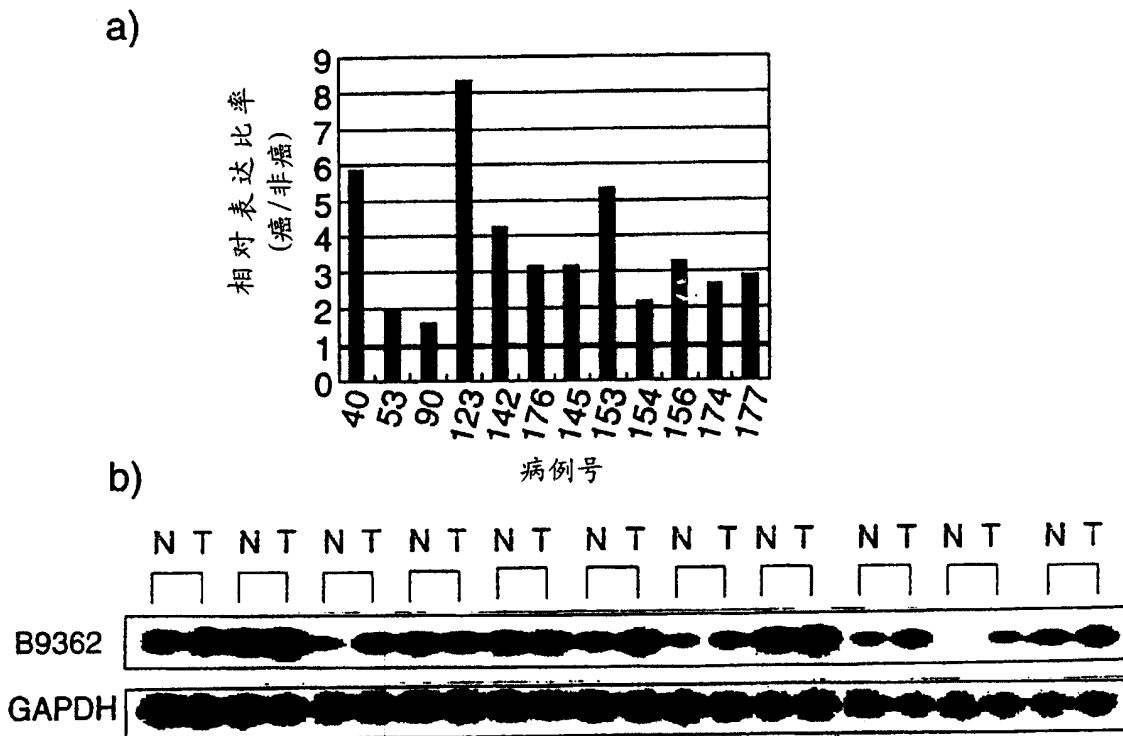


图 1

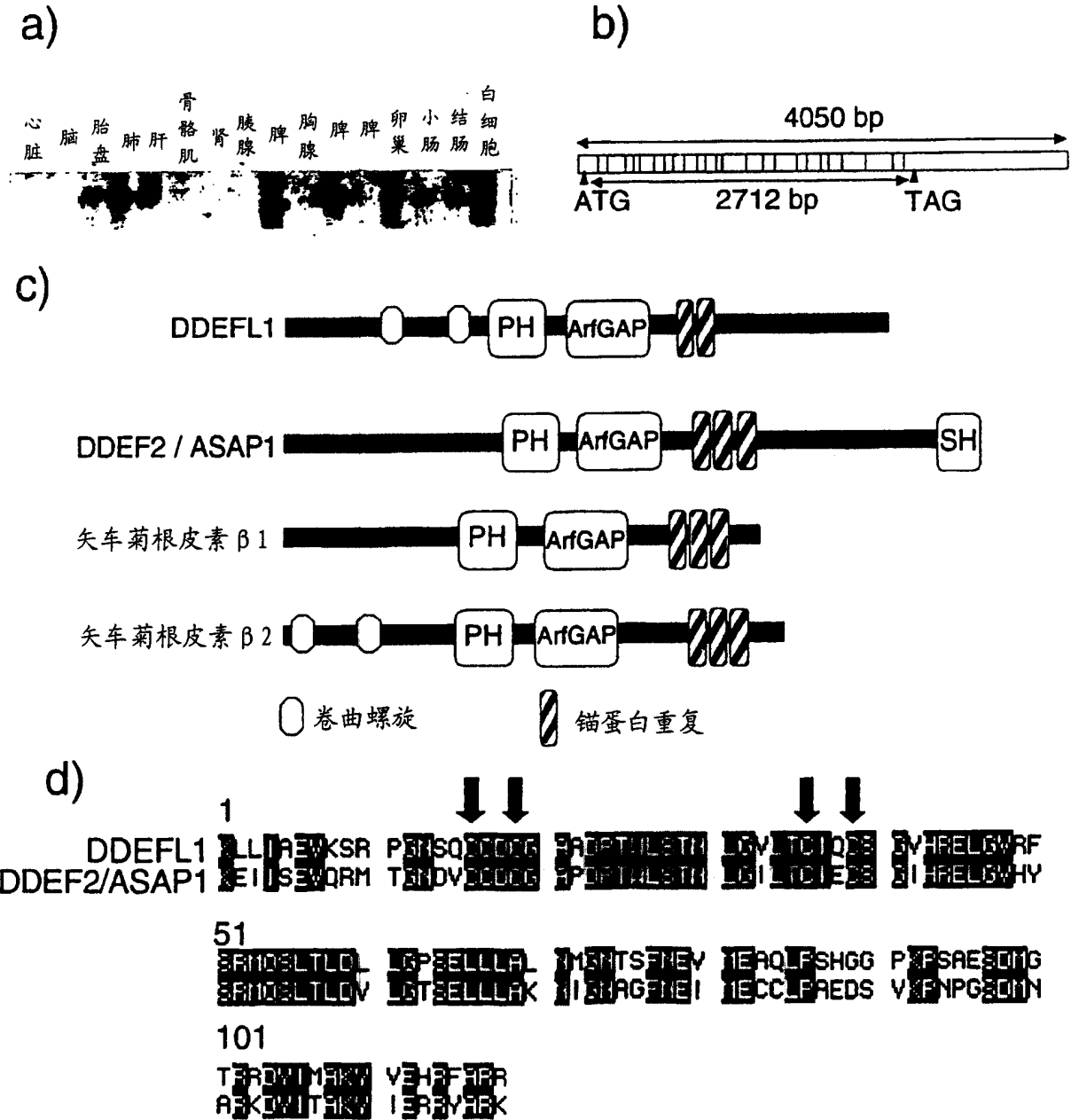


图 2

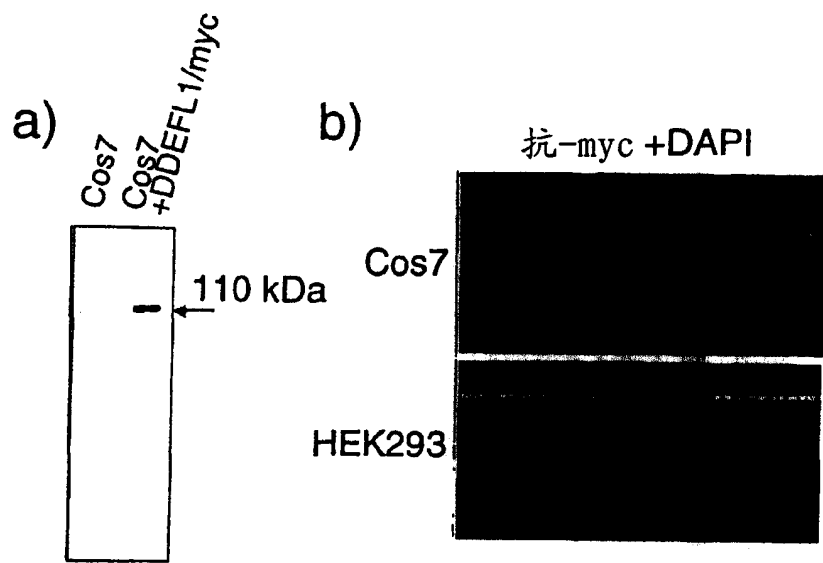


图 3

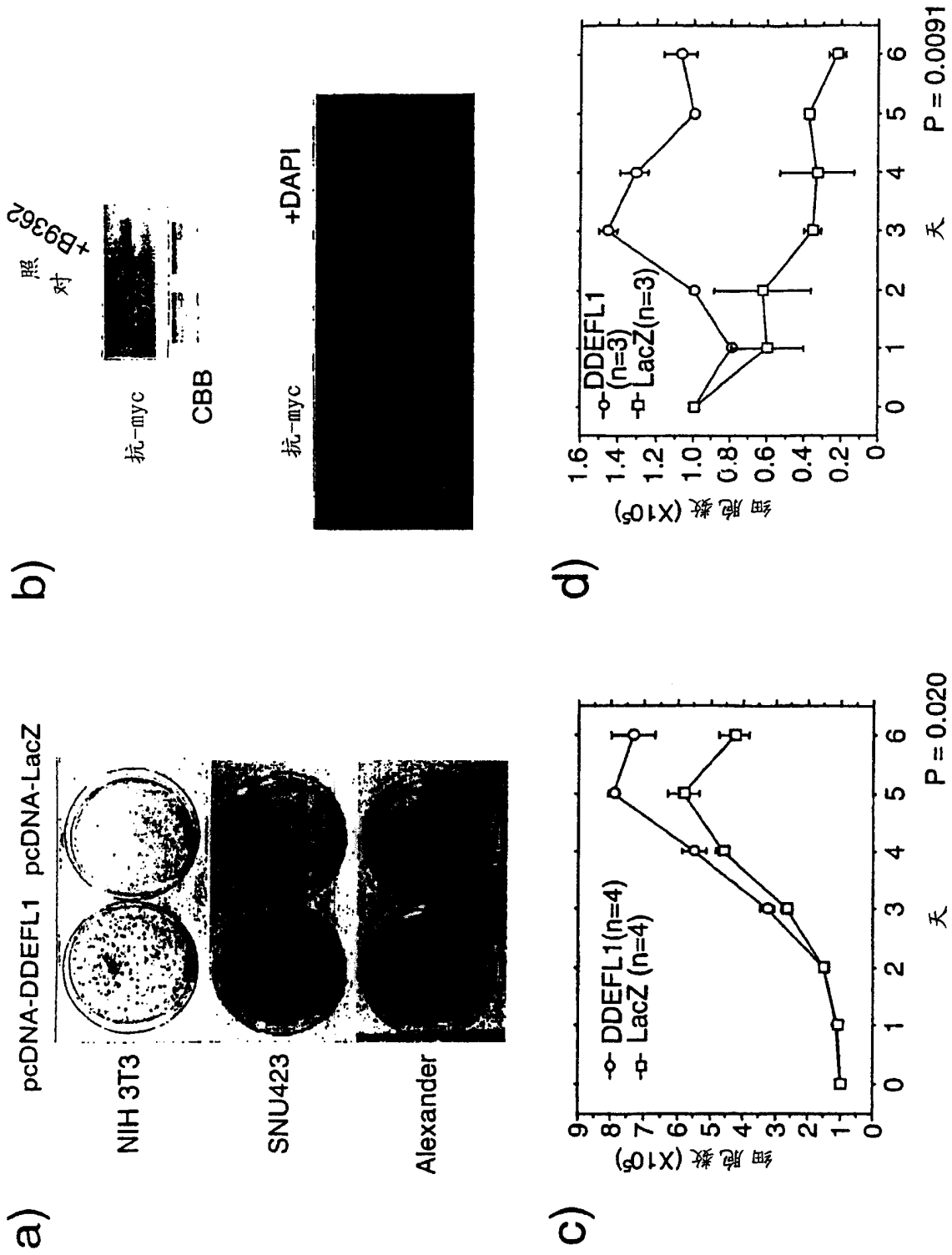


图 4

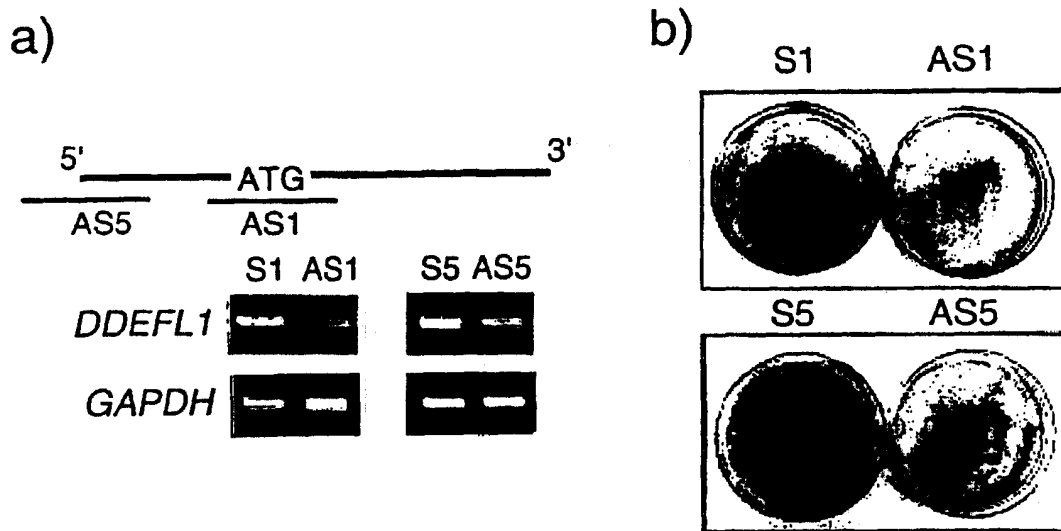


图 5

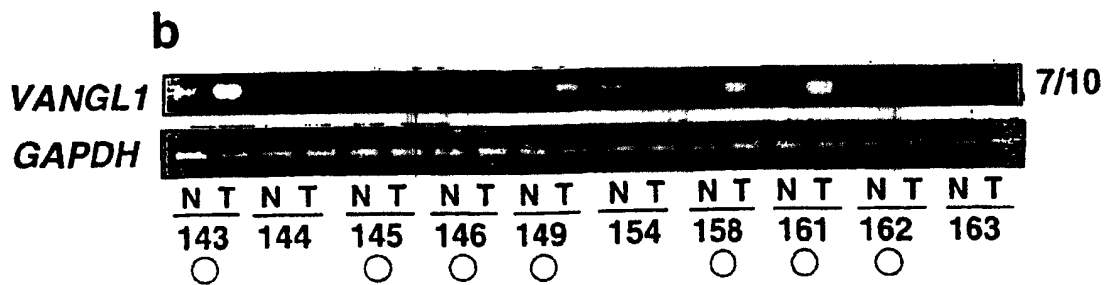
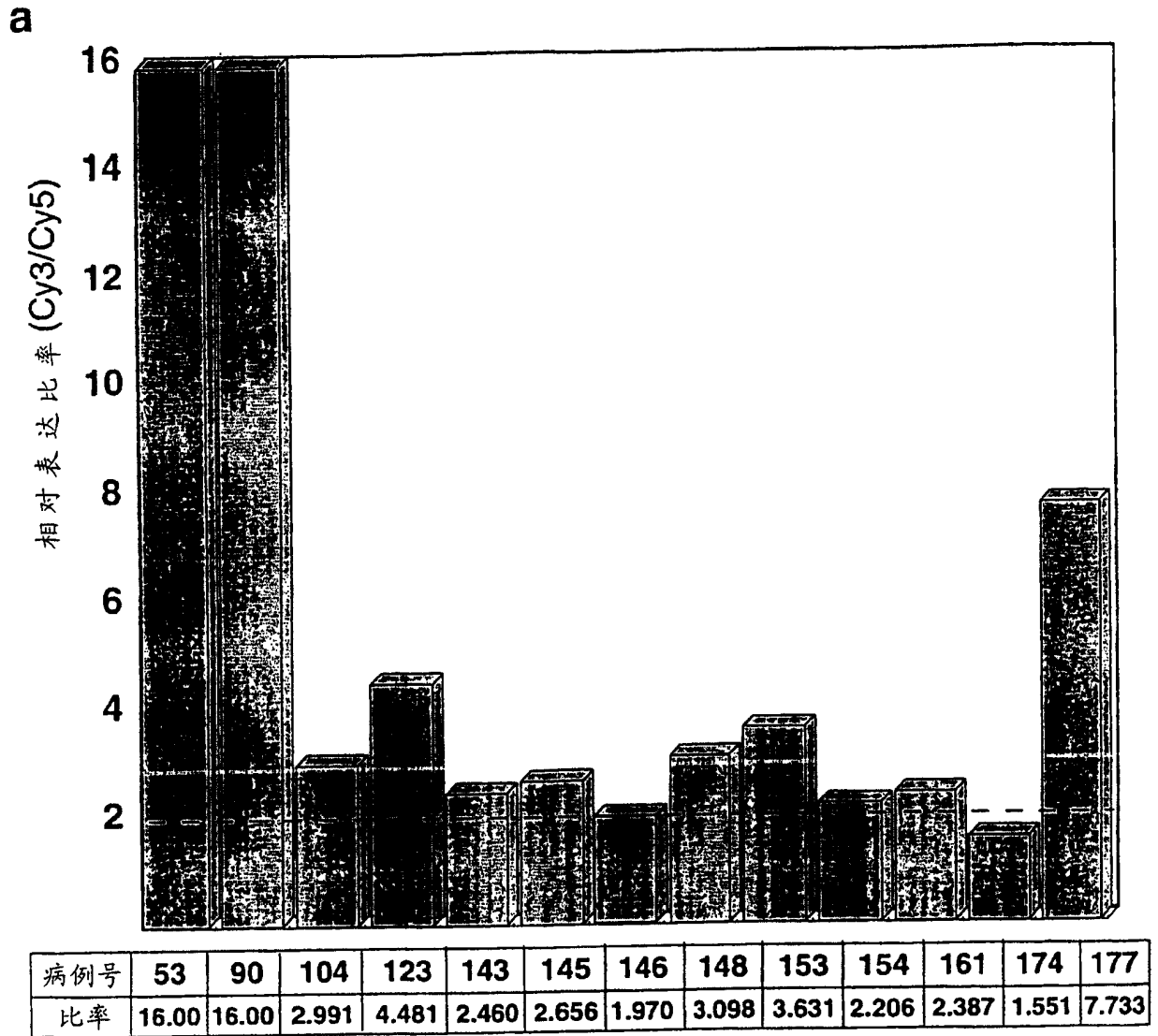


图 6

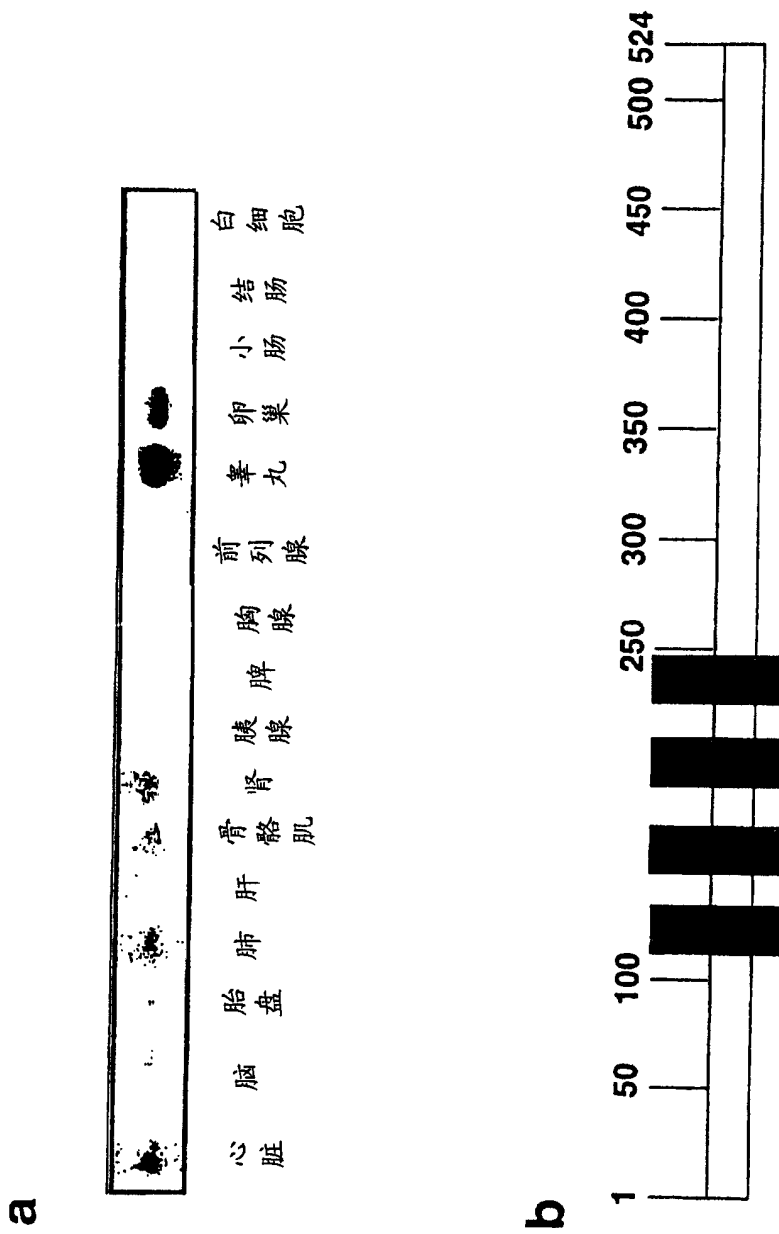


图 7

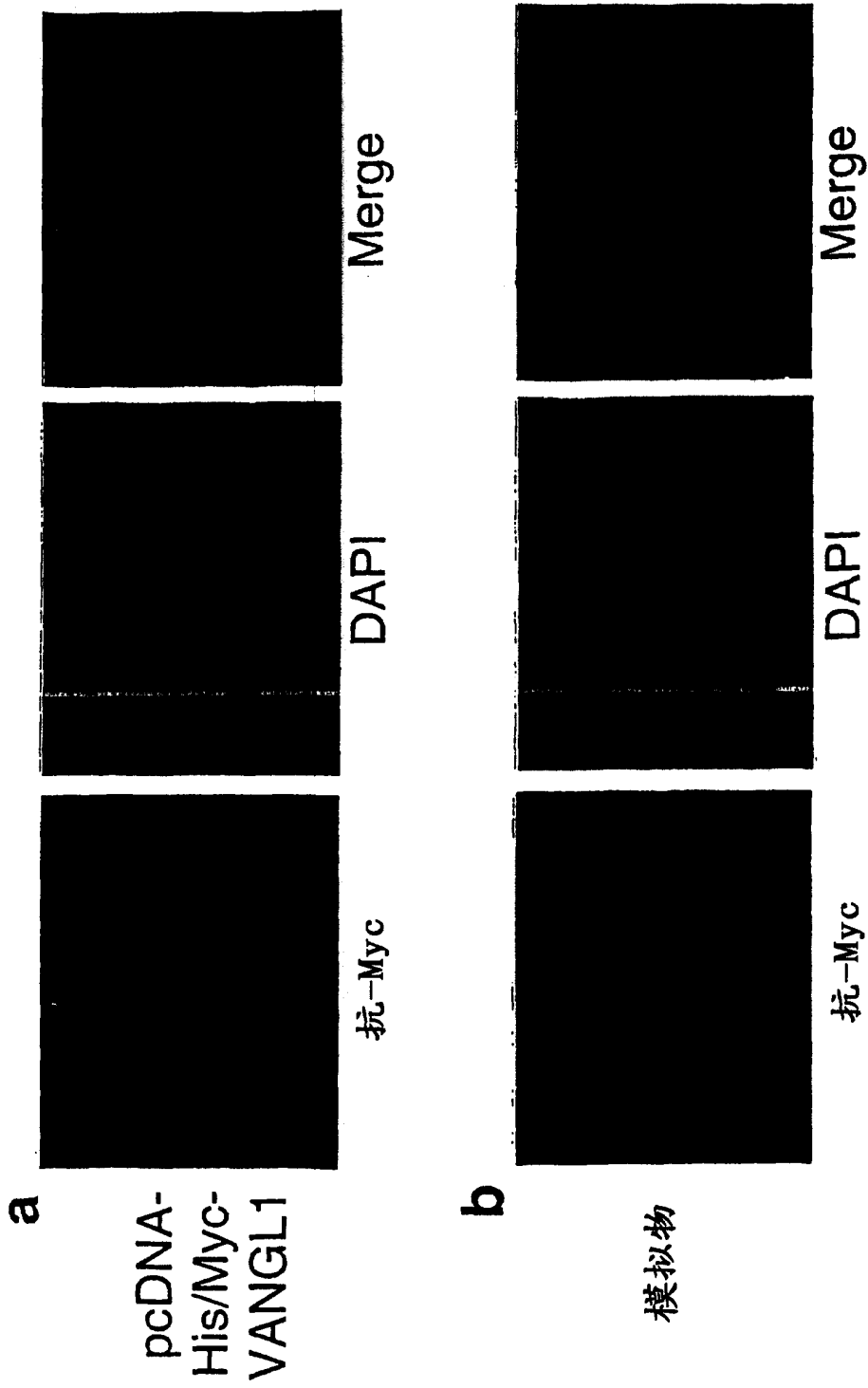


图 8

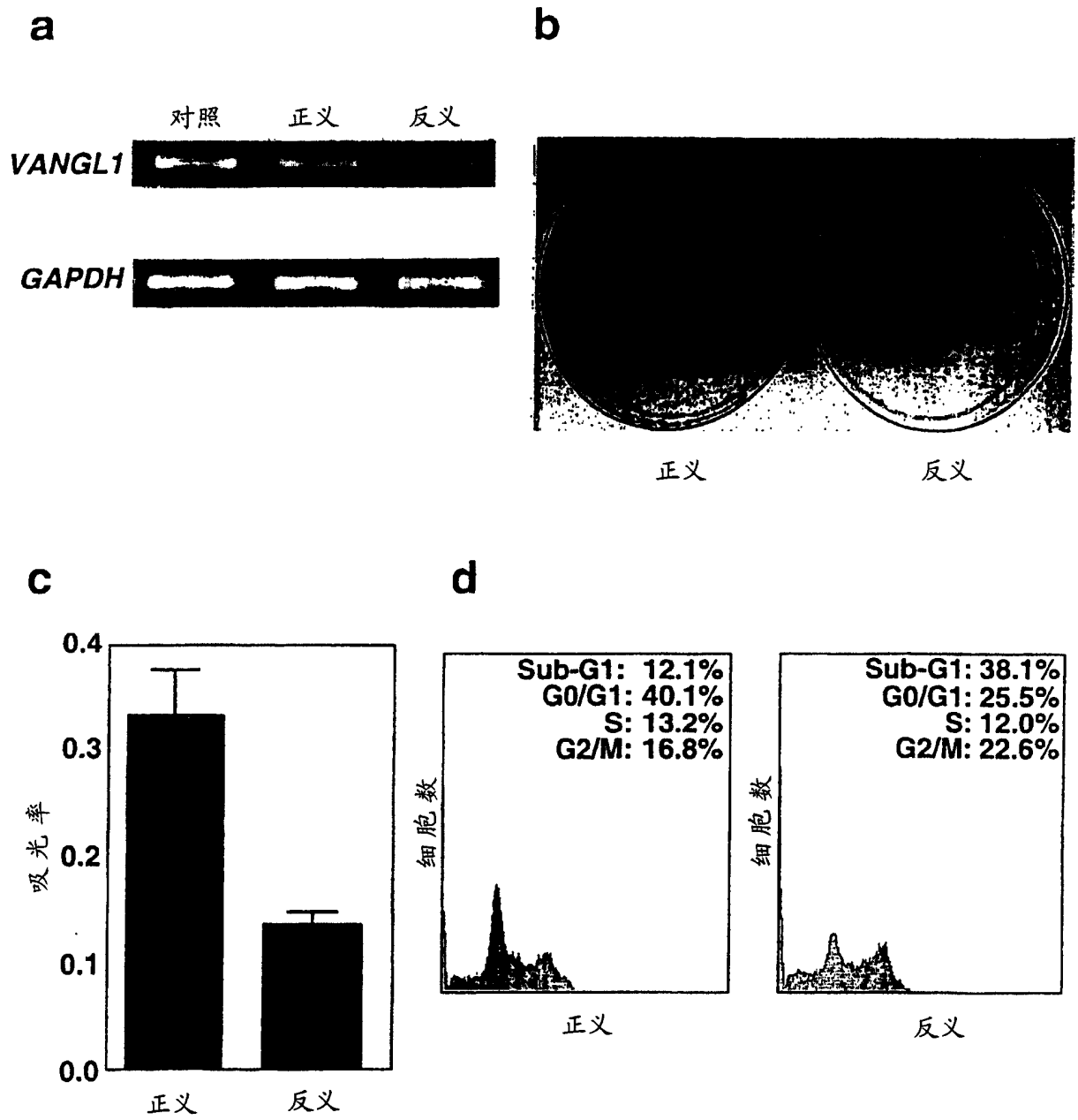


图 9

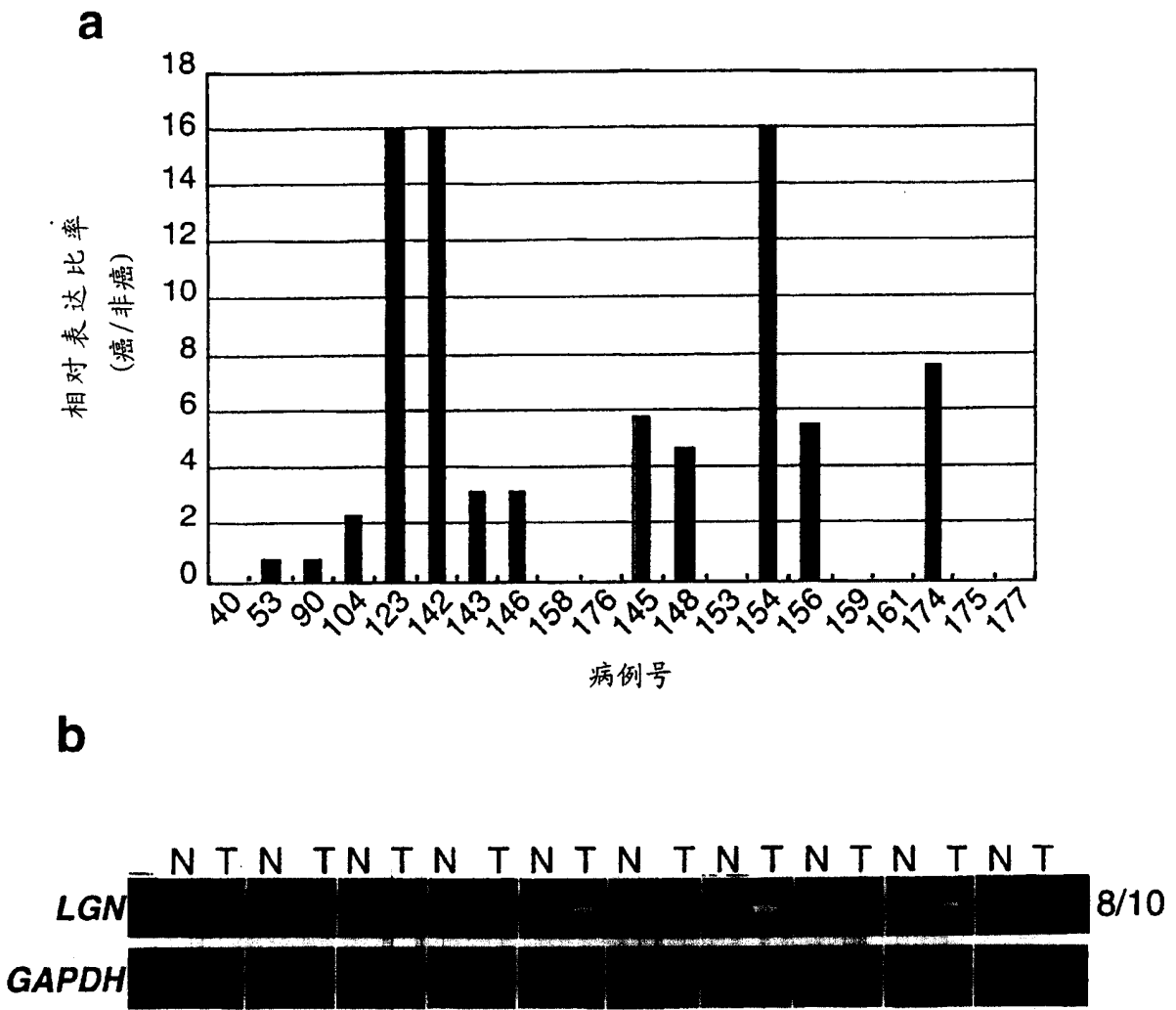


图 10

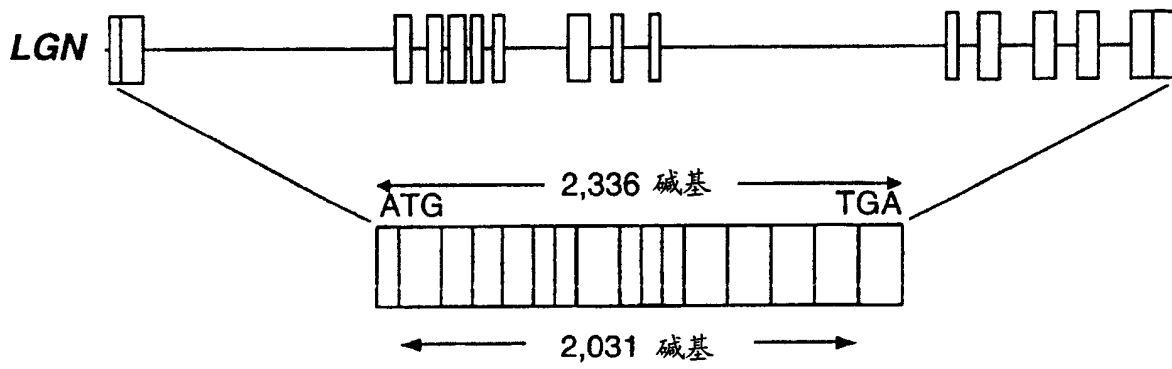


图 11



图 12

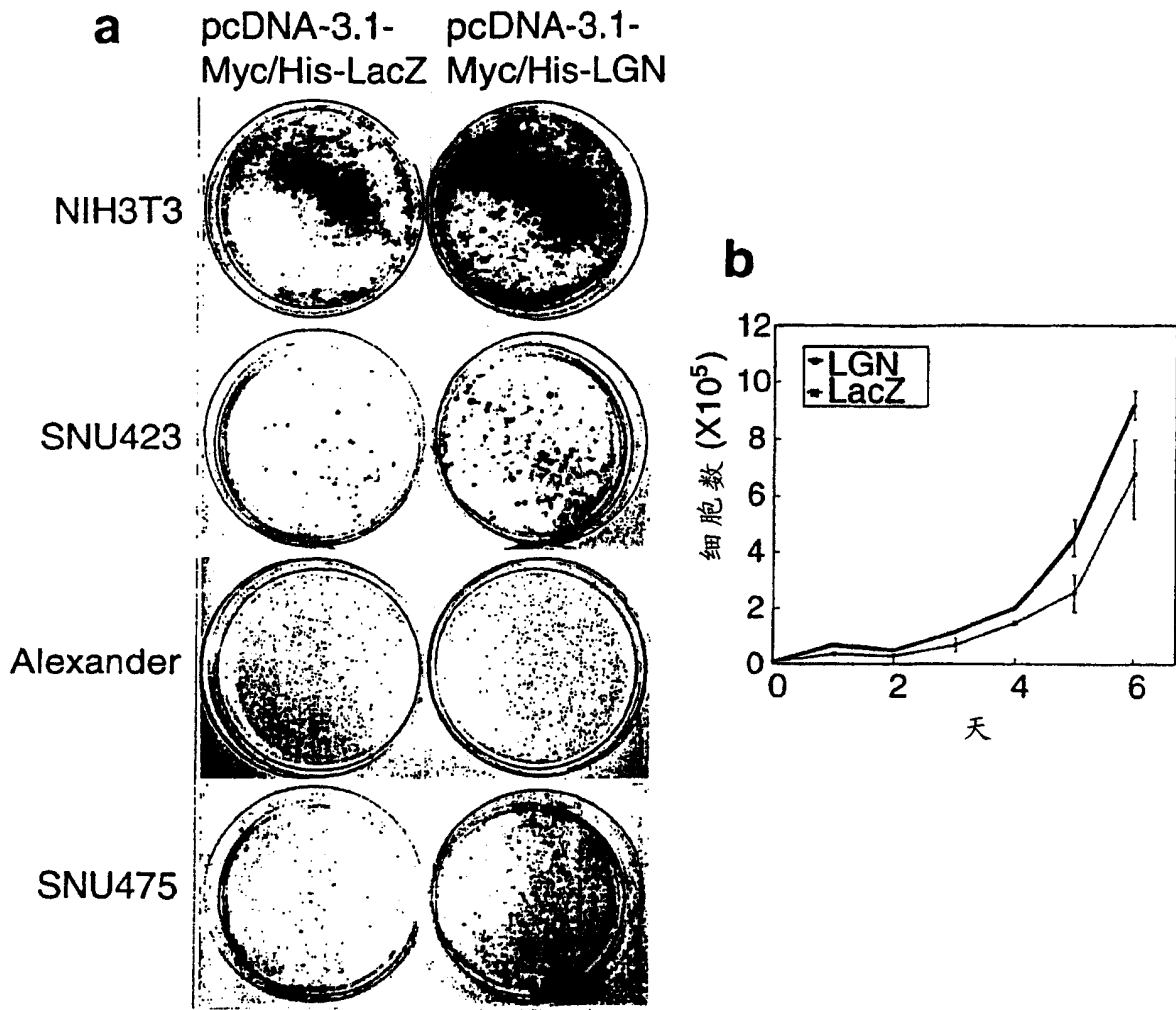


图 13

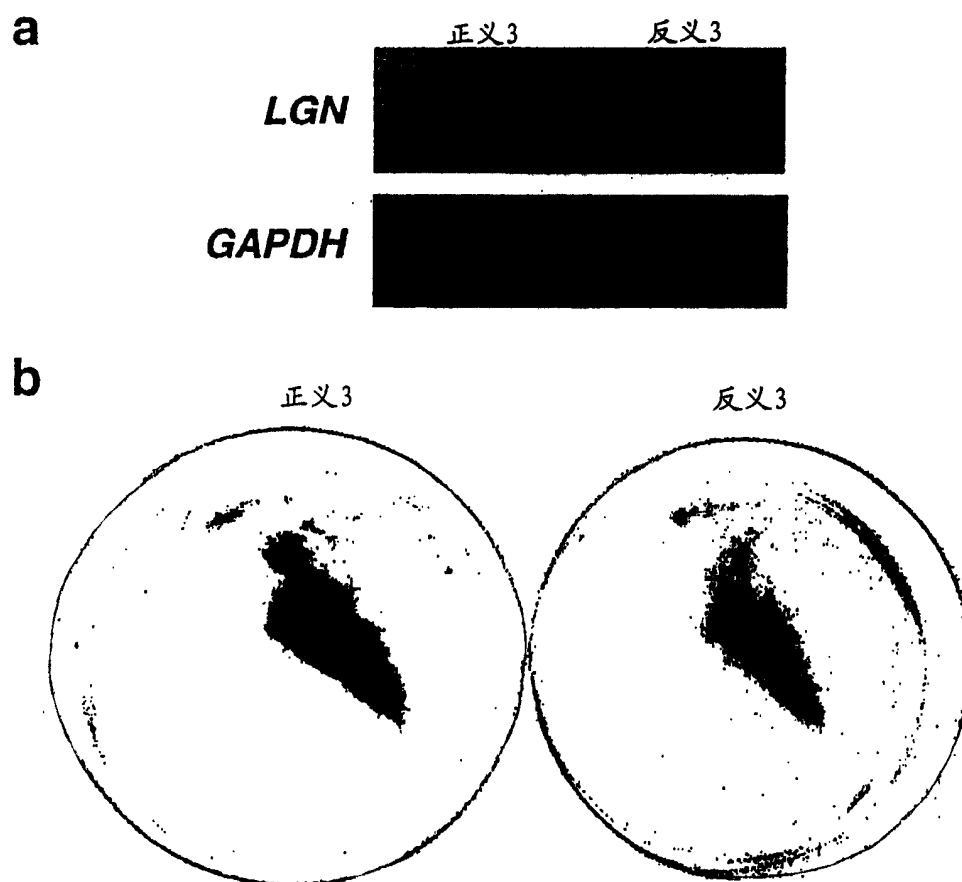


图 14