



(12) 发明专利

(10) 授权公告号 CN 1890381 B

(45) 授权公告日 2011.06.22

(21) 申请号 200480034701.1

(22) 申请日 2004.09.24

(30) 优先权数据

60/505,571 2003.09.24 US

(85) PCT申请进入国家阶段日

2006.05.24

(86) PCT申请的申请数据

PCT/JP2004/014438 2004.09.24

(87) PCT申请的公布数据

W02005/028676 EN 2005.03.31

(73) 专利权人 肿瘤疗法科学股份有限公司

地址 日本神奈川县

(72) 发明人 中村佑辅 片桐丰雅 中鹤修一

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 巫肖南 封新琴

(51) Int. Cl.

C12Q 1/68(2006.01)

C12N 9/10(2006.01)

(56) 对比文件

US 20030104418 A1, 2003.06.05, 全文.

US 20030157544 A1, 2003.08.21, 全文.

Perou C M, et al..Molecular portraits of human breast tumours..《Nature》.2000, 第406卷(第6797期), 747-752.

审查员 苏林

权利要求书 1 页 说明书 76 页

序列表 10 页 附图 13 页

(54) 发明名称

诊断乳腺癌的方法

(57) 摘要

本文描述了检测和诊断乳腺癌(BRC)的方法。在一个实施方案中,诊断方法包括测定在BRC细胞和正常细胞之间有区别的BRC-相关基因的表达水平。在另一个实施方案中,诊断方法包括测定BRC细胞中,在DCIS细胞和IDC细胞之间有区别的BRC-相关基因的表达水平。本发明还提供了使用BRC-相关基因预测和预防乳腺癌转移的方法,所述基因在具有淋巴结转移的乳腺癌细胞中有着独特的经改变的表达模式。最后,本发明还提供了筛选可用于治疗乳腺癌的治疗剂的方法,治疗乳腺癌的方法和接种受试者以使其免患乳腺癌的方法。

1. 反义组合物在制备用于治疗或预防受试者的乳腺癌的药物组合物中的用途,所述反义组合物含有与 BRC No. 456 基因的编码序列互补的核苷酸序列。

2. siRNA 组合物在制备用于治疗或预防受试者的乳腺癌的药物组合物中的用途,其中所述 siRNA 组合物包含 BRC No. 456 的 siRNA。

3. 权利要求 2 的用途,其中所述 siRNA 含有有义链,所述有义链含有选自 SEQ ID NO : 25, 28 和 31 的核苷酸序列。

4. 抗体或其免疫活性片段在制备用于治疗或预防受试者的乳腺癌的药物组合物中的用途,所述抗体抗 BRC No. 456 编码的多肽。

5. (a) 由 BRC No. 456 基因的核酸编码的多肽, (b) 所述多肽的免疫活性片段, 或 (c) 编码所述多肽的多核苷酸在制备用于治疗或预防受试者的乳腺癌的疫苗中的用途。

6. 权利要求 1 的用途,其中所述乳腺癌是 IDC。

7. 权利要求 2 的用途,其中所述乳腺癌是 IDC。

8. 权利要求 7 的用途,其中所述 siRNA 含有有义链,所述有义链含有选自 SEQ ID NO : 25, 28 和 31 的核苷酸序列。

9. 权利要求 4 的方法,其中所述乳腺癌是 IDC。

10. 权利要求 5 的方法,其中所述乳腺癌是 IDC。

11. 治疗或预防乳腺癌的组合物,所述组合物含有药物有效量的针对 BRC No. 456 基因的反义多核苷酸或 siRNA,其中所述 siRNA 含有有义链,所述有义链含有选自 SEQ ID NO : 25, 28 和 31 的核苷酸序列。

12. 权利要求 11 的组合物,其中所述乳腺癌是 IDC。

## 诊断乳腺癌的方法

[0001] 本申请要求 2003 年 9 月 24 日提交的美国临时申请流水号 60/505, 571 的权利 (该申请的全部内容在此列入本文作为参考)。

### 发明领域

[0002] 本发明涉及检测和诊断乳腺癌的方法以及治疗和预防乳腺癌和乳腺癌转移的方法。

### [0003] 发明背景

[0004] 乳腺癌是一种遗传异质疾病,是妇女中最常见的恶性肿瘤。据报道全世界每年估计约有 800000 个新病例出现 (Parkin DM, Pisani P, Ferlay J(1999). CA Cancer J Clin 49 :33-64)。乳房切除术是公认的治疗该病的首选方案。尽管可手术切除原发性肿瘤,但诊断时无法检测到的微转移会导致患处或较远的位点出现肿瘤复发 (Saphner T, Tommey DC, Gray R(1996). J Clin Oncol, 14, 2738-2749.)。通常在手术后施用细胞毒性剂作为辅助疗法,目的是杀死那些残留的或恶变前的细胞。

[0005] 用常规的化学治疗剂治疗经常需要依靠经验,大多基于组织学肿瘤参数,对特异性的疗法缺乏了解。因此,靶向药物正在成为乳腺癌的基础疗法。已证实靶向药物的两种代表性药物三苯氧胺和芳香酶抑制剂用作转移乳腺癌患者的辅助治疗剂或化学预防剂有良好的反应 (Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N(1998). J Natl Cancer Inst, 90, 1371-1388 ;Cuzick J(2002). Lancet 360, 817-824)。然而,缺点是仅有表达雌激素受体的患者对这些药物敏感。最近,人们甚至关注到与这些药物有关的副作用,具体提到长期用三苯氧胺治疗可能会导致子宫内膜癌,而且,在按处方服用芳香酶抑制剂的患者中,绝经后妇女会遭受骨折的有害副作用 (Coleman RE(2004). Oncology. 18(5Suppl 3), 16-20)。由于出现副作用和药物抗性,显然需要根据已鉴定的作用机理为选择性的灵活 (smart) 药物寻找新的分子靶标。

[0006] 乳腺癌是一种复杂的疾病,该病与多个遗传改变有关。据报道,乳腺癌的发生是一个多步骤的过程,该过程在广义上等同于正常细胞的转化,所述转化经过非典型性导管增生、原位导管癌 (DCIS) 和浸润性导管癌 (IDC) 这几个阶段,然而,对这些异常是否是乳腺肿瘤发生的原因仍知之甚少。有证据表明:仅有部分恶化前损害会发展成为浸润性癌症,而其它损害会自发消退。有关导致原发性乳腺癌的发生、发展和转移形成的分子参与的解释是目的在于预防和治疗乳腺癌的新策略的主要焦点。

[0007] 与传统的组织病理学方法能够提供的信息相比,由 cDNA 微阵列分析产生的基因-表达分布图能提供更多有关各个癌症特性的细节。对该信息的指望在于它能够改善治疗肿瘤疾病的临床策略和开发新药 (Petricoin, E. F., 3rd, Hackett, J. L., Lesko, L. J., Puri, R. K., Gutman, S. I., Chumakov, K., Woodcock, J., Feigal, D. W., Jr., Zoon, K. C., and Sistare, F. D. Medical applications of microarray technologies: a regulatory science perspective. Nat Genet, 32 Suppl :474-479, 2002)。为此,本发明人通过 cDNA 微

阵列分析了得自多种组织的肿瘤的表达分布图 (Okabe, H. et al., Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. *Cancer Res*, 61:2129-2137, 2001. ; Hasegawa, S. et al., Genome-wide analysis of gene expression in intestinal-type gastric cancers using a complementary DNAmicroarray representing 23,040 genes. *Cancer Res*, 62:7012-7017, 2002. ; Kaneta, Y. et al., and Ohno, R. Prediction of Sensitivity to STI571 among Chronic Myeloid Leukemia Patients by Genome-wide cDNA Microarray Analysis. *Jpn J Cancer Res*, 93:849-856, 2002. ; Kaneta, Y. et al., Genome-wide analysis of gene-expression profiles in chronic myeloid leukemia cells using a cDNAmicroarray. *Int J Oncol*, 23:681-691, 2003. ; Kitahara, O. et al., Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. *Cancer Res*, 61:3544-3549, 2001. ; Lin, Y. et al. Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. *Oncogene*, 21:4120-4128, 2002. ; Nagayama, S. et al., Genome-wide analysis of gene expression in synovial sarcomas using a cDNAmicroarray. *Cancer Res*, 62:5859-5866, 2002. ; Okutsu, J. et al., Prediction of chemosensitivity for patients with acute myeloid leukemia, according to expression levels of 28 genes selected by genome-wide complementary DNAmicroarray analysis. *Mol Cancer Ther*, 1:1035-1042, 2002. ; Kikuchi, T. et al., Expression profiles of non-small cell lung cancers on cDNA microarrays: identification of genes for prediction of lymph-node metastasis and sensitivity to anti-cancer drugs. *Oncogene*, 22:2192-2205, 2003. )。

[0008] 最近,有人使用 cDNA 微阵列详细检查了数千个基因的表达水平,结果在不同类型的乳腺癌中发现了不同的模式 (Sgroi, D. C. et al., In vivo gene expression profile analysis of human breast cancer progression. *Cancer Res*, 59:5656-5661, 1999. ; Sorlie, T. et al., Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*, 98:10869-10874, 2001. ; Kauraniemi, P. et al., New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNAmicroarrays. *Cancer Res*, 61:8235-8240, 2001. ; Gruvberger, S. et al., S. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res*, 61:5979-5984, 2001. ; Dressman, M. et al., Gene expression profiling detects gene amplification and differentiates tumor types in breast cancer. *Cancer Res*, 63:2194-2199, 2003. )。

[0009] 通过对乳腺癌的基因表达分布图进行研究,鉴定出可用作候选诊断标记或预后分布图的基因。然而,这些主要得自肿瘤块的数据并不能充分反映乳腺癌发生过程中的表达改变,因为乳腺癌细胞以实体块的形式存在,所述实体块具有严重的炎症反应并含有多种

细胞组分。因此,以前公开的微阵列数据可能反映的是异源分布图。

[0010] 被设计用于揭示癌症发生机制的研究促使人们鉴定出某些抗-肿瘤剂的分子靶。例如,起初被开发用于抑制与 Ras 有关的生长-信号传导途径、其活化取决于翻译后法尼基化的法尼基转移酶抑制剂 (FTI) 能在动物模型中有效治疗依赖于 Ras 的肿瘤 (He et al., *Cell* 99:335-45(1999))。类似地,以拮抗原癌基因受体 HER2/neu 为目的,使用抗癌药物和抗-HER2 单克隆抗体 trastuzumab 的组合对人体进行的临床试验已使乳腺癌患者获得改善的临床反应并使其全部存活 (Lin et al., *Cancer Res* 61:6345-9(2001))。最终,已开发出能选择性失活 bcr-abl 融合蛋白的酪氨酸激酶抑制剂 STI-571 以治疗慢性髓性白血病,在所述疾病中,bcr-abl 酪氨酸激酶的组成型活化在白细胞转化中起着至关重要的作用。此类药剂被设计用于抑制具体基因产物的致癌活性 (Fujita et al., *Cancer Res* 61:7722-6(2001))。因此,在癌细胞中一般被上调的基因产物显然可用作潜在的靶,该靶可用于开发新的抗癌剂。

[0011] 经证实, CD8+ 细胞毒 T 淋巴细胞 (CTL) 能识别 MHC I 类分子上呈递的肿瘤相关抗原 (TAA) 的表位肽并裂解肿瘤细胞。第一例 TAA 是 MAGE 家族,由于这一发现,使用免疫学方法已发现很多其它的 TAA (Boon, *Int J Cancer* 54:177-80(1993); Boon and van der Bruggen, *J Exp Med* 183:725-9(1996); van der Bruggen et al., *Science* 254:1643-7(1991); Brichard et al., *J Exp Med* 178:489-95(1993); Kawakami et al., *J Exp Med* 180:347-52(1994))。目前,临床上将一些新发现的 TAA 作为免疫疗法的靶进行开发。迄今为止所发现的 TAA 包括 MAGE (van der Bruggen et al., *Science* 254:1643-7(1991)), gp100 (Kawakami et al., *J Exp Med* 180:347-52(1994)), SART (Shichijo et al., *J Exp Med* 187:277-88(1998)) 和 NY-ES0-1 (Chen et al., *Proc Natl Acad Sci USA* 94:1914-8(1997))。另一方面,经证实在肿瘤细胞中特异性过表达的基因产物可作为诱导细胞免疫应答的靶被识别。所述基因产物包括 p53 (Umano et al., *Brit J Cancer* 84:1052-7(2001)), HER2/neu (Tanaka et al., *Brit J Cancer* 84:94-9(2001)), CEA (Nukaya et al., *Int J Cancer* 80:92-7(1999)) 等。

[0012] 尽管已在与 TAA 有关的基础和临床研究中取得显著进步 (Rosenberg et al., *Nature Med* 4:321-7(1998); Mukherji et al., *Proc Natl Acad Sci USA* 92:8078-82(1995); Hu et al., *Cancer Res* 56:2479-83(1996)), 但目前可以使用的能治疗腺癌,包括结肠直肠癌的候选 TAA 的数目仍然很有限。局限于在癌细胞中表达且在癌细胞中表达量高的 TAA 可能是理想的候选免疫治疗靶。此外,鉴定能诱导有效而特异性的抗-肿瘤免疫应答的新型 TAA 有望促进在临床上针对多种类型的癌症使用肽接种策略 (Boon and van der Bruggen, *J Exp Med* 183:725-9(1996); van der Bruggen et al., *Science* 254:1643-7(1991); Brichard et al., *J Exp Med* 178:489-95(1993); Kawakami et al., *J Exp Med* 180:347-52(1994); Shichijo et al., *J Exp Med* 187:277-88(1998); Chen et al., *Proc Natl Acad Sci USA* 94:1914-8(1997); Harris, *J Natl Cancer Inst* 88:1442-5(1996); Butterfield et al., *Cancer Res* 59:3134-42(1999); Vissers et al., *Cancer Res* 59:5554-9(1999); van der Burg et al., *J Immunol* 156:3308-14(1996); Tanaka et al., *Cancer Res* 57:4465-8(1997); Fujie et al., *Int J Cancer* 80:169-72(1999); Kikuchi et al., *Int J Cancer* 81:459-66(1999); Oiso et al., *Int J*

Cancer81 :387-94(1999))。

[0013] 据多篇文献报道,得自某些健康供体的经肽刺激的外周血单核细胞(PBMC)对肽作出反应,产生显著水平的IFN- $\alpha$ ,但在 $^{51}\text{Cr}$ -释放试验中很少以局限于HLA-A24或-A0201的方式对肿瘤细胞发挥细胞毒性(Kawano et al., Cancer Res 60 :3550-8(2000); Nishizaka et al., Cancer Res 60 :4830-7(2000); Tamura et al., Jpn J Cancer Res 92 :762-7(2001))。然而,在日本人和白种人中,HLA-A24和HLA-A0201是常见的HLA等位基因(Date et al., Tissue Antigens47 :93-101(1996); Kondo et al., J Immunol 155 :4307-12(1995); Kubo et al., J Immunol 152 :3913-24(1994); Imanishi et al., Proceeding of the eleventh International Histocompatibility Workshop and Conference Oxford University Press, Oxford, 1065(1992); Williams et al., Tissue Antigen 49 :129(1997))。因此,由这些HLA呈递的癌症抗原肽对于治疗日本人和白种人的癌症特别有用。另外,已知使用高浓度的肽通常可在体外诱导低亲和力的CTL,在抗原呈递细胞(APC)上产生高水平的特异性肽/MHC复合物,该复合物会有效激活这些CTL(Alexander-Miller et al., Proc Natl Acad Sci USA 93 :4102-7(1996))。

[0014] 因此,为了掌握与癌症有关的致癌机制并鉴定出潜在的靶用以开发新的抗癌剂,本发明人使用展现了23,040个基因的cDNA微阵列对在纯化的乳腺癌细胞群体中发现的基因表达分布图进行了大规模的基因组范围内的分析,其中所述乳腺癌包括12个原位导管癌(DCIS)和69个浸润性导管癌(IDC)。

[0015] 发明概述

[0016] 本发明基于发现与乳腺癌(BRC)相关的基因表达模式。本文将这些在乳腺癌中差异表达的基因统称为“BRC核酸”或“BRC多核苷酸”,将其相应的编码多肽称为“BRC多肽”或“BRC蛋白”。

[0017] 因此,本发明提供了通过检测得自患者的生物样品,如组织样品中BRC-相关基因的表达水平,以诊断或测定受试者对乳腺癌的易感性的方法。术语“BRC-相关基因”指的是特征在于表达水平在BRC细胞和正常细胞之间有差别的基因。正常细胞是得自乳腺组织的细胞。在本发明的上下文中,BRC-相关基因是表3-8所列出的基因(即BRC编号为123-512的基因)。表达水平的改变,例如与基因的正常对照水平相比,基因的表达水平提高或降低表示受试者患有BRC或有发展成为BRC的风险。

[0018] 在本发明的上下文中,术语“对照水平”指的是在对照样品中检测到的蛋白质表达水平,其包括正常对照水平和乳腺癌对照水平。对照水平可以是得自单个参照群体或得自多个表达模式的单个表达模式。例如,对照水平可以是以前检测过的细胞的表达模式数据库。“正常对照水平”指的是在已知未患乳腺癌的正常健康个体或一群个体中检测到的基因表达水平。正常个体是没有乳腺癌临床症状的个体。另一方面,“BRC对照水平”指的是在患有BRC的人群中发现的BRC-相关基因的表达分布图。

[0019] 与正常对照水平相比,在受试样品中检测到表3,5和7所列的一个或多个BRC-相关基因(即BRC编号为123-175,374-398和448-471的基因)表达水平的提高表示受试者(样品的提供者)患有BRC或有患BRC的风险。相反,与正常对照水平相比,在受试样品中检测到表4,6和8所列的一个或多个BRC-相关基因(即BRC编号为176-373,399-447和472-512的基因)表达水平的降低表示所述受试者患有BRC或有患BRC的风险。

[0020] 或者,将样品中的一系列 BRC- 相关基因的表达与相同基因系列的 BRC 对照水平相比较。样品表达和 BRC 对照表达之间的相似性表示受试者(样品的提供者)患有 BRC 或有患 BRC 的风险。

[0021] 根据本发明,当基因表达与对照水平相比增加或降低 10%, 25%, 50% 时,可认为基因表达水平是“经改变的”。或者,如果与对照水平相比,基因表达增加或降低至少 0.1 倍,至少 0.2 倍,至少 1 倍,至少 2 倍,至少 5 倍或至少 10 倍或更多倍时,可认为表达水平是“增加的”或“降低的”。通过例如在微阵列上检测 BRC- 相关基因探针与得自患者的组织样品的基因转录物之间的杂交即可检测表达。

[0022] 在本发明的上下文中,得自患者的组织样品可以是得自受试者,例如已知或被怀疑患有 BRC 的患者的任何组织。例如,组织可含有上皮细胞。更具体地,组织可以是得自乳腺导管癌的上皮细胞。

[0023] 本发明还提供了 BRC 参照表达分布图,其中包含表 3-8 所列的两个或多个 BRC- 相关基因的基因表达水平。或者,BRC 参照表达分布图可包含表 3,5 和 7 或表 4,6 和 8 所列的两个或多个 BRC- 相关基因的表达水平。

[0024] 本发明还提供了通过使表达 BRC- 相关基因的受试细胞与受试化合物接触,并测定 BRC- 相关基因的表达水平或其基因产物的活性来鉴定药剂的方法,所述药剂能抑制或增强 BRC- 相关基因,例如表 3-8 所列 BRC- 相关基因的表达或活性。受试细胞可以是上皮细胞,如得自乳腺癌的上皮细胞。与上调的 BRC- 相关基因或其基因产物的正常对照水平或活性相比,所述基因的表达水平或其基因产物的活性的降低表示受试药剂是 BRC- 相关基因的抑制剂,并可用于减轻 BRC 的症状,例如降低表 3,5 和 7 所列的一种或多种 BRC- 相关基因的表达。或者,与下调的 BRC- 相关基因或其基因产物的正常对照水平或活性相比,所述基因的表达水平或其基因产物的活性的增加表示受试药剂是 BRC- 相关基因表达或功能的增强剂,并可用于减轻 BRC 的症状,例如减轻表 4,6 和 8 所列的一种或多种 BRC- 相关基因的低表达。

[0025] 本发明还提供了试剂盒,其中含有能与一种或多种 BRC 核酸或 BRC 多肽结合的检测试剂。本发明还提供了能与一种或多种 BRC 核酸结合的核酸微阵列。

[0026] 本发明的治疗方法包括在受试者中治疗或预防 BRC 的方法,该方法包括给受试者施用反义组合物的步骤。在本发明的上下文中,反义组合物能降低特异性靶基因的表达。例如,反义组合物可含有与选自表 3,5 和 7 所列的 BRC- 相关基因的 BRC- 相关基因序列互补的核苷酸。或者,本发明的方法可包括给受试者施用小的干扰 RNA (siRNA) 组合物的步骤。在本发明的上下文中,siRNA 组合物能降低选自表 3,5 和 7 所列的 BRC- 相关基因的 BRC 核酸的表达。在另一个方法中,通过给受试者施用核酶组合物以在受试者中治疗或预防 BRC。在本发明的上下文中,核酸-特异性核酶组合物能降低选自表 3,5 和 7 所列的 BRC- 相关基因的 BRC 核酸的表达。实际上,已证实 siRNA 对表中所列 BRC- 相关基因的抑制作用。例如,在实施例部分已清楚地证明针对表 7 中的 BRC-456 (GenBank 登录号为 AF237709, TOPK ;源自 T-LAK 细胞的蛋白质激酶) 的 siRNA 可抑制乳腺癌细胞的细胞增殖。因此,在本发明中,表 3,5 和 7 所列的 BRC- 相关基因,特别是 BRC-456 是优选的乳腺癌治疗靶。其它治疗方法包括给受试者施用化合物的方法,所述化合物能增加表 4,6 和 8 所列的一种或多种 BRC- 相关基因的表达或由表 4,6 和 8 所列的一种或多种 BRC- 相关基因编码的多肽的活性。

[0027] 本发明还包括疫苗和接种方法。例如,在受试者中治疗或预防 BRC 的方法包括给受试者施用疫苗,所述疫苗含有由选自表 3,5 和 7 所列 BRC- 相关基因的核酸编码的多肽或所述多肽的免疫活性片段。在本发明的上下文中,免疫活性片段是长度短于全长天然蛋白质,但其诱导的免疫应答与全长蛋白质所诱导的相类似的多肽。例如,免疫活性片段的长度应至少为 8 个残基,并能刺激免疫细胞,如 T 细胞或 B 细胞。通过检测细胞增殖、细胞因子(如 IL-2)合成或抗体产生即可测定免疫细胞刺激。

[0028] 另外,本发明还提供了治疗或预防乳腺癌转移的靶分子。根据本发明,将表 11 所列基因(即 BRC 编号为 719-752 的基因)鉴定为在具有淋巴结转移的乳腺癌细胞中具有独特的经改变表达模式的基因。因此,通过抑制选自 VAMP3, MGC11257, GSPT1, DNM2, CFL1, CLNS1A, SENP2, NDUFS3, NOP5/NOP58, PSMD13, SUOX, HRB2, LOC154467, THTPA, ZRF1, LOC51255, DEAF1, NEU1, UGCGL1, BRAF, TUFM, FLJ10726, DNAJB1, AP4S1 和 MRPL40 的上调基因或其基因产物的表达或活性即可治疗或预防乳腺癌转移。或者,通过增强癌细胞中 UBA52, GenBank Acc# AA634090, CEACAM3, C21orf97, KIAA1040, EEF1D, FUS, GenBank Acc# AW965200 和 KIAA0475 的表达或活性即可治疗或预防乳腺癌转移。

[0029] 本发明还提供了预测乳腺癌转移的方法。具体地说,所述方法包括测定选自表 11 所列基因的标记基因的表达水平的步骤。本文将这些标记基因鉴定为在具有淋巴结转移的患者的乳腺癌细胞中具有独特的经改变表达模式的基因。因此,通过测定在得自受试者的样品中测得的表达水平是否与参照样品中淋巴结转移阳性病例或阴性病例的平均表达水平接近,即可预测受试者中的乳腺癌转移。

[0030] 在上调基因中,我们鉴定出被称为源自 T-LAK 细胞的蛋白质激酶(TOPK)的 A7870,它在能获得表达数据的 39 例乳腺癌中的 30 例(77%)中,特别是在 36 例浸润性导管癌样品的 29 例(81%)中过表达 3 倍以上。随后的半-定量 RT-PCR 也证实了与包括乳腺导管细胞或正常乳腺的正常人器官相比,A7870 在 12 个临床乳腺癌样品的 7 个中和 20 个乳腺癌细胞系的 17 个中被上调。Northern 印迹分析揭示出 A7870 转录物仅在乳腺癌细胞系和正常人睾丸和胸腺中表达。用 TOPK 抗体进行的免疫细胞化学染色证实:在乳腺癌细胞系 T47D, BT20 和 HBC5 的胞质和核膜周围观察到内源性 A7870 的亚细胞定位。用小的干扰 RNA(siRNA) 处理乳腺癌细胞能有效抑制 A7870 的表达,并能抑制乳腺癌细胞系 T47D 和 BT-20 的细胞/肿瘤生长,这表明该基因在细胞生长增殖中起着关键作用。这些发现暗示着 A7870 的过表达可能参与乳腺肿瘤的发生,可能是特异性治疗乳腺癌患者的理想策略。

[0031] 除非另有说明,本文所用的所有技术和科学术语都与本发明所属技术领域中的普通技术人员所常规理解的意思相同。尽管在实施或检验本发明时可使用与本文所述类似或等同的方法和材料,但下文将描述适当的方法和材料。本文提及的所有出版物、专利申请、专利和其它参考文献都全文引入本文作为参考。在存在冲突的情况下,以本申请的说明书,包括其中的定义为准。另外,本文描述的材料、方法和实施例仅是为了阐明而非限制本发明。

[0032] 本文所述方法的一个优点是在检测到乳腺癌明显的临床症状之前即可鉴定出该疾病。从下文的详细描述和权利要求书中可明显看出本发明的其它特征和优点。

[0033] 附图简述

[0034] 图 1 显示了显微解剖前(A道),显微解剖后(B道)和经显微解剖的细胞(C道)



的图像。使用激光微束显微解剖 (LMM) 技术对 DCIS, IDC 细胞和正常乳腺导管上皮细胞进行显微解剖。从经苏木精和伊红染色的切片上显微解剖得自每个样品的 DCIS 细胞 (10326T 病例), IDC 细胞 (10502T) 和正常乳腺导管上皮细胞 (10341N)。

[0035] 图 2 显示了跨越 102 个样品的 710 个基因的未经监督的二维等级聚类分析结果。图 2(A) 中的每一横行表示乳腺癌患者, 每一纵列的柱形显示出单个基因。标以红色和绿色的每个孔的颜色分别表示高于和低于跨越所有样品的基因的转录物水平平均值的转录物水平。星号标记表示主要历史类型, “#” 形标记表示相同病例中的次要历史类型。方形表示完全相同的病例 (10149a1 和 10149a1T)。黑色的方形表示未改变的表达。ER 指的是通过 EIA 测定的 ER 状态, LN 指的是淋巴结转移状态, ESR1 指的是在此微阵列中 ESR1 的表达分布图。图 2(B) 显示了跨越 16 个样品的 89 个基因的二维等级聚类分析, 所述样品显微解剖自 8 位乳腺癌患者并具有 2 个分化的损害。图 2(C) 显示出使用 25 个基因的聚类分析, 所述基因在高度 - 和低度分化的浸润性导管癌细胞之间显示出差异表达。

[0036] 图 3 显示出使用通过随机 - 排列测验选择的 97 个基因进行的经监督的基因等级聚类分析。横行中显示出 41 个 ER- 阳性样品和 28 个 ER- 阴性样品 (选自绝经前患者)。在纵列的柱形中, 根据相对表达率的相似性将 97 个基因聚类至不同的分支。优先以类似于 ESR1 表达水平以及图 2(A) 的方式表达下方主分支中的基因。上方分支中的基因与 ESR1 成反比。

[0037] 图 4 显示出在相对于正常导管细胞的 DCIS 和相对于 DCIS 的 IDC 中具有经改变的表达的基因。图 4(A) 表示通常在 DCIS 和 IDC 中被上调或下调的 251 个基因的基因簇。图 4(B) 表示在从 DCIS 转变为 IDC 时表达有所升高或降低的 74 个基因的基因簇。图 4(C) 表示在 IDC 中被特异性上调或下调的 65 个基因的基因簇。

[0038] 图 5 表示用半 - 定量 RT-PCR 证实高表达基因的结果。具体地说, 通过半 - 定量 RT-PCR 检查 5 个基因 (高度分化的 12 个病例中的 AI261804, AA205444 和 AA167194, 以及低度分化的 12 个病例中的 AA676987 和 H22566) 和 GAPDH (内部对照) 的表达。微阵列信号对应于半 - 定量 RT-PCR 实验的结果。正常乳腺导管细胞制备自该微阵列所用 15 位绝经前患者的正常导管上皮细胞。MG 指的是完整的人乳腺。

[0039] 图 6 显示出半 - 定量 RT-PCR 的结果。图中显示出 A7870 在得自 (a) 12 位乳腺癌患者, (b) 乳腺癌细胞系 (HBC4, HBC5, HBL100, HCC1937, MCF7, MDA-MB-231, SKBR3, T47D, YMB1, BT-20, BT-474, BT-549, HCC1143, HCC1500, HCC1599, MDA-MB-157, MDA-MB-435S, MDA-MB-453, OCUB-F 和 ZR-75-1) 的肿瘤细胞和正常人组织中的表达水平。

[0040] 图 7 显示出 (a) 多种人组织, 以及 (b) 乳腺癌细胞系和正常的人重要器官中 A7870 转录物的 Northern 印迹分析结果。

[0041] 图 8 显示出 (a) 外源性 A7870 在经转染的 -COS7 细胞以及 (b) 外源性 A7870 在 T47D, BT-20 和 HBC5 细胞中的亚细胞定位。

[0042] 图 9 显示出使用通过随机 - 排列测验选择的 206 个基因进行的经监督的基因等级聚类分析。横行中显示出 69 个样品 (选自 IDC 患者)。在纵列的柱形中, 根据相对表达率的相似性将 97 个基因聚类至不同的分支。分支 1 和分支 2 中的基因优先以类似于低度分化型和高度分化型的表达水平被表达。

[0043] 图 10(A) 显示出使用 34 个基因进行的二维等级聚类分析结果, 所述基因是在用于

建立预测评分系统的随机 - 排列测验之后,通过分类评价和 leave-one-out 测验选择得到的。上方主分支中的基因优先在涉及淋巴结转移的病例中表达;下方分支中的基因在淋巴结阴性病例中的表达更高。图 10(B) 显示出图 7(A) 中出现的基因的强度,以将非 - 转移的(淋巴结 - 阴性)肿瘤与转移的(淋巴结 - 阳性)肿瘤分开。方形代表淋巴结 - 阳性病例;三角形表示阴性病例。17 个中空的方形表示淋巴结 - 阳性试验例,20 个中空的三角形表示淋巴结 - 阴性试验例,不将这些试验例用于建立预测评分。图 10(C) 表示转移预测评分和手术后的临床信息之间的相关性。

[0044] 详细描述

[0045] 除非另有说明,本文所用术语“一个”或“一种”指的是至少一个或至少一种。

[0046] 乳腺癌细胞一般以实体肿块的形式存在,其具有高度炎性反应并含有多种细胞组分。因此,以前公开的微阵列数据可能反映的是异源分布图。

[0047] 考虑到这些问题,本发明人通过激光微束显微解剖法(LMM)制备纯化的乳腺癌细胞群体和正常的乳腺上皮导管细胞群体,并且使用展示了 23,040 个基因的 cDNA 微阵列分析了 81 个 BRC,包括 12 个原位导管癌(DCIS)和 69 个浸润性导管癌(IDC)在基因组范围内的基因表达分布图。这些数据不仅应提供与乳腺癌发生有关的重要信息,还应有利于鉴定候选基因,所述候选基因的产物可用作诊断标记,和/或用作分子靶标以治疗乳腺癌患者和提供临床相关信息。

[0048] 本发明部分基于下述发现:在 BRC 患者的上皮细胞和癌之间,多个核酸的表达模式有所改变。使用全面的 cDNA 微阵列系统即可鉴定基因表达的差异。

[0049] 使用展示了 23,040 个基因的 cDNA 微阵列以及激光显微解剖法分析得自 81 个 BRC,包括 12 个 DCIS 和 69 个 IDC 的癌细胞的基因表达分布图。通过比较得自被诊断为 BRC 的患者的癌细胞与单纯用激光显微解剖选择的正常导管上皮细胞的表达模式,将 102 个基因(示于表 3,5 和 7)鉴定为在 BRC 细胞中一般会上调的基因,从中选出 100 个基因作为本发明的 BRC- 相关基因。类似地,将 288 个基因(示于表 4,6 和 8)鉴定为在 BRC 细胞中一般会下调的基因。另外,选择具有检测患者血清或唾液中的癌症相关蛋白质之潜能的候选分子标记,并且发现一些潜在的靶,它们可用于开发针对人 BRC 的信号 - 抑制策略。其中,表 3 和 4 提供了在包括 DCIS 和 IDC 的 BRC 与正常组织之间表达有所改变的基因列表。表 3 和表 4 分别显示了通常在 DCIS 和 IDC 中被上调或下调的基因。表 5 和表 6 分别列出了在 DCIS 转变为 IDC 时表达有所增加或降低的基因。另外,表 7 和表 8 分别列出了与正常组织相比,在 IDC 中通常被上调或下调的基因。

[0050] 本文鉴定的差异表达基因能用作 BRC 诊断标记,并可作为 BRC 基因靶以用于诊断 BRC,可改变所述基因的表达以治疗或减轻 BRC 症状。或者,本文鉴定的在 DCIS 和 IDC 之间差异表达的基因能用作区分 IDC 和 DCIS 的标记以及诊断用的 BRC 基因靶,可改变该基因的表达以治疗或减轻 IDC 症状。

[0051] 表 3-8 概述了这些在 BRC 患者中表达水平受调制(即增加或降低)的基因,本文将其统称为“BRC- 相关基因”,“BRC 核酸”或“BRC 多核苷酸”,将其相应的编码多肽称为“BRC 多肽”或“BRC 蛋白质”。除非另有说明,“BRC”指的是本文公开的任何序列(例如表 3-8 列出的 BRC- 相关基因)。对于先前已被描述过的基因,还一并提供了数据库登录号。

[0052] 通过测定细胞样品中多种基因的表达即可诊断 BRC。类似地,测定这些基因对多种

药剂作出反应后的表达,即可鉴定出治疗 BRC 的药剂。

[0053] 本发明包括检测(例如测定)表 3-8 列出的至少一种,直至所有 BRC- 相关基因的表达。使用 GenBank™ 数据库入口提供的已知序列的序列信息,用本领域普通技术人员众所周知的技术即可检测和测定 BRC- 相关基因。例如,可使用对应于 BRC- 相关基因的序列数据库入口内的序列构建探针,以在例如 Northern 印迹杂交分析中检测对应于 BRC- 相关基因的 RNA 序列。探针一般包括参照序列的至少 10 个,至少 20 个,至少 50 个,至少 100 个或至少 200 个核苷酸。作为另一个例子,可使用序列构建引物,从而在例如基于扩增的检测方法,如基于逆转录的聚合酶链反应中特异性扩增 BRC 核酸。

[0054] 然后,将受试细胞群体,如得自患者的组织样品中的一种或多种 BRC- 相关基因的表达水平与参照群体中相同基因的表达水平相比较。参照细胞群体包括一种或多种待比较参数已被获知的细胞,即乳腺导管癌细胞(如 BRC 细胞)或正常乳腺导管上皮细胞(如非-BRC 细胞)。

[0055] 与参照细胞群体相比,受试细胞群体的基因表达模式是否能表示 BRC 或对 BRC 的易感性取决于参照细胞群体的组成。例如,如果参照细胞群体由非-BRC 细胞组成,受试细胞群体和参照细胞群体之间基因表达模式的相似性表示受试细胞群体是非-BRC。反之,如果参照细胞群体由 BRC 细胞组成,受试细胞群体和参照细胞群体之间基因表达分布图的相似性表示受试细胞群体包括 BRC 细胞。

[0056] 如果与参照细胞群体中相应 BRC 标记基因的表达水平相差的倍数大于 1.1,大于 1.5,大于 2.0,大于 5.0,大于 10.0 或相差更多倍,即可认为受试细胞群体中 BRC 标记基因的表达水平是“经改变的”。

[0057] 受试细胞群体和参照细胞群体之间的差异基因表达可归一化至对照核酸,例如持家基因。例如,对照核酸是已知不会根据细胞的癌或非癌状态而改变的核酸。使用对照核酸的表达水平归一化受试和参照群体中的信号水平。例如,对照基因包括但不限于  $\beta$ -肌动蛋白、甘油醛 3-磷酸脱氢酶和核糖体蛋白 P1。

[0058] 可将受试细胞群体与多个参照细胞群体相比较。其中每个参照群体的已知参数可以不同。因此,受试细胞群体可与已知含有例如 BRC 细胞的第一个参照细胞群体和已知含有例如非-BRC 细胞(正常细胞)的第二个参照细胞群体相比较。受试细胞可包括在已知含有或疑似含有 BRC 细胞的受试者的组织型样品或细胞样品中。

[0059] 受试细胞可得自身体组织或体液,例如生物液体(如血液或唾液)。例如,受试细胞可纯化自乳腺组织。优选受试细胞群体含有上皮细胞。上皮细胞优选得自己知或疑似为乳腺导管癌的组织。

[0060] 参照细胞群体中的细胞应得自与受试细胞的来源组织类似的组织类型。任选参照细胞群体是细胞系,例如 BRC 细胞系(即阳性对照)或正常非-BRC 细胞系(即阴性对照)。或者,对照细胞群体可得自分子信息数据库,所述信息来源于所检测的参数或疾病已被获知的细胞。

[0061] 优选受试者为哺乳动物。例如,哺乳动物包括但不限于人、非-人灵长类动物、小鼠、大鼠、狗、猫、马或奶牛。

[0062] 可使用本领域已知的方法,在蛋白质或核酸水平上检测本文所公开基因的表达。例如,可以使用 Northern 杂交分析测定基因表达,所述分析使用了特异性识别这些核酸序

列中的一种或多种序列的探针。或者,可使用基于逆转录的 PCR 检测法,例如使用特异于差异表达的基因序列的引物来测定基因表达。也可以在蛋白质水平上测定表达,即通过测定本文所述基因编码的多肽水平或其生物活性来测定表达。所述方法是本领域众所周知的,包括但不限于例如免疫测定法,该方法利用了抗所述基因编码的蛋白质的抗体。由所述基因编码的蛋白质的生物活性一般是众所周知的。

#### [0063] 诊断乳腺癌

[0064] 在本发明的上下文中,通过测定受试细胞群体(即得自患者的生物样品)中一种或多种 BRC 核酸的表达水平来诊断 BRC。优选受试细胞群体含有上皮细胞,例如得自乳腺组织的细胞。也可以由血液或其它体液,如尿液测定基因表达。其它生物样品可用于测定蛋白质水平。例如,可通过免疫测定法或其它常规的生物学试验测定待诊断受试者的血液或血清中的蛋白质水平。

[0065] 测定受试细胞或生物样品中一种或多种 BRC- 相关基因,如表 3-8 列出的基因的表达,并与和一种或多种受试 BRC- 相关基因有关的正常对照表达水平相比较。正常对照水平是通常在已知未患 BRC 的人群中发现的 BRC- 相关基因的表达分布图。在得自患者的组织样品中,一种或多种 BRC- 相关基因表达水平的改变(例如增加)表示受试者患有 BRC 或有发展成为 BRC 的风险。例如,与正常对照水平相比,受试群体中由表 3,5 和 7 列出的一种或多种上调 BRC- 相关基因表达的增加表示受试者患有 BRC 或有发展成为 BRC 的风险。反之,与正常对照水平相比,受试群体中由表 4,6 和 8 列出的一种或多种下调 BRC- 相关基因表达的降低表示受试者患有 BRC 或有发展成为 BRC 的风险。

[0066] 与正常对照水平相比,受试群体中一种或多种 BRC- 相关基因的改变表示受试者患有 BRC 或有发展成为 BRC 的风险。例如,一系列 BRC- 相关基因(表 3-8 列出的基因)中至少 1%,至少 5%,至少 25%,至少 50%,至少 60%,至少 80%,至少 90% 或更多的改变表示受试者患有 BRC 或有发展成为 BRC 的风险。

#### [0067] 鉴定 BRC 的组织病理学分化:

[0068] 本发明提供了鉴定受试者中 BRC 的组织病理学分化的方法,所述方法包括下述步骤:

[0069] (a) 检测收集自被治疗的受试者的组织样品中一种或多种标记基因的表达水平,其中一种或多种标记基因选自表 1 和 10 列出的基因;和

[0070] (b) 将检测到的一种或多种标记基因的表达水平与和高度分化的病例及低度分化的病例相关的表达水平相比较;

[0071] (c) 以使当检测到的一种或多种标记基因的表达水平与高度分化的病例类似时,将组织样品确定为高度分化型,而当检测到的一种或多种标记基因的表达水平与低度分化的病例类似时,将组织样品确定为低度分化型。

[0072] 在本发明中,用于鉴定 BRC 的组织病理学分化的标记基因可以是选自表 1 和 10 所示的 231 个基因中的至少一个基因。基因的核苷酸序列和由其编码的氨基酸序列是本领域已知的。基因的登录号可参见表 1 和 10。

#### [0073] 鉴定抑制或增强 BRC- 相关基因表达的药剂:

[0074] 通过使表达 BRC- 相关上调基因的受试细胞群体与受试药剂接触,然后测定 BRC- 相关基因的表达水平或其基因产物的活性,即可鉴定出能抑制 BRC- 相关基因的表

达或其基因产物活性的药剂。与缺乏受试药剂时的表达或活性水平相比,当存在药剂时,BRC- 相关基因的表达水平或其基因产物的活性水平的降低表示该药剂是 BRC- 相关上调基因的抑制剂,并可用于抑制 BRC。

[0075] 或者,通过使表达 BRC- 相关基因的受试细胞群体与受试药剂接触,然后测定 BRC- 相关下调基因的表达水平或活性,即可鉴定出能增强 BRC- 相关下调基因的表达或其基因产物活性的药剂。与缺乏受试药剂时的表达或活性水平相比,BRC- 相关基因的表达水平或其基因产物的活性水平的增加表示该受试药剂增强了 BRC- 相关下调基因的表达或其基因产物的活性。

[0076] 受试细胞群体可以是任何表达 BRC- 相关基因的细胞。例如,受试细胞群体可含有上皮细胞,例如得自乳腺组织的细胞。另外,受试细胞可以是得自癌细胞的无限增殖细胞系。或者,受试细胞可以是被 BRC- 相关基因转染的细胞,或是被得自 BRC- 相关基因的、与报道基因可操作相连的调节序列(例如启动子序列)转染的细胞。

[0077] 评价在受试者中治疗 BRC 的效力:

[0078] 本文所鉴定的差异表达的 BRC- 相关基因也能监测 BRC 的疗程。在此方法中,由接受 BRC 治疗的受试者提供受试细胞群体。必要时,在治疗之前、当中和 / 或之后的不同时间点由受试者获得受试细胞群体。然后测定细胞群体中一种或多种 BRC- 相关基因的表达,并与包括 BRC 状态已知的细胞的参照细胞群体相比较。在本发明的上下文中,参照细胞应未曾暴露于有意义的治疗之中。

[0079] 如果参照细胞群体不含有 BRC 细胞,受试细胞群体和参照细胞群体中 BRC- 相关基因表达的相似性表示受关注的治疗是有效的。然而,受试细胞群体和正常对照参照细胞群体中 BRC- 相关基因表达的差异表示临床效果或预后较差。类似地,如果参照细胞群体含有 BRC 细胞,受试细胞群体和参照细胞群体中 BRC- 相关基因表达之间的差异表示受关注的治疗是有效的,而受试群体和癌症对照参照细胞群体中 BRC- 相关基因表达的相似性表示临床效果或预后较差。

[0080] 另外,可将在治疗后获自受试者的生物样品中测定的一种或多种 BRC- 相关基因的表达水平(即治疗后的水平)与在开始治疗之前获自受试者的生物样品中测定的一种或多种 BRC- 相关基因的表达水平(即治疗前的水平)相比较。如果 BRC- 相关基因是上调基因,治疗后样品中表达水平的降低表示受关注的治疗是有效的,而治疗后样品中表达水平的增加或维持则表示临床效果或预后较差。反之,如果 BRC- 相关基因是下调基因,治疗后样品中表达水平的增加表示受关注的治疗是有效的,而治疗后样品中表达水平的降低或维持则表示临床效果或预后较差。

[0081] 本文所用术语“有效的”表示治疗导致受试者体内因病而被上调的基因的表达降低,因病而被下调的基因的表达增加,或乳腺导管癌的大小、流行性或转移潜力降低。当将有意义的治疗用于预防时,术语“有效的”指的是治疗能延迟或防止乳腺肿瘤的形成,或者能延迟、防止或减轻临床 BRC 的症状。可使用标准的临床方案评估乳腺癌。

[0082] 另外,可与任何已知的诊断或治疗 BRC 的方法相关联来测定效力。例如,通过鉴定症状的异常,如体重减轻、腹痛、背痛、食欲减退、恶心、呕吐、全身不适、虚弱和黄疸即可诊断 BRC。

[0083] 选择适用于具体个体的 BRC 治疗剂:

[0084] 个体基因组成的差异可导致其代谢多种药物的相对能力出现差异。通过使受试者细胞的基因表达模式发生变化,即从癌症状态的基因表达模式特征转变为非-癌状态的基因表达模式特征,在受试者体内被代谢以用作抗-BRC 药剂的药剂即可证明自己的效力。因此,本文公开的差异表达的 BRC- 相关基因可以使推定的 BRC 治疗或预防性抑制剂在选定受试者的受试细胞群体中得以被检测,从而测定该药剂对所述受试者而言是否是合适的 BRC 抑制剂。

[0085] 为了鉴定出适用于具体受试者的 BRC 抑制剂,可将得自该受试者的受试细胞群体暴露于治疗剂,然后测定表 3-8 所列的一种或多种 BRC- 相关基因的表达。

[0086] 在本发明方法的上下文中,受试细胞群体含有表达 BRC- 相关基因的 BRC 细胞。优选受试细胞是上皮细胞。例如,可在候选药剂的存在下保温受试细胞群体,测定受试细胞群体的基因表达模式,并与一种或多种参照分布图,例如 BRC 参照表达分布图或非-BRC 参照表达分布图进行比较。

[0087] 相对于含有 BRC 的参照细胞群体而言,受试细胞群体中表 3,5 和 7 所列的一种或多种 BRC- 相关基因表达的降低,或表 4,6 和 8 所列的一种或多种 BRC- 相关基因表达的增加表示该药剂具有治疗潜力。

[0088] 在本发明的上下文中,受试药剂可以是任何化合物或组合物。例举的受试药剂包括但不限于免疫调制剂。

[0089] 鉴定治疗剂的筛选试验:

[0090] 本文公开的差异表达的 BRC- 相关基因也可用于鉴定治疗 BRC 的候选治疗剂。本发明的方法包括筛选候选治疗剂,以测定该药剂是否能将表 3-8 所列的一种或多种 BRC- 相关基因以 BRC 状态为特征的表达分布图转变为以非-BRC 状态为特征的基因表达模式。

[0091] 在本发明的方法中,将细胞暴露于一种受试药剂或多种受试药剂(依次或联合),并测定细胞中由表 3-8 所列的一种或多种 BRC- 相关基因的表达。将在受试群体中测定的 BRC- 相关基因的表达分布图与未暴露于受试药剂的参照细胞群体中相同 BRC- 相关基因的表达水平进行比较。

[0092] 能刺激低表达基因的表达或抑制过表达基因的表达的药剂具有潜在的临床价值。可在动物或受试者中进一步检测该药剂防止乳腺导管癌生长的能力。

[0093] 在另一个实施方案中,本发明提供了筛选候选药剂的方法,所述药剂可作用于治疗 BRC 的潜在靶标。正如上文所详细讨论的,通过控制标记基因的表达水平或其基因产物的活性,即可控制 BRC 的发生和发展。因此,通过将所述表达水平和活性用作癌或非-癌状态指标的筛选方法,即可鉴定出能作用于治疗 BRC 的潜在靶标的候选药剂。在本发明的上下文中,所述筛选方法包括例如下述步骤:

[0094] a) 使受试化合物与由选自表 3,4,5,6,7 或 8 所列基因的多核苷酸编码的多肽接触;

[0095] b) 检测多肽和受试化合物之间的结合活性;和

[0096] c) 选择能结合多肽的受试化合物。

[0097] 或者,本发明的筛选方法可包括下述步骤:

[0098] a) 使候选化合物与表达一种或多种标记基因的细胞接触,其中一种或多种标记基因选自表 3,4,5,6,7 或 8 所列的基因;和

[0099] b) 选择能降低选自表 3,5 和 7 所列基因的一种或多种标记基因的表达水平,或能增加选自表 4,6 和 8 所列基因的一种或多种标记基因的表达水平的候选化合物。

[0100] 表达标记基因的细胞包括例如由 BRC 建立的细胞系;所述细胞可用于本发明的上述筛选。

[0101] 或者,本发明的筛选方法可包括下述步骤:

[0102] a) 使受试化合物与由选自表 3,4,5,6,7 或 8 所列基因的多核苷酸编码的多肽接触;

[0103] b) 检测步骤 (a) 的多肽的生物活性;和

[0104] c) 选择与缺乏受试化合物时检测的生物活性相比,能抑制由选自表 3,5 和 7 所列基因的多核苷酸所编码多肽的生物活性,或与缺乏受试化合物时检测的生物活性相比,能增强由选自表 4,6 和 8 所列基因的多核苷酸所编码多肽的生物活性的化合物。

[0105] 可使用标记基因的核苷酸序列,以重组蛋白质的形式获得可用于本发明的筛选方法的蛋白质。基于与标记基因及其编码的蛋白质有关的信息,本领域技术人员可选择蛋白质的任何生物活性作为筛选指标,并可选择任何适当的测定方法检测选定的生物活性。

[0106] 或者,本发明的筛选方法可包括下述步骤:

[0107] a) 使候选化合物与导入了载体的细胞接触,所述载体含有一种或多种标记基因的转录调节区和在所述转录调节区的控制之下表达的报告基因,其中一种或多种标记基因选自表 3,4,5,6,7 或 8 所列的基因;

[0108] b) 测定所述报告基因的表达或活性;和

[0109] c) 当所述标记基因是选自表 3,5 和 7 所列基因的上调标记基因时,选择与对照相比能降低所述报告基因的表达或活性的候选化合物,或当所述标记基因是选自表 4,6 和 8 所列基因的下调标记基因时,选择与对照相比能增强所述报告基因的表达水平的候选化合物。

[0110] 适当的报告基因和宿主细胞是本领域众所周知的。通过使用标记基因的转录调节区即可制备适用于本发明筛选方法的报道构建体。当本领域技术人员已知标记基因的转录调节区时,可通过使用以前公开的序列信息来制备报道构建体。当标记基因的转录调节区尚未被鉴定时,可基于标记基因的核苷酸序列信息从基因组文库中分离含有转录调节区的核苷酸区段。

[0111] 通过筛选分离出的化合物可用作候选药剂以开发药物,所述药物可抑制标记基因的表达或由标记基因编码的蛋白质的活性,并可用于治疗或预防乳腺癌。

[0112] 另外,通过本发明的筛选方法可以获得的化合物还包括这样一类化合物,其中通过添加、缺失和/或取代改变了抑制标记基因所编码蛋白质之活性的化合物的部分结构。

[0113] 当将通过本发明的方法分离出的化合物作为药物施用于人和其它哺乳动物,如小鼠、大鼠、豚鼠、兔、猫、狗、绵羊、猪、牛、猴、狒狒和黑猩猩时,可直接施用分离的化合物,或使用已知的药物制备方法将其配制成某种剂型。例如,根据需要,可以糖衣片剂、胶囊、酞剂和微囊剂型口服药物,或者以溶于水或任何其它药物可接受液体的无菌溶液或悬浮液注射剂形式非口服给药。例如,可以常规药物实践所需的单位剂量形式将化合物与药学可接受载体或介质混合,所述载体或介质具体地为无菌水、生理盐水、植物油、乳化剂、悬浮剂、表面活性剂、稳定剂、增味剂、赋形剂、载体、防腐剂、结合剂等。这些制品中所含活性成分的量

能提供所示范围内的适当剂量。

[0114] 可与片剂和胶囊混合的添加剂的例子包括但不限于：如明胶、玉米淀粉、黄耆胶和阿拉伯胶的结合剂；如晶体纤维素的赋形剂；如玉米淀粉、明胶和藻酸的膨胀剂；如硬脂酸镁的润滑剂；如蔗糖、乳糖或糖精的甜味剂；如胡椒薄荷、*Gaultheria adeno-thrix* 油和櫻桃的增味剂。当单位剂量形式是胶囊时，上述成分中还可包括液体载体，如油。可使用适于注射的载体，如蒸馏水，根据常规药物实践配制注射用的无菌混合物。

[0115] 生理盐水、葡萄糖和其它包括助剂的等渗液体，如 D- 山梨糖醇、D- 甘露糖、D- 甘露醇和氯化钠可用作注射用的水溶液。它们可与适当的增溶剂，如醇，例如乙醇；聚醇，如丙二醇和聚乙二醇；和非离子型表面活性剂，如 Polysorbate 80 (TM) 和 HCO-50 一起使用。

[0116] 芝麻油或豆油可用作油质液体，并可与苯甲酸苄酯或苄醇一起用作增溶剂，并可与如磷酸缓冲液和醋酸钠缓冲液的缓冲液；如盐酸普鲁卡因的止痛剂；如苄醇和酚的稳定剂；和 / 或抗氧化剂一起配制。可将制备好的注射液注入适当的安瓿中。

[0117] 可使用本领域技术人员众所周知的方法给患者施用本发明的药物组合物，例如以动脉内、静脉内或经皮注射给药或鼻内、经支气管、肌内或口服给药。给药剂量和方法随患者体重和年龄以及给药方法的变化而改变；然而，本领域技术人员可常规选择适当的给药方法。如果所述化合物可由 DNA 编码，可将 DNA 插入基因治疗载体，给患者施用所述载体以进行治疗。给药剂量和方法随患者体重、年龄和症状的变化而改变；但本领域技术人员可适当选择给药剂量和方法。

[0118] 例如，尽管与本发明的蛋白质结合并调节其活性的化合物的剂量取决于症状，但当给正常成人（体重为 60kg）口服给药时，剂量一般为约 0.1mg 至约 100mg/ 天，优选为约 1.0mg 至约 50mg/ 天，更优选为约 1.0mg 至约 20mg/ 天。

[0119] 当以注射液形式给正常成人（体重为 60kg）非肠道施用化合物时，尽管根据患者、靶器官、症状和给药方法的不同有一些差异，但合适的静脉内注射剂量为约 0.01mg 至约 30mg/ 天，优选为约 0.1 至约 20mg/ 天，更优选为约 0.1 至约 10mg/ 天。当给其它动物给药时，通过转换成 60kg 体重的量可以常规计算出适当的剂量。

[0120] 鉴定乳腺癌转移治疗剂的筛选试验：

[0121] 本发明提供了可用于治疗或预防乳腺癌转移的靶分子。可根据上文所述的针对 BRC 的方法，使用与 BRC 转移相关的标记基因进行本发明的针对 BRC 转移的筛选试验。

[0122] 在本发明中，选自表 11 所列基因的标记基因可用于筛选。表中所示的 34 个基因与淋巴结转移有关。在淋巴结 - 阳性肿瘤中，这些基因中的 25 个基因 (+) 被相对上调，9 个基因 (-) 被下调（表 11 和图 10）。通过本发明获得的能抑制一种或多种上调基因的表达或其基因产物的活性的药剂可用于治疗或预防具有淋巴结转移的 BRC。或者，通过本发明获得的能增强一种或多种下调基因的表达或其基因产物的活性的药剂也可用于治疗或预防具有淋巴结转移的 BRC。

[0123] 在本发明中，按照与鉴定能抑制或增强 BRC- 相关基因表达的药剂相同的方式，即可鉴定出能调节表 11 所列基因的表达水平的药剂。或者，按照与鉴定能抑制或增强 BRC- 相关基因产物的药剂相同的方式，也可鉴定出能调节其基因产物的活性的药剂。

[0124] 评估患有乳腺癌的受试者的预后情况：

[0125] 本发明还提供了评估患有 BRC 的受试者的预后情况的方法，所述方法包括在一系



列疾病阶段将受试细胞群体中的一种或多种 BRC- 相关基因的表达与得自患者的参照细胞群体中的相同 BRC- 相关基因的表达相比较的步骤。通过比较受试细胞群体与参照细胞群体中一种或多种 BRC- 相关基因的基因表达,或全程比较得自受试者的受试细胞群体的基因表达模式,即可评估受试者的预后情况。

[0126] 例如,与正常对照相比,一种或多种上调的 BRC- 相关基因,如表 3,5 或 7 所列基因的表达的增加,或与正常对照相比,一种或多种下调的 BRC- 相关基因,如表 4,6 或 8 所列基因的表达的降低表示预后不佳。相反,与正常对照相比,表 3-8 所列的一种或多种 BRC- 相关基因表达的相似性表示受试者的预后较好。优选通过比较选自表 3,4,5,6,7 和 8 所列基因的基因表达分布图来评估受试者的预后情况。可使用分类评分 (CS) 来比较表达分布图。

[0127] 试剂盒:

[0128] 本发明还包括 BRC- 检测试剂,如特异性结合或鉴定一种或多种 BRC 核酸的核酸,如与 BRC 核酸的一部分互补的寡核苷酸序列,或能与一种或多种由 BRC 核酸编码的蛋白质结合的抗体。可以试剂盒的形式将检测试剂包装在一起。例如,可将检测试剂包装于分开的容器中,例如,核酸或抗体(与固体基质结合,或与使其同基质结合的试剂分开包装),对照试剂(阳性和/或阴性)和/或可测标记。试剂盒中也可包括进行试验所用的说明书(如书面说明书、磁带、VCR、CD-ROM 等)。试剂盒的试验方式可以是本领域已知的 Northern 杂交或夹心 ELISA。

[0129] 例如,可将 BRC 检测试剂固定于固体基质,如多孔条上以形成至少一个 BRC 检测位点。多孔条的测定或检测区可包括多个位点,每个位点都含有核酸。测试条也可含有阴性和/或阳性对照的位点。或者,对照位点可位于与测试条不同的条上。任选不同的检测位点可含有不同量的固定化核酸,即第一个检测位点量较多,随后的位点量较少。通过添加受试样品,表现出可测信号的位点数目定量显示了样品中存在的 BRC 的量。检测位点的形状可以是任何可适当检测的形状,一般为跨越测试条宽度的条状或点状。

[0130] 或者,试剂盒可含有核酸底物微阵列,所述阵列含有一种或多种核酸。微阵列上的核酸可特异性鉴定一种或多种由表 3-8 所列 BRC- 相关基因表示的核酸序列。利用与微阵列测试条或芯片结合的水平,可鉴定出 2,3,4,5,6,7,8,9,10,15,20,25,40 或 50 种或更多种由表 3-8 所列 BRC- 相关基因表示的核酸的表达。底物微阵列可位于例如固体底物,如美国专利 5,744,305(其全部内容列入本文作为参考)描述的“芯片”上。

[0131] 微阵列和核酸混合物(pluralities):

[0132] 本发明还包括含有一种或多种核酸的核酸底物微阵列。微阵列上的核酸特异性地对应于一种或多种由表 3-8 所列 BRC- 相关基因表示的核酸序列。通过检测与微阵列结合的核酸,即可鉴定出 2,3,4,5,6,7,8,9,10,15,20,25,40 或 50 种或更多种由表 3-8 所列 BRC- 相关基因表示的核酸的表达水平。

[0133] 本发明还包括分离的核酸混合物(即两种或多种核酸的混合物)。核酸可以在液相或固相中,例如固定于如硝酸纤维素膜的固体支持物上。核酸混合物包括一种或多种由表 3-8 所列 BRC- 相关基因表示的核酸。在多个实施方案中,核酸混合物包括 2,3,4,5,6,7,8,9,10,15,20,25,40 或 50 种或更多种由表 3-8 所列 BRC- 相关基因表示的核酸。

[0134] 抑制乳腺癌的方法:

[0135] 本发明还提供了通过降低表 3,5 和 7 所列的一种或多种 BRC- 相关基因的表达(或

其基因产物的活性),或增加表 4,6 和 8 所列的一种或多种 BRC- 相关基因的表达(或其基因产物的活性)以治疗或减轻受试者的 BRC 症状的方法。可以给患有 BRC 或有发展成 BRC 的风险(或对 BRC 易感)的受试者预防性或治疗性地施用适当的治疗化合物。使用标准的临床方法或通过检测表 3-8 所列的一种或多种 BRC- 相关基因的异常表达水平或其基因产物的异常活性,即可鉴定出所述受试者。在本发明的上下文中,适当的治疗剂包括例如细胞周期调节、细胞增殖和蛋白质激酶活性的抑制剂。

[0136] 本发明的治疗方法包括增加一种或多种基因产物的表达、功能或同时增加这两者的步骤,相对于获得 BRC 细胞的同一组织类型的正常细胞而言,所述基因在 BRC 细胞中的表达降低(“下调”或“低表达”基因)。在这些方法中,用有效量的能增加受试者体内一种或多种低表达(下调)基因量的化合物治疗受试者。可以全身性给药或局部给药。适当的治疗化合物包括低表达基因的多肽产物,其生物活性片段,以及编码低表达基因并具有允许在 BRC 细胞中表达的表达控制元件的核酸;例如,能增加对 BRC 细胞而言为内源性的基因的表达水平的药剂(即上调低表达基因的表达)。施用该化合物能抵销受试者的乳腺细胞中异常低表达基因的作用,并能改善受试者的临床状况。

[0137] 或者,本发明的治疗方法包括降低在乳腺细胞中表达异常增加的基因(“上调”或“过表达”基因)的一种或多种基因产物的表达、功能或同时降低这两者的步骤。可按照本领域已知的几种方法中的任一种来抑制表达。例如,通过给受试者施用能抑制或拮抗过表达基因之表达的核酸,如能破坏过表达基因之表达的反义寡核苷酸或小的干扰 RNA 来抑制表达。

[0138] 反义核酸:

[0139] 如上所述,可使用对应于表 3,5 和 7 所列 BRC- 相关基因的核苷酸序列的反义核酸来降低基因的表达水平。对应于在乳腺癌中被上调的表 3,5 和 7 所列 BRC- 相关基因的反义核酸可用于治疗乳腺癌。具体地说,本发明的反义核酸可通过下述步骤起作用:同表 3,5 和 7 所列 BRC- 相关基因或其对应的 mRNA 结合,从而抑制基因的转录或翻译,促进 mRNA 降解和/或抑制表 3,5 和 7 所列 BRC- 相关基因所编码蛋白质的表达,从而抑制蛋白质的功能。本文所用术语“反义核酸”包含与靶序列完全互补的核苷酸以及具有一个或多个核苷酸错配的核苷酸,只要反义核酸能与靶序列特异性杂交即可。例如,本发明的反义核酸包括与跨越至少 15 个连续核苷酸的区域的高源性至少为 70%或更高,优选至少为 80%或更高,更优选至少为 90%或更高,甚至更优选至少为 95%或更高的多核苷酸。可使用本领域已知的算法测定高源性。

[0140] 本发明的反义核酸通过下述机制对产生由 BRC- 相关标记基因编码的蛋白质的细胞发挥作用,即与编码蛋白质的 DNA 或 mRNA 结合,抑制其转录或翻译,促进 mRNA 降解并抑制蛋白质的表达,从而导致蛋白质功能受抑制。

[0141] 通过与对核酸无活性的适当基质混合,可将本发明的反义核酸制成外用制剂,如涂抹剂(liniment)或泥罨敷剂(poultice)。

[0142] 必要时,也可通过加入赋形剂、等渗剂、增溶剂、稳定剂、防腐剂、止痛剂等,将本发明的反义核酸配制成片剂、粉剂、颗粒剂、胶囊、脂质体胶囊、注射剂、溶液、滴鼻剂和冻干剂。通过下述的已知方法即可制备所述制剂。

[0143] 通过直接施用于病痛部位或注射至血管以使其到达病痛部位,将本发明的反义核

酸施用给患者。也可使用反义核酸-固定介质增加持久性和膜-通透性。所述介质包括但不限于脂质体、聚-L-赖氨酸、脂质、胆固醇、脂质转染试剂,或其衍生物。

[0144] 可根据患者的身体状况适当调整本发明的反义核酸衍生物的剂量,并按所需的量使用所述剂量。例如,可以施用的剂量范围为 0.1 至 100mg/kg,优选为 0.1 至 50mg/kg。

[0145] 本发明的反义核酸抑制本发明蛋白质的表达,从而可用于抑制本发明蛋白质的活性。另外,含有本发明的反义核酸的表达-抑制剂是有用的,因为它们可抑制本发明蛋白质的生物活性。

[0146] 本发明的方法可用于改变细胞中上调的 BRC- 相关基因的表达,例如由细胞的恶性转化引起的上调。在靶细胞中使 siRNA 与对应于表 3,5 和 7 所列 BRC- 相关基因之一的转录物结合,导致细胞产生的蛋白质减少。寡核苷酸的长度至少为 10 个核苷酸,可以与天然转录物一样长。优选寡核苷酸的长度为 19-25 个核苷酸。最优选寡核苷酸的长度少于 75, 50, 25 个核苷酸。

[0147] 本发明的反义核酸包括经修饰的寡核苷酸。例如,可使用硫代寡核苷酸赋予寡核苷酸以核酶抗性。

[0148] 还可使用针对标记基因的 siRNA 降低标记基因的表达水平。本文中的术语“siRNA”指的是能防止靶 mRNA 翻译的双链 RNA 分子。可使用将 siRNA 导入细胞的标准技术,包括将 DNA 用作模板,由其转录 RNA 的技术。在本发明的上下文中,siRNA 含有针对上调标记基因,如表 3,5 和 7 所列 BRC- 相关基因的有义核酸序列和反义核酸序列。构建 siRNA 使得单个转录物具有来自靶基因的有义和互补反义序列,如发夹结构。

[0149] 如表 3,5 和 7 所列的 BRC- 相关基因的 siRNA 与靶 mRNA 杂交,从而通过与通常为单链的 mRNA 转录物结合,籍此干扰翻译,进而干扰蛋白质的表达,来降低或抑制由表 3,5 和 7 所列 BRC- 相关基因编码的多肽的产生。在本发明的上下文中,siRNA 的长度优选少于 500, 200, 100, 50 或 25 个核苷酸。更优选 siRNA 的长度为 19-25 个核苷酸。例举的用于产生 TOPK siRNA 的核酸序列包括作为靶序列的 SEQ ID NO:25, 28 和 31 的核苷酸序列。为了增强 siRNA 的抑制活性,可在靶序列反义链的 3' 末端添加核苷酸“u”。添加的“u”的数目至少为 2,一般为 2 至 10,优选为 2 至 5。添加的“u”在 siRNA 反义链的 3' 末端形成单链。

[0150] 可以能结合 mRNA 转录物的形式将如表 3,5 和 7 所列的 BRC- 相关基因的 siRNA 直接导入细胞。或者,可用载体承载编码 siRNA 的 DNA。

[0151] 例如,通过以允许两条链表达(通过转录 DNA 分子)的方式,将 BRC- 相关基因的靶序列克隆至表达载体中即可产生载体,所述表达载体中具有侧翼于该序列的可操作相连的调节序列 (Lee, N. S., Dohjima, T., Bauer, G., Li, H., Li, M.-J., Ehsani, A., Salvaterra, P., and Rossi, J. (2002) Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nature Biotechnology* 20:500-505.)。通过第一个启动子(例如克隆 DNA 3' 端的启动子序列)转录对 BRC- 相关基因的 mRNA 而言为反义的 RNA 分子,通过第二个启动子(例如克隆 DNA 5' 端的启动子序列)转录对 BRC- 相关基因的 mRNA 而言为有义链的 RNA 分子。有义和反义链在体内杂交以产生 siRNA 构建体,用于沉默 BRC- 相关基因。或者,可利用两个构建体产生 siRNA 构建体的有义和反义链。克隆的 BRC- 相关基因可编码具有二级结构,如发夹的构建体,其中单个转录物具有得自靶基因的有义序列和互补的反义序列。

[0152] 由任意核苷酸序列组成的环序列可位于有义和反义序列之间以形成发夹环结构。因此,本发明还提供了具有通式 5' -[A]-[B]-[A']-3' 的 siRNA,其中 [A] 是与选自表 3, 5 或 7 的基因的序列相对应的核糖核苷酸序列, [B] 是由 3 至 23 个核苷酸组成的核糖核苷酸序列, [A'] 是由 [A] 的互补序列组成的核糖核苷酸序列。区域 [A] 与 [A'] 杂交,然后形成由区域 [B] 组成的环。环序列的长度优选为 3 至 23 个核苷酸。环序列可选自例如下述序列 ([http://www.ambion.com/techlib/tb/tb\\_506.html](http://www.ambion.com/techlib/tb/tb_506.html))。另外,由 23 个核苷酸组成的环序列也可提供有活性的 siRNA (Jacque, J. -M., Triques, K., and Stevenson, M. (2002) Modulation of HIV-1 replication by RNA interference. Nature 418 :435-438.)。

[0153] CCC, CCACC 或 CCACACC :Jacque, J. M, Triques, K., and Stevenson, M (2002) Modulation of HIV-1 replication by RNA interference. Nature, Vol. 418 :435-438。

[0154] UUCG :Lee, N. S., Dohjima, T., Bauer, G., Li, H., Li, M. -J., Ehsani, A., Salvaterra, P., and Rossi, J. (2002) Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. Nature Biotechnology 20 :500-505. Fruscoloni, P., Zamboni, M., and Tocchini-Valentini, G. P. (2003) Exonucleolytic degradation of double-stranded RNA by an activity in *Xenopus laevis* germinal vesicles. Proc. Natl. Acad. Sci. USA 100 (4) :1639-1644。

[0155] UUCAAGAGA :Dykxhoorn, D. M., Novina, C. D., and Sharp, P. A. (2002) Killing the messenger :Short RNAs that silence gene expression. Nature Reviews Molecular Cell Biology 4 :457-467。

[0156] 因此,环序列可选自 CCC, UUCG, CCACC, CCACACC 和 UUCAAGAGA。优选的环序列是 UUCAAGAGA (DNA 为 “ttcaagaga”)。适用于本发明上下文的发夹 siRNA 包括例如:

[0157] 对 TOPK-siRNA 而言

[0158] gaacgauauaaagccagcc-[b]-ggcuggcuuuauaucguuc (针对 SEQ ID NO :25 的靶序列);

[0159] cuggaugaucauaccaga-[b]-ucugguaucauccag (针对 SEQ ID NO :28 的靶序列);

[0160] guguggcuugcguaaaauaa-[b]-uuauuuacgcaagccacac (针对 SEQ ID NO :31 的靶序列)

[0161] 使用可得自 Ambion 网站 ([http://www.ambion.com/techlib/misc/siRNA\\_finder.html](http://www.ambion.com/techlib/misc/siRNA_finder.html)) 的 siRNA 设计计算机程序设计适当 siRNA 的核苷酸序列。计算机程序基于下述方案选择 siRNA 合成所用的核苷酸序列。

[0162] 选择 siRNA 靶位点:

[0163] 1. 自目标转录物的 AUG 起始密码子开始,扫描下游的 AA 二核苷酸序列。记录每个 AA 和 3' 方向邻接的 19 个核苷酸的出现,以作为潜在的 siRNA 靶位点。Tuschl 等不推荐针对 5' 和 3' 非翻译区 (UTR) 和起始密码子附近的区域 (75 个碱基以内) 设计 siRNA,因为所述区域中的调节蛋白结合位点较多。UTR- 结合蛋白和 / 或翻译起始复合物可干扰 siRNA 内切核酸酶复合物的结合。

[0164] 2. 将潜在的靶位点与人基因组数据库进行比较,无需考虑任何与其它编码序列具有显著同源性的靶序列。可使用能在 NCBI 服务器 ([www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)) 中找

到的 BLAST 进行同源性检索。

[0165] 3. 选择合格的靶序列用于合成。在 Ambion 网站, 优选沿着需评价之基因的长度选择几个靶序列。

[0166] 侧翼于 BRC- 相关基因序列的调节序列可以相同或不同, 以独立地或按时序或空间来调制其表达。通过将 BRC- 相关基因模板分别克隆至载体即可在细胞内转录 siRNA, 所述载体含有例如小的核 RNA (snRNA) U6 的 RNA polIII 转录单位或人 H1 RNA 启动子。为了将载体导入细胞, 可使用转染 - 增强剂。FuGENE (Rochediagnostices), Lipofectamin 2000 (Invitrogen), Oligofectamin (Invitrogen) 和 Nucleofactor (Wako pure Chemical) 可用作转染 - 增强剂。

[0167] 本发明的反义寡核苷酸或 siRNA 能抑制本发明多肽的表达, 从而可用于抑制本发明多肽的生物活性。含有本发明的反义寡核苷酸或 siRNA 的表达 - 抑制剂也可用于抑制本发明多肽的生物活性。因此, 含有本发明的反义寡核苷酸或 siRNA 的组合物可用于治疗乳腺癌。

[0168] 抗体:

[0169] 或者, 可通过施用能与基因产物结合或要不然能抑制其功能的化合物来抑制在 BRC 中过表达的基因的一种或多种基因产物的功能。例如, 化合物可以是能结合一种或多种过表达基因产物的抗体。

[0170] 本发明提及了抗体, 特别是抗上调标记基因所编码蛋白质的抗体, 或所述抗体的片段的用途。本文所用术语“抗体”指的是具有特殊结构的免疫球蛋白分子, 它仅与用于合成抗体的抗原 (即上调标记基因的基因产物) 或与之密切相关的抗原相互作用 (即结合)。另外, 抗体可以是抗体片段或经修饰的抗体, 只要它们能结合一种或多种由标记基因编码的蛋白质即可。例如, 抗体片段可以是 Fab, F(ab')<sub>2</sub>, Fv 或单链 Fv (scFv), 其中通过适当接头连接得自 H 和 L 链的 Fv 片段 (Huston et al., Proc Natl Acad Sci USA 85: 5879-5883 (1988))。更具体地, 通过用酶, 如木瓜蛋白酶或胃蛋白酶处理抗体即可产生抗体片段。或者, 可构建编码抗体片段的基因, 将其插入表达载体, 并在适当宿主细胞中表达 (参见例如 Co et al., J Immunol 152: 2968-2976 (1994); Better and Horwitz, Methods Enzymol 178: 476-496 (1989); Pluckthun and Skerra, Methods Enzymol 178: 497-515 (1989); Lamoyi, Methods Enzymol. 121: 652-663 (1986); Rousseaux et al., Methods Enzymol 121: 663-669 (1986); Bird and Walker, Trends Biotechnol 9: 132-137 (1991))。

[0171] 通过与多种分子, 如聚乙二醇 (PEG) 偶联即可修饰抗体。本发明提供了这种经修饰的抗体。通过对抗体进行化学修饰即可获得经修饰的抗体。这些修饰方法是本领域的常规技术。

[0172] 或者, 抗体可包括具有非人抗体可变区和人抗体恒定区的嵌合抗体, 或包括人源化抗体, 其含有非人抗体互补决定区 (CDR)、人抗体框架区 (FR) 和恒定区。使用已知技术即可制备出所述抗体。

[0173] 经过临床开发以及一系列抗癌药物的获准, 已证实针对癌细胞中出现的特定分子改变的癌症疗法, 所述抗癌药物如治疗晚期乳腺癌的 trastuzumab (Herceptin), 治疗慢性骨髓性白血病的 imatinib methylate (Gleevec), 治疗非 - 小细胞肺癌 (NSCLC) 的

gefitinib(Iressa) 和治疗 B- 细胞淋巴瘤和外套细胞淋巴瘤的 rituximab( 抗 -CD20mAb) (Ciardiello F, Tortora G. A novel approach in the treatment of cancer :targeting the epidermal growth factor receptor. ClinCancer Res. 2001 Oct ;7(10) :2958-70. Review. ;Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that over-expresses HER2. N Engl J Med. 2001 Mar 15 ;344(11) :783-92. ;Rehwald U, Schulz H, Reiser M, Sieber M, Staak JO, Morschhauser F, Driessen C, Rudiger T, Muller-Hermelink K, Diehl V, Engert A. Treatment of relapsed CD20+Hodgkin lymphoma with the monoclonal antibody rituximab is effective and well tolerated :results of a phase 2 trial of the German Hodgkin Lymphoma Study Group. Blood. 2003 Jan 15 ;101(2) :420-424. ;Fang G, Kim CN, Perkins CL, Ramadevi N, Winton E, Wittmann S and Bhalla KN. (2000). Blood, 96, 2246-2253. )。与传统的抗癌药剂相比, 这些药物在临床上是有有效的, 并且能被较好地耐受, 原因在于它们仅靶向转化细胞。因此, 所述药物不仅能改善癌症患者的存活和 生活质量, 也能证实分子靶向癌症疗法的概念。另外, 当联合使用时, 靶向药物可增强标准化学疗法的效力 (Gianni L. (2002). Oncology, 63 Suppl 1, 47-56. ; Klejman A, Rushen L, Morrione A, Slupianek A and Skorski T. (2002). Oncogene, 21, 5868-5876. )。因此, 未来的癌症治疗可能会将常规药物与针对肿瘤细胞的不同特征, 如血管生成和侵入的靶 - 特异性药剂联合使用。

[0174] 这些调制方法可来自体内地进行或在体外 (如通过在药剂的存在下培养细胞) 或体内 (如通过给受试者施用药剂) 进行。所述方法包括施用蛋白质或蛋白质组合或核酸分子或核酸分子组合作为疗法, 以抵制差异表达基因的异常表达或其基因产物的异常活性。

[0175] 可以用拮抗 (即降低或抑制) 过表达基因之活性的治疗剂来治疗特征在于基因和基因产物的表达水平或生物活性 (相对于未患病的受试者而言) 分别有所增加的疾病。可以治疗性或预防性地施用能拮抗活性的治疗剂。

[0176] 因此, 可用于本发明上下文的治疗剂包括例如 (i) 过表达或低表达基因的多肽或其类似物、衍生物、片段或同系物 ;(ii) 抗过表达基因或基因产物的抗体 ;(iii) 编码过表达或低表达基因的核酸 ;(iv) “功能异常的” 反义核酸或核酸 (即在一个或多个过表达基因的核酸内进行异源插入所致) ;(v) 小的干扰 RNA (siRNA) ;或 (vi) 调制剂 (即改变过表达或低表达多肽与其结合配对物之间的相互作用的抑制剂、激动剂和拮抗剂)。功能异常的反义分子被用于通过同源重组来 “敲除” 多肽的内源性功能 (参见例如 Capecchi, Science 244 :1288-1292 1989)。

[0177] 可以用能增加活性的治疗剂 (即为活性的激动剂) 治疗特征在于 (相对于未患病的受试者而言) 生物活性有所降低的疾病。可以治疗性或预防性的方式施用能上调活性的治疗剂。可以使用的治疗剂包括但不限于能增加生物利用度的多肽 (或其类似物、衍生物、片段或同系物) 或激动剂。

[0178] 通过获得患者的组织样品 (例如得自活检组织), 并在体外检测样品中的 RNA 或肽水平、表达肽 (或表达有所改变的基因的 mRNA) 的结构和 / 或活性来定量肽和 / 或 RNA, 即可容易地检测增加或降低的水平。本领域众所周知的方法包括但不限于免疫测定法 (例

如通过 Western 印迹分析、免疫沉淀后进行十二烷基硫酸钠 (SDS) 聚丙烯酰胺凝胶电泳、免疫细胞化学等) 和 / 或检测 mRNA 表达的杂交试验 (例如 Northern 试验、点印迹、原位杂交等)。

[0179] 在疾病的明显临床症状表现出来之前进行预防性给药, 从而预防疾病或延迟其进程。

[0180] 本发明的治疗方法可包括使细胞与能调制差异表达基因的一种或多种基因产物活性的药剂接触的步骤。能调制蛋白质活性的药剂的例子包括但不限于核酸、蛋白质、所述蛋白质的天然关连配体、肽、肽模拟物和其它小分子。例如, 适当的药剂可刺激一种或多种差异低表达基因的一种或多种蛋白质活性。

[0181] 抗乳腺癌预防接种:

[0182] 本发明还涉及治疗或预防受试者的乳腺癌的方法, 其包括给所述受试者施用疫苗的步骤, 所述疫苗含有由选自表 3, 5 和 7 所列 BRC- 相关基因 (即上调基因) 的核酸编码的多肽, 所述多肽的免疫活性片段, 或编码所述多肽的多核苷酸或其片段。施用多肽能在受试者体内诱导抗 - 肿瘤免疫力。为了诱导抗 - 肿瘤免疫力, 可给有此需要的受试者施用由选自表 3, 5 和 7 所列 BRC- 相关基因的核酸编码的多肽, 所述多肽的免疫活性片段, 或编码所述多肽或其片段的多核苷酸。另外, 由选自表 5 和 7 所列 BRC- 相关基因的核酸编码的多肽可分别诱导防止乳腺癌和 IDC 侵入的抗肿瘤免疫力。多肽或其免疫活性片段可用作抗 BRC 的疫苗。在一些情况下, 可以与 T 细胞受体 (TCR) 结合的形式或由抗原呈递细胞 (APC), 如巨噬细胞、树状细胞 (DC) 或 B- 细胞呈递的方式施用蛋白质或其片段。由于 DC 强的抗原呈递能力, 在 APC 中最优选使用 DC。

[0183] 在本发明中, 抗 BRC 疫苗指的是通过给动物接种能诱导抗 - 肿瘤免疫力的物质。根据本发明, 由表 3, 5 和 7 所列 BRC- 相关基因编码的多肽或其片段被认为是受限于 HLA-A24 或 HLA-A\*0201 的表位肽, 它可诱导针对表达表 3, 5 和 7 所列 BRC- 相关基因之 BRC 细胞的强有力的特异性免疫应答。因此, 本发明还包括使用多肽诱导抗 - 肿瘤免疫力的方法。一般说来, 抗 - 肿瘤免疫力包括如下所述的免疫应答:

[0184] - 诱导抗肿瘤的细胞毒淋巴细胞,

[0185] - 诱导识别肿瘤的抗体, 和

[0186] - 诱导产生抗 - 肿瘤的细胞因子。

[0187] 因此, 当将某种蛋白质接种于动物可诱导任一种免疫应答时, 就可以确定该蛋白质具有诱导抗 - 肿瘤免疫力的作用。通过观察宿主免疫系统在体内或体外针对蛋白质的反应, 即可检测出蛋白质诱导的抗 - 肿瘤免疫力。

[0188] 例如, 检测细胞毒 T 淋巴细胞诱导的方法是众所周知的。具体地说, 进入活体的外源物质通过抗原呈递细胞 (APC) 的作用被呈递至 T 细胞和 B 细胞。以抗原特异性方式对 APC 呈递的抗原作出反应的 T 细胞因受抗原的刺激而分化成细胞毒 T 细胞 (或细胞毒 T 淋巴细胞; CTL), 然后进行增殖 (也称作 T 细胞激活)。因此, 通过用 APC 将某种肽呈递至 T 细胞, 并检测 CTL 的诱导即可评价肽导致的 CTL 诱导。另外, APC 具有激活 CD4+T 细胞、CD8+T 细胞、巨噬细胞、嗜酸性粒细胞和 NK 细胞的作用。由于 CD4+T 细胞和 CD8+T 细胞对抗 - 肿瘤免疫力也很重要, 因此, 使用这些细胞的激活作用作为标志, 即可评价肽的抗 - 肿瘤免疫力诱导作用。

[0189] 评价以树状细胞 (DC) 作为 APC 的 CTL 诱导作用的方法是本领域众所周知的。DC 是具有最强 CTL 诱导作用的代表性 APC。在此方法中,起初使受试多肽与 DC 接触,然后使该 DC 与 T 细胞接触。与 DC 接触后检测到具有针对感兴趣细胞的细胞毒作用的 T 细胞,这表明受试多肽具有诱导细胞毒 T 细胞的活性。例如,使用  $^{51}\text{Cr}$ - 标记的肿瘤细胞的裂解作为标志,可以检测抗肿瘤的 CTL 活性。或者,使用  $^3\text{H}$ - 胸苷摄取活性或 LDH (乳糖脱氢酶)- 释放作为标志评价肿瘤细胞损害程度的方法也是众所周知的。

[0190] 除了 DC 外,也可将外周血单核细胞 (PBMC) 用作 APC。据报道,在 GM-CSF 和 IL-4 的存在下培养 PBMC 可以增强 CTL 诱导。类似地,已证实在匙孔血蓝蛋白 (KLH) 和 IL-7 的存在下培养 PBMC 可以诱导 CTL。

[0191] 经这些方法证实具有 CTL 诱导活性的受试多肽是具有 DC 激活作用继而具有 CTL 诱导活性的多肽。因此,能诱导针对肿瘤细胞的 CTL 的多肽可用作抗肿瘤疫苗。另外,通过与多肽接触获得诱导抗肿瘤 CTL 的能力的 APC 可用作抗肿瘤疫苗。另外,因 APC 呈递多肽抗原而获得细胞毒性的 CTL 也可用作抗肿瘤疫苗。使用 APC 和 CTL 所导致的抗 - 肿瘤免疫力的肿瘤治疗方法被称为细胞免疫治疗。

[0192] 通常,当使用多肽进行细胞免疫治疗时,已知混合多种具有不同结构的多肽并将它们与 DC 接触可以增加 CTL- 诱导的效率。因此,当用蛋白质片段刺激 DC 时,使用多种类型的片段的混合物较为有利。

[0193] 或者,通过观察对抗肿瘤抗体产生的诱导可以进一步证实多肽对抗 - 肿瘤免疫力的诱导。例如,当在用多肽免疫的实验室动物中诱导出抗多肽的抗体时,以及在肿瘤细胞的生长受所述抗体抑制时,多肽能够诱导抗 - 肿瘤免疫力。

[0194] 通过施用本发明的疫苗可诱导抗 - 肿瘤免疫力,抗 - 肿瘤免疫力的诱导可治疗和预防 BRC。抗癌疗法或对癌症发生的预防包括下述步骤中的任何一种,如对癌细胞生长的抑制、癌症的退化和对癌症发生的抑制。癌症的治疗或预防还包括患癌个体死亡率和发病率的降低、血液中肿瘤标记水平的降低、与癌症相伴随的可测症状的减轻等。优选所述治疗和预防作用在统计学上是显著的。例如,与未施用疫苗的对照相比,优选在 5% 或更低的显著性水平上观察到疫苗抗细胞增殖性疾病的治疗或预防作用。例如,Student' s t- 试验、Mann-Whitney U- 试验或 ANOVA 可用于统计学分析。

[0195] 上述具有免疫活性的蛋白质或编码该蛋白质的载体可与佐剂混合。佐剂指的是当与具有免疫活性的蛋白质一起 (或相继) 施用能增强抗蛋白质的免疫应答的化合物。佐剂的例子包括但不限于霍乱毒素、沙门氏菌毒素、明矾等,但并不局限于此。另外,本发明的疫苗可与药物可接受载体适当混合。所述载体的例子包括无菌水、生理盐水、磷酸缓冲液、培养液等。另外,必要时疫苗还可含有稳定剂、悬浮剂、防腐剂、表面活性剂等。可全身或局部施用疫苗。可通过单次给药施用疫苗,或通过多次给药加强免疫。

[0196] 当使用 APC 或 CTL 作为本发明的疫苗时,可通过例如来自体内的方法治疗或预防肿瘤。更具体地,收集接受治疗或预防的受试者的 PBMC,使来自体内的细胞与多肽接触,诱导 APC 或 CTL 之后,可给受试者施用细胞。也可通过在来自体内的 PBMC 中导入编码多肽的载体来诱导 APC。在给药前可克隆在体外诱导的 APC 或 CTL。通过克隆和培养具有较高靶细胞损害活性的细胞,可更有效地进行细胞免疫治疗。另外,按此方式分离的 APC 和 CTL 不仅可用于针对细胞所来源的个体的细胞免疫治疗,还可用于针对来自其它个体的相似类型



肿瘤的细胞免疫治疗。

[0197] 另外,本发明提供了用于治疗或预防细胞增殖性疾病,如癌症的药物组合物,其含有药物有效量的本发明多肽。药物组合物可用于产生抗肿瘤免疫力。

[0198] 抑制 BRC 或恶性 BRC 的药物组合物:

[0199] 在本发明的上下文中,适当的药剂包括适于口服、直肠、鼻、局部(包括颊和舌下)、阴道或非肠道(包括肌内、皮下和静脉内)给药或经吸入或吹入给药的药剂。优选静脉内给药。任选以分开的剂量单位包装药剂。

[0200] 适于口服给药的药剂包括各含预定量活性成分的胶囊、扁囊剂或片剂。适当的制剂还包括粉剂、颗粒剂、溶液、悬浮液和乳剂。任选以大丸剂、糖剂或糊剂的形式施用活性成分。口服给药的片剂和胶囊可含有常规的赋形剂,如结合剂、填充剂、润滑剂、分解剂和/或湿润剂。任选通过用一种或多种配方成分压制或塑型来制备片剂。通过在适当的机器中压制可自由-流动(如粉末状或颗粒状)的活性成分,即可制备经压制的片剂,其中所述活性成分任选与结合剂、润滑剂、无活性的稀释剂、润滑剂、表面活性剂和/或分散剂混合。通过在适当的机器中对已用无活性的液体稀释剂湿润的粉末状化合物的混合物进行塑型即可制备出成型的片剂。可根据本领域众所周知的方法包被片剂。口服液体制剂可以采取例如含水或含油悬浮液、溶液、乳剂、糖浆或酞剂的形式,也可以是使用前需用水或其它适当载体重建的干燥产品。所述液体制剂可含有常规的添加剂,如悬浮剂、乳化剂、非-水载体(可包括食用油)和/或防腐剂。可任选地配制片剂以缓释或控释其中的活性成分。片剂包装可以包括每月服用一片的片剂。

[0201] 适于非肠道给药的药剂包括含水和不含水的无菌注射溶液,其任选含有抗-氧化剂、缓冲液、抑菌剂和使制剂与欲用药的受体的血液等渗的溶质;以及含水和不含水的无菌悬浮液,其中包括悬浮剂和/或增稠剂。可用单位剂量或多-剂量的容器,如密封的安瓿和小管来提供药剂,也可将药剂储藏于冻干条件,仅需在临用前添加无菌液体载体,例如盐水和注射用水即可。或者,以持续灌注的形式提供药剂。临时用的注射溶液和悬浮液可制备自上述类型的无菌粉末、颗粒和片剂。

[0202] 适于直肠给药的药剂包括含标准载体,如可可脂或聚乙二醇的栓剂。适于口腔局部给药,例如颊或舌下给药的药剂包括在诸如蔗糖和金合欢胶或黄耆胶的经调味基质中含有活性成分的锭剂,以及在诸如明胶和甘油或蔗糖和金合欢胶的基质中含有活性成分的香锭。为了进行鼻内给药,可将本发明的化合物用作液体喷雾剂、可分散粉剂或滴剂。可用还含有一种或多种分散剂、增溶剂和/或悬浮剂的含水或不含水基质配制滴剂。

[0203] 为了经吸入给药,可以方便地由吹入器、喷雾器、加压填充物或其它能传递气溶胶喷雾剂的便利工具来递送化合物。加压填充物可含有适当的推进剂,如二氯二氟甲烷、三氯氟甲烷、二氯四氟乙烷、二氧化碳或其它适当的气体。对于加压气溶胶而言,通过提供阀来递送经计量的量即可测定剂量单位。

[0204] 或者,为了经吸入或吹入给药,化合物可采取干粉组合物的形式,例如其可以是化合物和适当粉末基质,如乳糖或淀粉的粉末混合物。可按单位剂量形式,例如以胶囊、药筒、明胶或发疱填充物的形式提供粉末组合物,借助于吸入器或吹入器即可施用所述粉末。

[0205] 其它药剂包括能释放治疗剂的可植入装置和带粘性的补片。

[0206] 必要时可使用经修饰后能缓释活性成分的上述制剂。药物组合物也可含有其它活

性成分,如抗微生物剂、免疫抑制剂和 / 或防腐剂。

[0207] 应理解除了上文特别指出的成分外,本发明的药剂可包括其它与所述药剂的类型有关的本领域常规试剂。例如,适于口服给药的药剂可包括增味剂。

[0208] 优选的单位剂量药剂含有的活性成分为下文提及的有效剂量或其适当部分。

[0209] 对于上述的每种病而言,可按每天约 0.1 至约 250mg/kg 的剂量范围口服或经注射施用组合物,例如多肽和有机化合物。成人的剂量范围一般为约 5mg 至约 17.5g/ 天,优选为约 5mg 至约 10g/ 天,最优选为约 100mg 至约 3g/ 天。分开的单位中提供的片剂或递送药物的其它单位剂量形式可方便地含有在此剂量或多个所述剂量下有效的量,例如含有约 5mg 至约 500mg,通常为约 100mg 至约 500mg 的单位。

[0210] 所用的剂量取决于多个因素,包括受试者的年龄和性别,欲治疗的确切病症及其严重程度。根据疾病及其严重程度的不同,给药途径也有所不同。无论如何,本领域技术人员考虑到上述因素可常规计算出适当的和最佳的剂量。

[0211] 下文实施例描述了本发明的各个方面,但它并不会限制权利要求书中描述的本发明的范围。下文的实施例阐明了在 BRC 细胞中差异表达的基因的鉴定和表征。

#### [0212] 实施例

[0213] 评价得自患病组织(如 BRC 的上皮细胞)和正常组织的组织,以鉴定出在疾病状态(如 BRC)下差异表达的基因。按下述进行试验。

#### [0214] 患者和组织样品:

[0215] 在知情并同意的情况下,从在日本东京癌症研究所医院乳腺外科部接受治疗的患者中获得 81 名原发性乳腺癌患者(12 例原位导管癌和 69 例浸润性导管癌,2cm 至 5cm(T2),年龄范围为 21 至 68 岁,平均年龄 45 岁),相关情况已告诉所有患者并获得了她们的同意(表 12)。临床信息得自医学记录,病理学家根据组织病理学亚型和分级来诊断每个肿瘤。使用肿瘤组织评价肿瘤类型(根据世界卫生组织的分类体系和日本癌症学会的分类体系)。根据 JBCS TNM 分类体系判断临床阶段。在淋巴结-阳性和淋巴结-阴性的病例之间未观察到显著差异。由病理学家测定侵入血管的生长和广泛的淋巴细胞浸润的存在。通过 EIA 测定雌激素受体(ER)和孕酮受体(PgR)的表达(当少于 13fmol/mg 蛋白质时为 ER 阴性, BML)。将取自 15 名绝经前乳腺癌患者或 12 名绝经后患者的正常乳腺导管细胞的混合物分别用作正常对照。将所有样品立即冷冻并储藏于 -80°C。

#### [0216] 组织样品和 LMM:

[0217] 表 12 详细描述了有关肿瘤的临床和病理学信息。将样品包埋于 TissueTek OCT 介质(Sakura),然后储藏于 -80°C 待用。在低温恒温器中将冷冻样品连续切成 8- $\mu$ m 的薄片,并用苏木精和伊红染色以限定待分析区域。为了避免癌细胞和非癌细胞的交叉污染,根据厂商的方案,加以几处改动,通过 EZ Cut LMM System(SL Microtest GmbH)制备这两个细胞群体。为了使储藏过程和组织收集期间的的影响最小化,通过相同的程序小心地处理癌症组织。为了检查 RNA 的质量,在变性琼脂糖凝胶中电泳提取自每个病例残留组织的总 RNA,并通过核糖体 RNA 带的存在证实其质量。

#### [0218] RNA 提取和基于 T7 的 RNA 扩增:

[0219] 从每个经激光捕获的细胞群体中提取总 RNA,溶于 350  $\mu$ l RLT 裂解缓冲液(QIAGEN)。室温下用 30 单位 DNase I(QIAGEN)将提取的 RNA 处理 30 分钟。于 70°C 灭活 10 分

钟之后,根据厂商推荐的方案用 RNeasy Mini 试剂盒 (QIAGEN) 纯化 RNA。使用 Ampliscribe T7 转录试剂盒 (Epicentre Technologies) 对所有经 DNase I 处理的 RNA 进行基于 T7 的扩增。对每个样品而言,两轮扩增产生 28.8–329.4  $\mu$ g 经扩增的 RNA (aRNA),然而,当扩增得自 15 名绝经前患者或 12 名绝经后患者的正常样品的 RNA 时,总共分别产生了 2240.2  $\mu$ g 和 2023.8  $\mu$ g。分别在 Cy5-dCTP 和 Cy3-dCTP (Amersham Biosciences) 的存在下逆转录 2.5  $\mu$ g 得自每种癌细胞和非癌乳腺导管细胞的 aRNA 等份试样。

[0220] cDNA 微阵列:

[0221] 建立“基因组范围内的”cDNA 微阵列系统,其含有 23,040 个选自 UniGene 数据库 (build#131) (国立生物技术信息中心 (NCBI)) 的 cDNA。别处已描述了 cDNA 微阵列载玻片的制作 (Ono K, Tanaka T, Tsunoda T, Kitahara O, Kihara C, Okamoto A, Ochiai K, Katagiri T and Nakamura Y. Identification by cDNA Microarray of Genes Involved in Ovarian Carcinogenesis. *Cancer Res.*, 60, 5007–11, 2000.)。简单地说,以分离自多种人器官的 poly(A)+RNA 为模板,通过逆转录-PCR 扩增 cDNA;扩增子的长度为 200 至 1100bp,其中不含重复序列或 poly(A) 序列。使用 Lucidea 微阵列点样器 (Amersham Biosciences) 将 PCR 产物以双份点样至 7-型载玻片 (Amersham Bioscience) 上;在单个载玻片上以双份点样 4,608 或 9,216 个基因。制备三个不同套别的载玻片(总共 23,040 个基因),在每个载玻片上还点上了相同的 52 个持家基因和两种阴性-对照基因。

[0222] 杂交和数据获取:

[0223] 除了所有过程皆用 Automated Slide Processor (Amersham Biosciences) 进行外,根据先前描述的方法进行杂交和洗涤 (Giuliani, N., et al., V. Human myeloma cells stimulate the receptor activator of nuclear factor- $\kappa$ B ligand (RANKL) in T lymphocytes: a potential role in multiple myeloma bone disease. *Blood*, 100: 4615–4621, 2002.)。使用 Array Vision 计算机程序 (Amersham Pharmacia) 以光度学计算每个杂交信号的强度,并减去背景强度。调整每个靶点的 Cy5 (肿瘤) 和 Cy3 (对照) 的荧光强度,以使用得自 52 个持家基因的平均信号执行平均的 Cy5/Cy3 比率。由于得自低信号强度的数据较不可靠,测定每个载玻片上信号强度的截断值,当 Cy3 和 Cy5 染料所产生信号的强度低于截断值时,即从进一步的研究中排除这些基因。根据背景的波动,自动计算出每个表达水平的截断值。当 Cy5 和 Cy3 信号强度皆低于截断值时,将该样品中相应基因的表达评估为缺乏。将 Cy5/Cy3 比率计算为相对表达率。对于其它基因而言,使用每个样品的原始数据计算 Cy5/Cy3 比率。

[0224] 使用 Array Vision 软件 (Imaging Research, Inc., St. Catharines, Ontario, Canada),通过置换背景定量和分析 23,040 个点中的 Cy3 和 Cy5 信号强度。随后调整每个靶点的 Cy5 (肿瘤) 和 Cy3 (对照) 的荧光强度,以使微阵列上 52 个持家基因的平均 Cy3/Cy5 比率等于 1。由于得自低信号强度的数据较不可靠,按先前所述的方法测定每个载玻片上的截断值 (Ono, K., et al., Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Res*, 60: 5007–5011, 2000.)。当 Cy3 和 Cy5 染料所产生信号的强度低于截断值时,即从进一步的研究中排除这些基因 (Saito-Hisaminato, A., Katagiri, T., Kakiuchi, S., Nakamura, T., Tsunoda, T., and Nakamura, Y. Genome-wide profiling of gene expression in 29 normal human tissues with a cDNA microarray.

DNA Res, 9 :35-45, 2002.)。对于其它基因而言,使用每个样品的原始数据计算 Cy5/Cy3 比率。

[0225] 计算污染百分比:

[0226] 通过用 cDNA 微阵列测定 29 个正常人组织的基因表达分布图,发现围脂滴蛋白 (PLIN) 和脂肪酸结合蛋白 4 (FABP4) 仅在脂肪组织和乳腺组织中表达 (Saito-Hisaminato, A. et al., Genome-wide profiling of gene expression in 29 normal human tissues with a cDNA microarray. DNA Res, 9 :35-45, 2002.)。使用它们来评价经显微解剖的正常乳腺导管上皮细胞群体中存在的脂肪细胞的比例。分别在 Cy5-dCTP 和 Cy3-dCTP 的存在下逆转录分离自正常完整乳腺的 poly A+RNA (Clontech) 的每种 aRNA 以及经显微解剖的正常乳腺导管上皮细胞的每种 aRNA。在微阵列载玻片上杂交之后,计算 Cy5/Cy3 比率。由使用绝经前患者和绝经后患者的乳腺组织和经显微解剖的正常乳腺导管细胞获得的结果决定每个比率的平均值。

[0227] 根据基因表达分布图对 81 例乳腺癌的 102 个样品进行聚类分析:

[0228] 将不受监督的等级聚类方法用于基因和肿瘤。为了获得可再现的簇用以对 102 个样品进行分类,选择了 710 基因,已在 80% 的实验中获得了这些基因的有效数据,它们的表达率有所改变,其中标准误差大于 1.1。使用 M. Eisen 编写的可获自万维网的软件 ("Cluster" and "TreeView") (<http://genome-www5.stanford.edu/MicroArray/SMD/restech.html>) 进行分析。在使用聚类算法之前,将每个点的荧光率转化为对数,然后将每个样品的数据集中至中值以除去实验偏差,并使用平均连锁值。

[0229] 鉴定在 DCIS 和 IDC 之间上调或下调的基因:

[0230] 将每种基因的相对表达率 (Cy5/Cy3 强度比率) 归类为下述四个类别中的一类: (A) 上调 (表达率 > 2.0); (B) 下调 (表达率 < 0.5); (C) 不变 (表达率在 0.5 至 2.0 之间); 和 (D) 未表达 (或少量表达,但低于检测的截断值水平)。使用这些类别检测出一套基因,所述基因在样品中的表达率常常有所改变。为了检测在每组中通常被上调或下调的候选基因,首先筛选 23,040 个基因的整体表达模式,以选择出在 > 50% 被归类的组中存在的表达率 > 3.0 或 < 1/3 的基因。

[0231] 半-定量 RT-PCR:

[0232] 选出 5 个上调基因,通过使用半-定量 RT-PCR 实验检查其表达水平。使用随机引物 (Taniguchi, K., et al., Mutational spectrum of beta-catenin, AXIN1, and AXIN2 in hepatocellular carcinomas and hepatoblastomas. Oncogene, 21 :4863-4871, 2002.) 和 Superscript II (Life Technologies, Inc.) 逆转录 1- $\mu$ g 得自每个样品的 aRNA 等份试样以获得单链 cDNA。稀释每种 cDNA 混合物,随后用表 9 所示的引物套进行 PCR 扩增。将 GAPDH 的表达用作内部对照。最优化 PCR 反应中的循环次数以确保产物强度位于扩增的线性期内。

[0233] 鉴定负责乳腺癌的组织病理学状态、ER 状态和淋巴结转移的基因:

[0234] 使用下面两个标准选择有差别的基因: (1) 在至少 70% (ER 状态) 或 50% (组织病理学状态和淋巴结转移) 的病例中信号强度高于截断值水平; (2) 病例的  $|\text{Medr}-\text{Medn}| > 1$  (ER 状态) 或 0.5 (组织病理学状态和淋巴结转移), 其中 Med 表示由淋巴结-阳性病例或-阴性病例已转化成对数的相对表达率获得的中值。接着使用随机排列测验鉴定出在

一组 (A 组) 和另一组 (B 组) 之间差异表达的基因。由 A 组 (r) 和 B 组 (n) 病例中每个基因已转化成对数的相对表达率计算出平均值 ( $\mu$ ) 和标准 ( $\sigma$ ) 误差。每个基因的差异评分 (DS) 的定义如下:

$$[0235] \quad DS = (\mu_r - \mu_n) / (\sigma_r + \sigma_n)$$

[0236] 进行排列测验以估计各个基因区分 A 组和 B 组的能力;在两个类别之间 10,000 次地随机排列样品。由于每个基因的 DS 数据套显示出正常的分布,计算出 P 值以进行用户-受限的分组 (Golub, T. et al., Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science, 286: 531-537, 1999.)。

[0237] 计算淋巴结转移的预测评分:

[0238] 根据先前描述的方法计算预测评分 (Golub, T. et al., Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science, 286: 531-537, 1999.)。根据样品中的表达水平 ( $x_i$ ) 是否与参照样品中淋巴结-阴性或-阳性样品的平均表达水平接近,将每个基因 ( $g_i$ ) 表决为淋巴结-阴性或淋巴结-阳性。票数 ( $v_i$ ) 的数量反映了样品中的表达水平与两个类别的平均值的偏差:

$$[0239] \quad V_i = |x_i - (\mu_r + \mu_n) / 2|$$

[0240] 计算总票数以获得淋巴结-阴性 ( $V_r$ ) 和淋巴结-阳性 ( $V_n$ ) 的总票数,并按下述计算 PS 值:

[0241]  $PS = (V_r - V_n) / (V_r + V_n) \times 100$ , 反映出在淋巴结-阴性或淋巴结-阳性方向上胜利的差数。PS 值为 -100 至 100; PS 的绝对数值较高反映了较强的预测能力。

[0242] 评价分类和 Leave-One-Out 试验:

[0243] 按下述使用每个基因套中淋巴结-阴性 ( $PS_r$ ) 和淋巴结-阳性 ( $PS_n$ ) 的预测评分计算分类评分 (CS):

$$[0244] \quad CS = (\mu_{PS_r} - \mu_{PS_n}) / (\sigma_{PS_r} + \sigma_{PS_n})$$

[0245] CS 值较大表示通过预测-评分系统已将两组较好地分开。为了进行 leave-one-out 试验,扣留一个样品,使用其余样品计算排列 p-值和平均表达水平,随后通过计算其预测评分来评价扣留样品的类别。对 20 个样品的每一个重复此过程。

[0246] 细胞系

[0247] 人乳腺癌细胞系 HBL-100, HCC1937, MCF-7, MDA-MB-435s, YMB1, SKBR3, T47D, BT-20, BT-474, BT-549, HCC1143, HCC1500, HCC1599, MDA-MB-157, MDA-MB453, OUCB-F, ZR-75-1, COS-7 细胞系购自美国典型培养物保藏中心 (ATCC), 并按照其各自的保藏者推荐的方法进行培养。HBC4, HBC5 和 MDA-MB-231 细胞系由 Molecular Pharmacology, Cancer Chemotherapy Centre of the Japanese Foundation for Cancer Research 的 Yamori 博士惠赠。将所有细胞培养于适当的培养基中,即 RPMI-1640 (Sigma, St. Louis, MO), 它可用于 HBC4, HBC5, T47D, YMB1, OUCB-F, ZR-75-1, BT-549, HCC1143, HCC1500, HCC1599 和 HCC1937 (含 2mM L-谷氨酰胺); Dulbecco 改进的 Eagle 培养基 (Invitrogen, Carlsbad, CA), 它可用于 BT474, HBL 100, COS7; 含 0.1mM 必需氨基酸 (Roche), 1mM 丙酮酸钠 (Roche), 0.01mg/ml 胰岛素 (Sigma) 的 EMEM (Sigma), 它可用于 BT-20 和 MCF-7; McCoy (Sigma), 它可用于 SKBR3 (含 1.5mM L-谷氨酰胺); L-15 (Roche), 它可用于 MDA-MB-231, MDA-MB-157,

MDA-MB453 和 MDA-MB-435S。每种培养基都添加有 10% 胎牛血清 (Cansera) 和 1% 抗生素 / 抗真菌溶液 (Sigma)。于 37°C, 在不含 CO<sub>2</sub> 的湿润空气的气氛中维持 MDA-MB-231 和 MDA-MB-435S 细胞。于 37°C, 在含 5% CO<sub>2</sub> 的湿润空气的气氛中维持其它细胞系。临床样品 (乳腺癌和正常乳腺导管) 得自外科手术样品, 有关情况已让所有患者知情并同意。

#### [0248] Northern-印迹分析

[0249] 根据厂商说明, 使用 Rneasy 试剂盒 (QIAGEN) 从所有乳腺癌细胞系中提取总 RNA。用 DNase I (Nippon Gene, Osaka, Japan) 处理之后, 根据厂商说明, 使用 mRNA 纯化试剂盒 (Amersham Biosciences) 分离 mRNA。在 1% 变性的琼脂糖凝胶上分离 1- $\mu$ g 每种 mRNA 以及分离自正常成人乳腺 (Biochain)、肺、心脏、肝脏、肾脏、骨髓 (BD, Clontech, Palo Alto, CA) 的 polyA(+)RNA 的等份试样, 并转移至尼龙膜上 (乳腺癌-Northern 印迹)。使乳腺癌-和人多组织 Northern 印迹 (Clontech, Palo Alto, CA) 与通过 RT-PCR 制备的 A7870 经 [ $\alpha$  <sup>32</sup>P]-dCTP- 标记的 PCR 产物杂交 (见下文)。根据厂商推荐的方法进行预-杂交、杂交和洗涤。于 -80°C 用增感屏使印迹放射自显影 14 天。通过使用下述引物套进行 RT-PCR, 制备出针对 A7870 (320bp) 的特异性探针; 5'-AGACCCTAAAGATCGTCCTTCTG-3' (SEQ ID NO:13) 和 5'-GTGTTTTAAGTCAGCATGAGCAG-3' (SEQ ID NO:14), 用大量引发 (megeprime) 的 DNA 标记系统 (Amersham bioscience) 对所述探针进行放射性标记。

#### [0250] 免疫细胞化学染色

[0251] 为了构建 A7870 表达载体, 通过使用 KOD-Plus DNA 聚合酶 (Toyobo, Osaka, Japan) 进行 PCR 来扩增 A7870cDNA 的完整编码序列。将 PCR 产物插入 pCAGGSn3FH-HA 表达载体的 EocRI 和 Xho I 位点。通过 DNA 测序进一步证实该构建体 (pCAGGS-A7870-HA)。接着, 为了先检查外源性 A7870 的亚细胞定位, 我们以每孔  $1 \times 10^5$  个细胞的密度接种 COS7 细胞以进行外源性表达。24 小时之后, 根据厂商说明分别用 FuGENE 6 转染试剂 (Roche) 将 1  $\mu$ g pCAGGS-A7870-HA 瞬时转染至 COS7 细胞。然后, 用含有 4% 低聚甲醛的 PBS 固定细胞 15 分钟, 于 4°C 用含有 0.1% Triton X-100 的 PBS 透化细胞 2.5 分钟。随后, 于 4°C 用含有 3% BSA 的 PBS 覆盖细胞 12 小时以阻断非-特异性杂交。接着, 使经 A7870-HA- 转染的 COS7 细胞与稀释度为 1 : 1000 的小鼠抗-HA 抗体 (SANTA CRUZ) 和稀释度为 1 : 1000 的抗-TOPK 多克隆抗体 (Cell Signaling) 一起保温。用 PBS 洗涤之后, 用稀释度为 1 : 5000 的与 Alexa594- 偶联的抗-小鼠第二抗体 (Molecular Probe) 染色这两种经转染的细胞。

[0252] 我们进一步证实了内源性 A7870 蛋白在密度为每孔  $2 \times 10^5$  个细胞的乳腺癌细胞系 T47D, BT-20 和 HBC5 中的亚细胞定位。使细胞与稀释度为 1 : 1000 的兔抗-TOPK 多克隆抗体一起保温, 所述抗体由对应于人 PBK/TOPK c- 末端氨基酸的合成肽制成。用 PBS 洗涤之后, 用稀释度为 1 : 3000 的与 Alexa488- 偶联的抗-兔第二抗体 (Molecular Probe) 染色细胞。用 4',6'-二脒-2'-苯基吡啶二盐酸化物 (DAPI) 复染细胞核。在 TCS SP2 AOBs 显微镜 (Leica, Tokyo, Japan) 下获得荧光图像。

#### [0253] 使用 psiU6BX3.0 构建 A7870 特异性-siRNA 表达载体

[0254] 我们根据以前的报道, 使用 psiU6BX3.0 siRNA 表达载体建立基于载体的 RNAi 系统 (Shimokawa T, Furukawa Y, Sakai M, Li M, Miwa N, Lin YM, Nakamura Y (2003). Cancer Res, 63, 6116-6120)。通过将表 13 中的双链寡核苷酸克隆至 psiU6BX3.0 载体的 BbsI 位点, 制备出针对 A7870 的 siRNA 表达载体 (psiU6BX-A7870)。通过分别将下述双链寡核苷酸克

降至 psiU6BX3.0 载体的 BbsI 位点,制备出对照质粒 psiU6BX-SC 和 psiU6BX-LUC;针对 SC 的 5' -TCCC GCGC GCTTTGTAGGATTCGTTCAAGAGACGAATCCTACAAAGCGCGC-3' (SEQ ID NO:15) 和 5' -AAAAGCGCGCTTTGTAGGATTCGTTCTTGAACGAATCCTACAAAGCGCGC-3' (SEQ ID NO:16) (scrambled 对照);针对 LUC 的 5' -TCCC GTACGCGGAATACTTCGATTCAAGAGATCGAAGTATTCGCG GTACG-3' (SEQ ID NO:17) 和 5' -AAAACGTACGCGGAATACTTCGATCTCTTGAATCGAAGTATTCGCG TACG-3' (SEQ ID NO:18) (荧光素酶对照)。

#### [0255] A7870 的基因 - 沉默作用

[0256] 将人乳腺癌细胞系 T47D 或 BT-20 铺于 15-cm 平皿中 ( $4 \times 10^6$  个细胞 / 皿), 根据厂商推荐的方法, 使用 FuGENE6 试剂 (Roche) 将各 16  $\mu$ g 的 psiU6BX-LUC (荧光素酶对照), 作为阴性对照的 psiU6BX-SC (scrambled 对照) 和 psiU6BX-A7870 转染至所述细胞中。转染 24 小时之后, 再次重接种细胞以进行集落形成试验 ( $2 \times 10^6$  个细胞 / 10cm 平皿), RT-PCR ( $2 \times 10^6$  个细胞 / 10cm 平皿) 和 MTT 试验 ( $2 \times 10^6$  个细胞 / 孔)。我们在 T47D 或 BT-20 细胞中分别使用含有 0.7mg/ml 或 0.6mg/ml 新霉素 (Geneticin, Gibco) 的培养基选择出导入 A7870 的细胞。然后, 每两天换一次培养基, 持续 3 周。为了评价 siRNA 的功能, 在用新霉素选择 11 天之后, 从细胞中提取总 RNA, 然后使用针对 A7870 和 GAPDH 的下述特异性引物套进行半 - 定量 RT-PCR 以进一步证实 siRNA 的击倒作用: 5' -ATGGAATCCCATCACC ATCT-3' (SEQ ID NO:19) 和 5' -GGTTGAGCACAGGGTACTTTATT-3' (SEQ ID NO:20) 针对的是用作内部对照的 GAPDH, 而 5' -GCCTTCATCATCC AACATT-3' (SEQ ID NO:21) 和 5' -GGCAAATATGTCTGCCTTGT-3' (SEQ ID NO:22) 针对的是 A7870。

[0257] 另外, 在含有新霉素的选择培养基中分别将使用 T47D 或 BT-20 细胞系的能表达 siRNA 的转染子培养 23 天。用 4% 低聚甲醛固定之后, 用吉姆萨溶液染色经转染的细胞以评价集落形成。进行 MTT 试验以定量细胞生存力。在含新霉素的培养基中培养 10 天后, 以 0.5mg/ml 的浓度加入 MTT 溶液 (3-(4,5-二甲基-2-噻唑)-2,5-二苯基溴化四唑) (Sigma)。37°C 保温 2.5 小时之后, 加入酸 - SDS (0.01N HCl/10% SDS); 剧烈混合悬浮液, 然后于 37°C 保温过夜以溶解黑蓝色的晶体。用 Microplate Reader 550 (BioRad) 测定 570nm 下的光吸收值。为了评价 siRNA 的功能, 从选择 7 天之后的细胞中提取总 RNA, 在选择 10 天后, 根据厂商提供的方法, 使用 Cell Counting Kit-8 (Dojindo) 进行 MTT 试验。用 Microplate Reader 550 (BioRad) 测定 570nm 波长处的光吸收值。为了进行集落形成试验, 用 4% 低聚甲醛固定细胞 15 分钟, 然后用吉姆萨溶液 (Merck) 染色细胞。将每个实验重复 3 次。

#### [0258] 结果

##### [0259] 在乳腺癌的精确基因表达分布图的基础上进行分类分析:

[0260] 由于乳腺癌在肿瘤块中含有较少的癌细胞群体, 并且源自正常的上皮导管细胞, 因此需进行显微解剖以避免周围非 - 癌细胞或非 - 正常导管上皮细胞的污染。由于乳腺组织中的绝大多数细胞是脂肪细胞, 所以人们认为不适于使用整个乳腺组织来分析该器官中的癌症 - 特异性表达分布图。如图 1 所示, 从每个临床样品上经显微解剖得到 DCIS (病例 10326T), IDC (10502T) 和正常导管上皮 (10341N) 的代表性例子。这样就可以更精确地获得随后的基因表达分布图。如先前的文献所述, 通过测定在脂肪和乳腺组织中高表达的两个基因 (即 PLIN 和 FABP4) 的信号强度, 即可检查出污染了经显微解剖的正常乳腺

导管上皮细胞群体（用作通用对照）的脂肪细胞所占的比例（Saito-Hisaminato, A., et al., Genome-wide profiling of gene expression in 29 normal human tissues with a cDNA microarray. *DNA Res*, 9:35-45, 2002.）。当在含有大量脂肪细胞的整个乳腺组织中检查这些基因的信号强度时,这些基因信号强度的平均比率约为 99.4% ;经显微解剖的正常乳腺导管上皮细胞中的比率约为 0.6%（参见材料和方法中的污染百分比部分）。因此,据估计在显微解剖之后,对照细胞群体中污染的脂肪细胞的平均比率为 0.6%。首先,以基因在下述 102 个临床样品中表达模式的相似性为基础,使用未经监督的二维等级聚类算法对基因进行分组 :81 个经显微解剖的不同临床乳腺癌样品,10 个个体中的 11 个经显微解剖的不同组织类型,2 个完整的乳腺癌组织,6 个经显微解剖的正常乳腺导管细胞和 2 个完整的乳腺组织。用 710 个基因获得了可再现的基因簇（参见材料和方法）;图 2A 显示了它们跨越 102 个样品的表达模式。在样品轴中,根据其表达分布图将 102 个样品聚类成 3 个主要的组 (A, B 和 C 组)。然后,使分类与临床参数,特别是用 EIA 测定的雌激素受体 (ER) 相关联。在 55 个 ER- 阳性肿瘤中,45 个病例被聚类至肿瘤树形图的相同分支 (B 组),暗示着有处于 ER 状态的倾向。另外,在独立的实验中标记并杂交具有不同组织类型的 10 个病例中的 7 例 (样品 #10864, 10149, 10818, 10138, 10005, 10646 和 10435),它们被聚类至相同组内最接近的位置。特别是其中的一个重复病例 (10149a1 和 10149a1T) 也被聚类至最短的分支,该结果支持了微阵列数据的可再现性和可靠性。很显然,C 组除了一个经显微解剖的肿瘤病例外,还含有经显微解剖的非-癌细胞和乳腺癌全组织,暗示着该数据展现出准确的乳腺癌特异性-表达分布图。

[0261] 另外,从 8 位乳腺癌患者中经显微解剖获得 16 个具有 2 个分化损害的样品,跨越这些样品进行 89 个基因的二维等级聚类分析。结果,具有不同表型损害的乳腺癌样品紧密相邻 (图 2B)。接着,进行随机排列试验以鉴定出在与患者匹配的表型高度-或低度-分化的损害中差异表达的基因,其中所述损害得自经显微解剖的 8 个癌症样品。如图 2C 所示,使用 25 个显示出差异表达的基因进行的聚类分析可将高度-或低度-分化的浸润性导管癌细胞分开。这 25 个基因 (表 1) 包括一些关键的因子 :TNFSF11, ITGA5 和 NFAT5,以前的文献已报道过所述因子在侵入和细胞生长中可能起到的作用 (Giuliani, N., et al., Human myeloma cells stimulate the receptor activator of nuclear factor-kappa B ligand (RANKL) in T lymphocytes : a potential role in multiple myeloma bone disease. *Blood*, 100 :4615-4621, 2002. ; Sebastien J. et al., The role of NFAT transcription factors in integrin-mediated carcinoma invasion. *Nature cell biology*, 4 :540-544, 2002. ; Klein, S. et al., Alpha 5 beta 1 integrin activates an NF-kappa B-dependent program of gene expression important for angiogenesis and inflammation. *Mol Cell Biol*, 22 :5912-5922, 2002.)。

[0262] 接着,进行随机排列试验以鉴定出在 IDC 的 41 例 ER- 阳性肿瘤和 28 例 ER- 阴性肿瘤中差异表达的基因。所有样品均得自绝经前的患者。表中列出了 97 个能以低于 0.0001 的排列 P- 值区别 ER 阳性和阴性的基因 (参见“材料和方法”) (图 3 和表 2)。从中选出 96 个基因作为本发明的 BRC 相关基因。与 ER- 阴性组相比,在 ER- 阳性组中,这些基因中的 92 个基因的表达水平增加,其它 5 个的表达水平降低。在这些基因中, GATA 结合蛋白 3 (GATA3), 三叶因子 3 (TFF3), 细胞周期蛋白 D1 (CCND), MAPKK 同系物 (MAP2K4) 和金属蛋白



酶 1 的组织抑制剂 (TIMP1), 胰岛素受体底物 1 (IRS1), X-盒结合蛋白 1 (XBP1) GLI-Kruppel 家族成员 GLI3 (GLI3) 在 ER- 阳性中过表达 (表 2)。另外, 由于根据 p- 值的大小将雌激素受体 (ESR1) 列为第 6 个基因 (图 3 中下方的一组试验), 根据 ER 的表达分布图可以区分乳腺癌。

[0263] 鉴定在 DCIS 或 IDC 中通常被上调或下调的基因:

[0264] 为了进一步阐明乳腺癌的致癌机制, 分别研究在 DCIS 和 IDC 中通常被上调或下调的基因。77 例乳腺肿瘤 (8 例 DCIS 和 69 例绝经前患者) 中的基因表达分布图鉴定出 325 个表达常有改变的基因 (图 4A, 4B); 与其在正常乳腺导管细胞中的表达水平相比, 78 个基因的表达水平通常被上调 3 倍以上 (图 4A, 4B, 表 3, 5), 而 247 个基因在乳腺癌细胞中的表达降低 1/3 以下 (图 4A, 4B, 表 4, 6)。特别是如图 4B 所示, 当从 DCIS 转变为 IDC 时, 25 个基因的表达水平增加, 49 个基因的表达水平降低 (表 5 和 6)。表达有所增加的基因包括据报道在乳腺癌中过表达的纤连蛋白 (FN1) (Mackay, A. et al., cDNA microarray analysis of genes associated with ERBB2 (HER2/neu) overexpression in human mammary luminal epithelial cells. *Oncogene*, 22:2680-2688, 2003. ; Lalani, E. N. et al., Expression of the gene coding for a human mucin in mouse mammary tumor cells can affect their tumorigenicity. *J Biol Chem*, 266:15420-15426, 1991. ; 22. Martin-Lluesma, S., et al., A Role of Hecl in spindle checkpoint signaling and kinetochore recruitment of Mad1/Mad2. *Science*, 297:2267-2270, 2002.) (表 4)。另一方面, 表达有所降低的基因还包括已知能用作肿瘤抑制基因的 ST5 和 SCHIP1 (表 6)。

[0265] 接下来研究的是只在 IDC 中特异性地改变表达的基因。结果, 鉴定出 24 个上调基因 (图 4C, 表 7) 和 41 个下调基因 (图 4C, 表 8)。在上调基因中, 已知 ERBB2, CCNB1, BUB1B 参与乳腺癌的发生 (Latta, E. K., et al., The role of HER2/neu overexpression/amplification in the progression of ductal carcinoma in situ to invasive carcinoma of the breast. *Mod Pathol*, 15:1318-1325, 2002. ; Takeno, S., et al., Prognostic value of cyclin B1 in patients with esophageal squamous cell carcinoma. *Cancer*, 94:2874-2881, 2002. ; Slamon, D. J., et al., Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*, 235:177-182, 1987.)。下调基因中包括由 AXIN 诱导的基因 AXUD1, 该基因经常在肺、肝脏、结肠和肾癌中被下调 (Ishiguro, H., et al., Identification of AXUD1, a novel human gene induced by AXIN1 and its reduced expression in human carcinomas of the lung, liver, colon and kidney. *Oncogene*, 20:5062-5066, 2001.)。暗示着 AXUD1 也有可能参与乳腺癌的发生。

[0266] 通过半-定量 RT-PCR 证实选择的基因:

[0267] 为了进一步证实通过 cDNA 微阵列分析获得的表达数据的可靠性, 对 3 个在高度分化型报告病例中被高度上调的基因 (登录号为 AI261804, AA205444, AA167194), 以及 2 个在低度分化型报告病例中被高度上调的基因 (AA676987 和 H22566) 进行半-定量 RT-PCR 实验。RT-PCR 结果与在绝大多数受试病例中获得的微阵列分析结果高度一致 (图 5, 表 9)。

[0268] 将被称为源自 T-LAK 细胞的蛋白质激酶的 A7870 鉴定为乳腺癌细胞中的上调基因

[0269] 我们鉴定出 24 个在 IDC 中被上调的基因 (表 7)。其中, 我们更关注被称为源

自 T-LAK 细胞的蛋白质激酶 TOPK (Genbank 登录号为 NM 018492) 的 A7870, 它位于染色体 8p21.2 上, 由 8 个外显子组成的 mRNA 转录物的长度为 1899 个碱基。在能获得表达数据的 39 个乳腺癌病例的 30 例 (77%), 特别是在 36 例浸润性导管癌样品的 29 例 (81%) 中, A7870 的表达有所增加。为了进一步证实该基因在乳腺癌中的表达模式, 我们使用乳腺癌细胞系和包括正常乳腺细胞的正常人组织进行半-定量 RT-PCR 分析。结果, 我们发现: 与正常乳腺导管细胞和其它正常组织相比, 在 12 个临床乳腺癌样品 (高度-分化型) 的 7 个中, A7870 的表达有所增加 (图 6a), 在 20 个乳腺癌细胞系的 17 个中, A7870 被过表达 (图 6b)。为了进一步检查该基因的表达模式, 我们使用 A7870 的 cDNA 片段 (320bp) 为探针, 用多个人组织和乳腺癌细胞系进行 Northern 印迹分析 (图 7a)。结果, 我们观察到两个转录物 (约 1.9kb 和 1.8kb) 只在正常人睾丸和胸腺中表达。当我们用乳腺癌-Northern 印迹进一步检查这些转录物的表达模式时, 发现与正常人组织相比, 乳腺癌细胞系中的这两个转录物被特异性地过表达 (图 7b)。

#### [0270] 分离乳腺癌特异性-表达的 A7870 转录物

[0271] 通过对两个 A7870 转录物的测序分析, 由于 A7870 的两个变体含有相同的开放读码框 (ORF), 我们更关注 TOPK (Genbank 登录号为 NM018492), 它编码的蛋白质是与双重特异性促分裂原活化蛋白激酶激酶 (MAPKK) 家族相关的丝氨酸/苏氨酸激酶。SMART 计算机预测显示出 TOPK 含有 pfam, 32 至 320 个残基的 p 激酶基元, 暗示着该蛋白质可能参与在细胞形态发生和细胞生长中起作用的信号转导途径。

#### [0272] A7870 的亚细胞定位

[0273] 为了进一步检查 A7870 的鉴定, 我们检查了这些基因产物在哺乳动物细胞中的亚细胞定位。首先, 当我们将表达 A7870 蛋白的质粒 (pCAGGS-A7870-HA) 瞬时转染至 COS7 细胞时, 用抗-Hatag 抗体和 TOPK 多克隆抗体进行的免疫细胞化学分析揭示出外源性 A7870 蛋白位于所有经转染 COS7 细胞的胞质, 特别是在核膜周围有强信号存在 (图 8a)。另外, 我们使用抗-TOPK 多克隆抗体进行免疫细胞化学染色, 检查了内源性蛋白质的亚细胞定位。类似地, 在 T47D, BT-20 和 HBC5 细胞的胞质器和核周围也观察到 A7870 蛋白 (图 8b)。

#### [0274] 被设计用于降低 A7870 表达的小的干扰 RNA (siRNA) 的生长-抑制作用

[0275] 为了评估 A7870 的生长-促进作用, 我们通过基于哺乳动物载体的 RNA 干扰 (RNAi) 技术 (参见材料和方法), 在显示出 A7870 过表达的乳腺癌细胞系 T47D 和 BT-20 中击倒内源性 A7870 的表达。我们通过半-定量 RT-PCR 实验检查了 A7870 的表达水平。与对照 siRNA 构建体 (psiU6BX-LUC 或 -SC) 相比, A7870 (si1, si3 和 si4)-特异性 siRNA 能显著抑制表达。为了进一步证实 A7870-特异性 siRNA 的细胞生长抑制作用, 我们分别进行了集落-形成和 MTT 试验。结果, 导入 A7870 siRNA 构建体能抑制这些乳腺癌细胞的生长, 这与上文所述的该基因的表达有所降低的结果一致。通过 3 个独立的实验证实每个结果。因此, 我们的发现表明 A7870 对乳腺癌的细胞生长起着显著作用。

#### [0276] 鉴定在组织病理学类型中差异表达的基因和各个患者的表型差异:

[0277] 本发明的一个目标是发现在一些患者的不同表型中始终被上调或下调的基因。然而, 由于乳腺癌显示出不均一的和不同的表型, 通过图 2 显示的基因表达模式进行未经监督的分类不能清楚地分辨显微术获得的组织病理学分化。为了更准确地检查这一观察结果, 我们进行了随机-排列试验, 提取出 206 个能区分高度-分化和低度-分化病例的基

因。在 31 例高度 - 和 24 例低度 - 分化的癌症之间, 这 206 个区别基因的显著水平皆为  $P < 0.01$  (图 9, 表 10)。使用这 206 个基因的二维等级聚类分析也能对与 IDC 的不同组分有关的组进行分类 (高度 - 分化的, 中度 - 分化的和低度 - 分化的)。A 组的基因簇含有在低度 - 分化的样品中表达显著增加的基因 (横行中的分支 1); 细胞外基质结构 (COL1A2, COL3A1 和 P4HA2), 细胞粘附 (LOXL2, THBS2 和 TAGLN2), 而 B 组的基因簇含有主要在高度 - 分化和中度 - 分化的样品中表达有所增加的基因 (横行中的分支 2); 调节转录 (BTF, WTAP, HTATSF1), 细胞周期调节物 (CDC5L, CCT7)。然而, B 组中的两个低度 - 分化的样品 (样品 #10709 和 10781) 显示出与高度 - 分化型而不是低度 - 分化型类似的表达模式。一些高度 - 分化的样品显示出低度 - 分化型特征性的一些基因的共表达。

#### [0278] 开展对淋巴结转移的预测评分:

[0279] 在乳腺癌中, 对腋下淋巴结的侵入是最重要的预测因素 (Shek, L. L. and Godolphin, W. Model for breast cancer survival: relative prognostic roles of axillary nodal status, TNM stage, estrogen receptor concentration, and tumornecrosis. *Cancer Res*, 48 :5565-5569, 1988.)。为了研究出可使用选定基因的表达分布图获得评分参数以预测腋下淋巴结转移的等式, 我们比较了 20 个淋巴结 - 阳性病例和 20 个淋巴结 - 阴性病例的表达分布图。根据上文所述的标准, 首先选择 93 个排列  $p$ - 值小于 0.0001 的区别基因。然后, 获得候选列表中的前 34 个基因, 这些基因能最佳地区分淋巴结 - 阳性和淋巴结 - 阴性病例 (表 11)。如图 10A 所示, 使用这 34 个基因的等级聚类分析根据淋巴结状态将所有 40 个乳腺癌病例清楚地归类至两组中的一组。

[0280] 最终, 构建出预测 - 评分系统, 该系统可使用由 34 个基因组成的一套基因的表达分布图清楚地区分淋巴结 - 阳性病例和淋巴结 - 阴性病例。为了进一步证实该评分系统, 计算出构建评分系统时未使用的 20 个淋巴结 - 阳性病例和 20 个淋巴结 - 阴性病例的分值 (参见“材料和方法”)。当属于阳性 - 转移组和阴性转移组的 40 个患者的界限分值 15.8 被清楚地分开时 (图 10B), 分值大于 15.8 即为“阳性”, 分值等于或小于 15.8 即为“阴性”。为了进一步阐明该系统, 计算出原发性肿瘤转移的预测评分, 所述肿瘤包括 17 个淋巴结 - 阳性病例和 20 个淋巴结 - 阴性病例, 这些病例是选择区别基因的最初过程所未曾使用过的。如图 10B 和 10C 所示, 17 个淋巴结转移病例中的所有病例都具有根据本文定义的正分, 而 20 个不具有淋巴结转移的病例中有 18 例 (90%) 显示出负分。根据其淋巴结状态正确安放了 77 个病例中的 75 例 (97%), 但有 2 个淋巴结 - 阴性病例被放错了或被放在边界处或阳性区。

#### [0281] 讨论

[0282] 乳腺癌是多因素疾病, 是由遗传、环境和激素这几个因素相互作用引起的。尽管已描述过乳腺癌的不同病理学阶段, 但对这些阶段的分子差异仍知之甚少 (McGuire, W. L. Breast cancer prognostic factors: evaluation guidelines. *J Natl Cancer Inst*, 83 :154-155, 1991. ; Eifel, P., et al., National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000. *J Natl Cancer Inst*, 93 :979-989, 2001. ; Fisher, B., et al., Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl*

J Med, 347 :1233-1241, 2002.)。

[0283] 开发基因表达的基因组 - 范围内的分析以及分离乳腺癌纯的癌细胞群体的激光微束显微解剖 (LMM) 能够在多种肿瘤类型, 特别是乳腺癌中找到具有癌症 - 特异性分类、治疗和结果预测作用的分子 - 靶基因。

[0284] 由于脂肪细胞占乳腺组织的 90% 以上, 器官中始发癌症的上皮细胞对应于很小的百分比, 使用完整的癌组织和正常的完整乳腺进行的基因表达分布图分析会受到被检查组织中具体细胞混合物的显著影响; 脂肪细胞、成纤维细胞和炎性细胞成比例的差异可显著遮蔽参与乳腺癌发生的基因的特异性表达。因此, 使用 LMM 系统尽可能多地纯化得自外科手术样品的癌细胞和正常上皮细胞群体 (Hasegawa, S., et al. Genome-wide analysis of gene expression in intestinal-type gastric cancers using a complementary DNAmicroarray representing 23,040 genes. Cancer Res, 62 :7012-7017, 2002. ; Kitahara, et al., and Tsunoda, T. Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. Cancer Res, 61 :3544-3549, 2001. ; Kikuchi, T., et al. Expression profiles of non-small cell lung cancers on cDNA microarrays: identification of genes for prediction of lymph-node metastasis and sensitivity to anti-cancer drugs. Oncogene, 22 :2192-2205, 2003. ; Gjerdrum, L. M., et al., Laser-assisted microdissection of membrane-mounted paraffin sections for polymerase chain reaction analysis: identification of cell populations using immunohistochemistry and in situ hybridization. J Mol Diagn, 3 :105-110, 2001.) (图 1)。为了评价经显微解剖的细胞群体的纯度, 使用 cDNA 微阵列测得 29 个正常人组织的基因表达分布图, 籍此分析在脂肪组织和乳腺中高表达的 PLIN 和 FABP4 的表达 (Saito-Hisaminato, A., et al., Genome-wide profiling of gene expression in 29 normal human tissues with a microarray. DNA Res, 9 :35-45, 2002.)。经解剖之后, 估计正常乳腺导管上皮细胞中污染的脂肪细胞的比例小于 0.6%。特别是, 当检查 PLIN 的表达水平时 (Nishiu, J., et al., Isolation and chromosomal mapping of the human homolog of perilipin (PLIN), a rat adipose tissue-specific gene, by differential display method. Genomics, 48 :254-257, 1998.), 经过 LMM 技术处理的细胞群体的纯度可约为 100%。如图 2 所示, 未经监督的聚类分析表明: 通过 LMM 可将乳腺癌完整组织与经显微解剖的乳腺癌细胞分开, 而正常的乳腺导管细胞和乳腺被聚类至相同的分支。因此, 为了在一些研究中准确地获得乳腺癌特异性的表达分布图, 必需显微解剖乳腺癌细胞和始发乳腺癌的正常乳腺导管上皮细胞。联合使用 LMM 和 cDNA 微阵列分析提供了一种很有效的方法, 所述方法可阐明与乳腺癌的发生和发展有关的精确分子事件, 并有助于人们理解乳腺癌细胞多步骤癌症发生和肿瘤不均一性的机制。

[0285] 如图 2A 所示, 通过在表达分布图的基础上进行的未经监督的分类分析, 将原发性乳腺癌分成两组, 通过 EIA 证实它们与 ER 状态有关。发现 ER+ 和 ER- 肿瘤显示出非常不同的基因表达表型。该结果表明这两个组织类型不同的损害具有不同的可能在乳腺癌发生中起重要作用的生物特性, 该结果进一步表明可使用 ER 状态在辅助背景中确立激素疗法

的必要性 (Eifel, P., et al National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000. J Natl Cancer Inst, 93:979-989, 2001. ;Hartge, P. Genes, hormones, and pathways to breast cancer. N Engl J Med, 348:2352-2354, 2003.)。另外,通过经监督的统计学分析,选择出能区分 ER- 阳性和 ER- 阴性以研究依赖于激素的进程的基因亚群,并探查抗- 癌药物的新的分子- 靶标。通过随机排列试验鉴定出 97 个在由绝经前患者组成的这两组之间有着显著不同的表达的基因 (图 3)。这些基因中包括 MAP2K4,它是 SAPK 途径占主要地位的介质,还包括细胞周期蛋白 D1,它是在此研究和其它研究与乳腺癌中的 ER 表达密切相关的基因 (May, F. E. and Westley, B. R. Expression of human intestinal trefoil factor in malignant cells and its regulation by oestrogen in breast cancer cells. J Pathol, 182:404-413, 1997.)。雌激素是正常乳腺生长和分化的重要调节剂,它对乳腺癌的发生和发展也很重要 (Shek, L. L. and Godolphin, W. Model for breast cancer survival: relative prognostic roles of axillary nodal status, TNM stage, estrogen receptor concentration, and tumor necrosis. Cancer Res, 48:5565-5569, 1988.)。雌激素经由 ER 调节基因表达;然而,人们仍不能全面而详细地了解雌激素对下游基因靶的作用、辅因子的作用和其它信号转导途径之间的串话。由于在诊断时约有 2/3 的乳腺癌是 ER+, 受体的表达对其生物学和疗法有重要的暗示作用。由于最新的选择性雌激素受体调制剂 (SERM) 正被开发成针对 ER- 阳性乳腺癌患者的激素疗法,这些与 ER 状态有关的基因可能是 SERM 新的潜在的分子- 靶标 (Smith, I. E. and Dowsett, M. Aromatase inhibitors in breast cancer. N Engl J Med, 348:2431-2442, 2003.)。这些发现暗示着比较表达分布图和 ER- 状态可提供有用的信息,用以阐明不依赖于 ER 的乳腺癌细胞的细胞增殖和进程的激素调节。

[0286] 开发和针对乳腺癌和其它人类恶性肿瘤的基于分子的疗法需要对患者组织进行详细的分子遗传学分析。组织学证据表明在浸润性乳腺肿瘤之前存在几个前肿瘤状态。这些组织损害包括非典型性导管增生、非典型性小叶增生、原位导管癌 (DCIS) 和原位小叶癌 (Lakhani, S. R. The transition from hyperplasia to invasive carcinoma of the breast. J Pathol, 187:272-278, 1999.)。据认为这些损害攻击处于始发乳腺癌的正常乳腺上皮或末端导管小叶单元与最终的浸润性乳腺癌之间的组织连续体。有人提议用几个模型解释前肿瘤和肿瘤形成之间的基因异常。

[0287] 观察到多个基因在病理学分立期,如比较 DCIS 和 IDC 时通常显示出增加或降低的表达,总共鉴定出 325 个基因。这些基因构成乳腺癌病理学分级的分子基础,这些基因的表达水平与晚期肿瘤分级有关。还在 DCIS 和 IDC 中鉴定出 78 个通常被上调的基因 (表 3, 5) 和 247 个通常被下调的基因 (表 4, 6)。在上调基因中,注意到据报道在乳腺癌中过表达的 NAT1, HEC, GATA3 和 RAI3 在侵入前期潜在表达 (Geylan, Y. S., et al., Arylamine N-acetyltransferase activities in human breast cancer tissues. Neoplasia, 48:108-111, 2001. ;Chen, Y., et al., HEC, a novel nuclear protein rich in leucine heptad repeats specifically involved in mitosis. Mol Cell Biol, 17:6049-6056, 1997. ;Bertucci, F., et al., Gene expression profiling of primary breast carcinomas using arrays of candidate genes. Hum Mol Genet, 9:2981-2991, 2000. ;Cheng, Y. and Lotan, R. Molecular cloning and characterization of a

novelretinoic acid-inducible gene that encodes a putative G protein-coupled receptor. *J Biol Chem*, 273 :35008-35015, 1998.)。另一方面, 已知本发明所包括的下调基因 TGFBR2 能导致恶性肿瘤缩小 (Sun, L., et al., Expression of transforming growth factor beta type II receptor leads to reduced malignancy in human breast cancer MCF-7 cells. *J Biol Chem*, 269 :26449-26455, 1994.)。这些发现表明这些基因可能参与 DCIS 至 IDC 的转变。

[0288] 特别是鉴定出 25 个上调基因 (表 5) 和 49 个下调基因 (表 6), 它们根据由 DCIS 至 IDC 的转变而具有增加或降低的表达。上调元件的列表中包括在浸润性肿瘤发生中起重要作用的基因, 所述基因编码转录因子和参与信号转导途径和细胞周期的蛋白质。据报道 FoxM1 和细胞周期蛋白 B1 在多种肿瘤类型中过表达。FoxM1 的过表达刺激细胞周期蛋白 B1 的表达 (Leung TW, 2001)。CCNB1 是经过 G2 和有丝分裂的传代所需的细胞周期控制蛋白 (Pines, J. and Hunter, T. Cyclins A and B1 in the human cell cycle. *Ciba Found Symp*, 170 :187-196 ;discussion 196-204, 1992.)。TOP2A 抑制剂被广泛用作治疗肺癌的化学治疗剂 (Miettinen, H. E., et al., High topoisomerase II alpha expression associates with high proliferation rate and and poor prognosis in oligodendrogliomas. *Neuropathol Appl Neurobiol*, 26 :504-512, 2000.)。BUB1B 可能负责在有丝分裂的限制点对肿瘤进程起作用的染色体不稳定性表型和遗传不稳定性 (Bardelli, A., et al. Carcinogen-specific induction of genetic instability. *Proc Natl Acad Sci U S A*, 98 :5770-5775, 2001.)。已证实 MMP11 的表达对患者的存活具有直接的负面作用 (Boulay, A., et al. High cancer cell death in syngeneic tumors developed in host mice deficient for the stromelysin-3 matrix metalloproteinase. *Cancer Res*, 61 :2189-2193, 2001.)。ECM1 具有血管生成特性并由乳腺肿瘤细胞表达 (Han, Z., et al., Extracellular matrix protein 1 (ECM1) has angiogenic properties and is expressed by breast tumor cells. *Faseb J*, 15 :988-994, 2001.)。尽管大多数功能仍是未知的, 但对这些基因功能分析的评价可显示这些基因在介导侵入活性中起作用。

[0289] 在此报告中, 通过利用基因组范围内的 cDNA 微阵列测得的乳腺癌精确表达分布图, 我们分离出与正常人组织相比, 在乳腺癌细胞中显著过表达的新基因 A7870。另外, 我们还阐明了用 siRNA 处理乳腺癌细胞, 所述 siRNA 能有效抑制靶基因 A7870 的表达并能显著抑制乳腺癌细胞 / 肿瘤的生长。这些发现表明 A7870 可能在肿瘤细胞生长增殖中起关键作用, 可能是开发抗 - 癌药物的理想靶标。

[0290] 选择被称为 TOPK (MAPKK 家族的新成员) 的 A7870 进行研究, 因为它在乳腺癌中的表达显著增加。我们鉴定出显示出癌症特异性表达的约 1.8 和 1.9kb 的转录物。这些转录物具有不同的 5' UTR 序列, 但 ORF 序列相同。我们阐明了用 siRNA 处理乳腺癌细胞, 所述 siRNA 能有效抑制 A7870 的表达并能显著抑制乳腺癌细胞 / 肿瘤的生长。这些发现表明 A7870 可能在肿瘤细胞生长增殖中起关键作用, 可能是开发抗 - 癌药物的理想靶标。

[0291] 然而, 一些标准预测疾病进程和临床结果的能力是不完美的。更具攻击力的疾病的患者可获利于辅助化学疗法或激素疗法, 目前可根据下述标准的组合来鉴定这些患者: 年龄、肿瘤大小、腋下淋巴结状态、癌症的组织类型和病理学分级以及激素 - 受体状态。组织类型不同的肿瘤可通过基因亚群来归类, 这是一种能提供病理学相关信息的方法。大多

数研究者提出如果肿瘤显示出显著较高百分比的低度分化的组织学,该患者的预后较差。

[0292] 此项研究令人惊奇的结果是:在每个患者中,不同组织类型的表达分布图有显著的相似性。通过显微解剖和全面的基因表达分析,使用经监督的分析由高度-分化型和低度分化型乳腺癌细胞获得 mRNA 表达分布图,使用该分布图检查与肿瘤侵入和预后有关的基因表达变化。通过在表达分布图基础上进行的未经监督的分类分析,可将乳腺癌分成两组,已证实所述乳腺癌与不同的病理学损害有关。通过随机排列试验鉴定出 25 个基因,在由每个患者组成的这两组之间,所述基因的表达显著不同(图 2C)。在这些基因中,活化的 T-细胞的核因子 5(NFAT5)只限于促进癌细胞迁移,这使人们关注到由这些转录因子诱导的不同基因的可能性(Sebastien J. et al., The role of NFAT transcription factors in integrin-mediated carcinoma invasion. *Nature cell biology*, 4:540-544, 2002.)。血小板反应蛋白 2(THSB2)是细胞外基质蛋白,它似乎在细胞粘附和细胞迁移中起作用。基于 LMM 的方法的一个重要优点是能从一个样品中选择不同表型的癌细胞。基因表达模式的系统分析为肿瘤侵入的生物学和发病机理提供了一个窗口。

[0293] 另外,在乳腺癌患者中,淋巴结转移是肿瘤进程中的关键步骤,并且是较差预后的主要组分之一(Shek, L. L. and Godolphin, W. Model for breast cancer survival: relative prognostic roles of axillary nodal status, TNM stage, estrogen receptor concentration, and tumor necrosis. *Cancer Res*, 48:5565-5569, 1988.),但仅有少数患者在诊断时表现出临床上可测的转移。诊断时的淋巴结状态是未来复发和整体存活的最重要的衡量标准,它充其量只不过是一个不完美的代用品。例如,约 1/3 未见可测的淋巴结转移的患者在 10 年内会旧病复发(Saphner, T., et al., Annual hazard rates of recurrence for breast cancer after primary therapy. *J Clin Oncol*, 14:2738-2746, 1996.)。已证实岗哨淋巴结活检是研究腋下淋巴结的准确方法;它能使乳腺癌与外科手术有关的发病率显著降低,并能避免腋下解剖。其它参数,如核分级、患者年龄、肿瘤大小不能预测腋下淋巴结状态,也不可能通过岗哨淋巴结活检有效诊断淋巴结状态。因此,本发明对于在淋巴结-阳性和淋巴结-阴性肿瘤之间差异表达的基因亚群的鉴定能够改善临床诊断和对精确生物物理事件的了解。聚类分析(图 10)暗示着可将淋巴结转移病例与淋巴结未转移的病例分开。能用于根据淋巴结转移状态区分两个患者组的基因可用作转移的分子标记(Ramaswamy, S., et al., A molecular signature of metastasis in primary solid tumors. *Nat Genet*, 33:49-54, 2003.)。例如,在这 34 个基因中,以在脂质肉瘤中被易位的 TLS 为人所知的 FUS 在淋巴结-阴性癌症中有所降低,它被编码骨髓细胞白血病中的转录因子 ERG-1 的基因易位。野生型 FUS 的一个重要功能是基因组维持,特别是基因组稳定性的维持(Hicks, G. G., et al., Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nat Genet*, 24:175-179, 2000.)。与转移-阴性组相比,转移-阳性组中一些基因的表达水平增加。例如,关于 EEF1D, EF-1 $\delta$  在肿瘤中的较高表达暗示着体内恶性转化需要增加翻译因子 mRNA 和蛋白质合成以进入细胞周期和经过细胞周期进行转换。CFL1, Rho 蛋白信号转导和 Rho 家族 GTPase 调节细胞骨架和细胞迁移,并且常在肿瘤中被过表达(Yoshizaki, H., et al., Activity of Rho-family GTPases during cell division as visualized with FRET-based probes. *J Cell Biol*, 162:223-232, 2003. ;

Arthur, W. T., et al., Regulation of Rho family GTPases by cell-cell and cell-matrix adhesion. *Biol Res*, 35 :239-246, 2002.)。已证实 B-Raf 激酶 BRAF 能因刺激生长因子而导致 MEK 的磷酸化和活化。尽管一些基因的功能仍是未知的,对这些基因产物功能的了解可阐明它们在乳腺癌转移中的作用。

[0294] 目前,肿瘤复发的诱因和临床过程仍是未知的。另外,也不可能根据可获得的临床、病理学和遗传标记可靠地预测结果。尽管据信本发明使用这 34 个基因的表达分布图的预测评分系统可用于改善预后,但导入临床阶段需要使用大量病例来证实。无论如何,本发明似乎能提供与被常规组织学诊断误解的癌细胞生物特性有关的精确信息。

[0295] 针对癌细胞中发生的特异性分子变化的癌疗法已经通过下述抗癌药的临床研发和获批规定认可为有效:治疗晚期乳腺癌的 trastuzumab (Herceptin) (Cousens, L., et al. Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. *Science*, 230 :1132-1139, 1985.)。这些药物临床上是有用的,并且比传统抗癌剂具有更好的耐受性,因为它们仅仅靶向转化的细胞。因此,这类药物不仅能改善癌患者的存活率和生活质量,而且能验证分子靶向的癌疗法概念的正确性。另外,靶向药物当与常规化疗联用时,可以提高常规化疗的效力 (Gianni, L. and Grasselli, G. Targeting the epidermal growth factor receptor a new strategy in cancer treatment. *Suppl Tumori*, 1 :S60-61, 2002.; Klejman, A., et al., Phosphatidylinositol-3 kinase inhibitors enhance the anti-leukemia effect of STI571. *Oncogene*, 21 :5868-5876, 2002.)。因此,未来的癌治疗将很可能涉及将常规药物与针对肿瘤细胞不同特性如血管生成和侵袭性的特异性靶向试剂联用。另外,本发明阐明新的肿瘤标记,即在患有乳腺癌的妇女的血液或乳头吸出液中以异常量存在的物质可以足够可靠地常规用于检测早期乳腺癌。

[0296] 目前,对于晚期乳腺癌患者没有有效的疗法可以使用。因此,迫切需要新的治疗方法和个性化的治疗方法。本发明的癌症-特异性表达分布图,包括在乳腺癌中被上调和下调的基因应能提供有用的信息用于鉴定治疗患者所用的分子靶标。

[0297] 表 1 在组织表型的高度和低度分化型之间表达有所改变的基因列表

[0298]

| BRC 编号 | 登录号      | 标志        | 名称                                     | p-值     |
|--------|----------|-----------|--|---------|
| 1      | AF053712 | TNFSF11   | 肿瘤坏死因子(配体)超家族, 成员 11                   | 1.2E-06 |
| 2      | BF973104 | LOC201725 | 假拟蛋白 LOC201725                         | 3.2E-05 |
| 3      | AV752313 | KPNA6     | 核周蛋白 $\alpha 6$ (输入蛋白 $\alpha 7$ )     | 1.1E-04 |
| 4      | AK026898 | FOXP1     | Forkhead 盒 P1                          | 7.4E-04 |
| 5      | AA148107 | ITGA5     | 整联蛋白, $\alpha 5$ (纤连蛋白受体, $\alpha$ 多肽) | 7.9E-04 |
| 6      | AK001067 | NFAT5     | 活化 T-细胞的核因子 5, 紧张性-反                   | 8.2E-04 |



[0299]

|    |          |          | 应型   |         |
|----|----------|----------|--|---------|
| 7  | AB007919 | KIAA0450 | KIAA0450 基因产物  | 1.8E-03 |
| 8  | BG026429 | SFRS2    | 剪接因子, 富含精氨酸/丝氨酸的因子 2   | 2.0E-03 |
| 9  | M87770   | FGFR2    | 成纤维细胞生长因子受体 2 (细菌-表达的激酶, 角质形成细胞生长因子受体, 颅骨面骨发育不全 1, Crouzon 综合征, Pfeiffer 综合征, Jackson-Weiss 综合征) | 2.1E-03 |
| 10 | L02785   | SLC26A3  | 溶质载体蛋白家族 26, 成员 3  | 2.7E-03 |
| 11 | BF037402 |          | 智人, 克隆 MGC:17296<br>IMAGE:3460701, mRNA, 完整的编码序列   | 2.8E-03 |
| 12 | L12350   | THBS2    | 血小板反应蛋白 2  | 2.8E-03 |
| 13 | N36875   |          | 智人, 克隆 IMAGE:4994678, mRNA   | 3.8E-03 |
| 14 | AL135342 |          | EST, 与神经元线状蛋白[智人]<br>[H.sapiens]有点类似   | 4.3E-03 |
| 15 | AL049426 | SDC3     | 黏结蛋白聚糖 3 (N-黏结蛋白聚糖)  | 4.5E-03 |
| 16 | AW961424 | KIAA1870 | KIAA1870 蛋白  | 5.2E-03 |
| 17 | AA523117 | DC-TM4F2 | 类似于 TM4SF9 的 tetraspanin   | 5.5E-03 |
| 18 | Z11531   | EEF1G    | 真核翻译延伸因子 1 $\gamma$  | 6.1E-03 |
| 19 | AI423028 | SMARCD3  | 与 SWI/SNF 相关的, 与基质相连的, 依赖于肌动蛋白的染色质调节物, 亚族 d, 成员 3  | 6.8E-03 |
| 20 | AB002391 | MN7      | D15F37 (假基因)   | 7.1E-03 |
| 21 | D32050   | AARS     | 丙氨酰-tRNA 合成酶   | 7.2E-03 |
| 22 | BE876949 | RAB7     | RAB7, 成员 RAS 致癌基因家族  | 7.9E-03 |
| 23 | AW291083 |          | EST  | 8.0E-03 |
| 24 | AI568910 |          | EST  | 8.2E-03 |
| 25 | AK023480 | SRP72    | 信号识别颗粒 72kDa   | 8.7E-03 |

[0300] 表 2 在 ER- 阳性和 ER- 阴性肿瘤之间表达有所改变的基因列表

| [0301] | BRC 编号 | 登录号 | 标志 | 名称 | p-值 |
|--------|--------|-----|----|----|-----|
|        |        |     |    |    |     |

[0302]

|    |           |               |  |         |
|----|-----------|---------------|--|---------|
| 26 | AW949747  | GATA3         | GATA 结合蛋白 3  | 3.2E-20 |
| 27 | BE868254  | EST           | EST  | 2.2E-14 |
| 28 | AF037335  | CA12          | 碳酸酐酶 XII   | 1.6E-13 |
| 29 | BF724977  | ASB13         | 含锚蛋白重复序列和 SOCS 盒的 13                               | 8.5E-13 |
| 30 | NM_004636 | SEMA3B        | Sema 结构域, 免疫球蛋白结构域(Ig), 短的基本结构域, 分泌的, (脑信号蛋白) 3B   | 9.7E-13 |
| 31 | NM_000125 | ESR1          | 雌激素受体 1  | 1.2E-12 |
| 32 | M73554    | CCND1         | 细胞周期蛋白 D1 (PRAD1: 甲状旁腺腺瘤病 1)                       | 3.9E-12 |
| 33 | NM_005544 | IRS1          | 胰岛素受体底物 1  | 4.4E-12 |
| 34 | M14745    | BCL2          | B-细胞 CLL/淋巴瘤 2                                     | 5.1E-12 |
| 35 | BE826171  | BCMP11        | 乳腺癌膜蛋白 11  | 2.8E-11 |
| 36 | AI087270  | SIAH2         | seven in absentia 同系物 2 (果蝇)                       | 2.8E-11 |
| 37 | L07033    | HMGCL         | 3-羟甲基-3-甲基戊二酰辅酶 A 裂合酶(羟甲基戊二酸尿症)                    | 2.8E-11 |
| 38 | AB014523  | ULK2          | unc-51-样激酶 2 (C. elegans)                          | 4.0E-11 |
| 39 | AL137588  | DKFZp434K1210 | 假拟蛋白 DKFZp434K1210                                 | 5.2E-11 |
| 40 | AL137566  | EST           | 智人 mRNA; cDNA DKFZp586G0321 (得自克隆 DKFZp586G0321)   | 5.4E-11 |
| 41 | AF038421  | GFRA1         | GNDF 家族受体 $\alpha 1$                               | 8.4E-11 |
| 42 | AI194045  | FE65L2        | FE65-样蛋白 2   | 9.2E-11 |
| 43 | BG163478  | EST           | EST, 与 BAI1_HUMAN 脑-特异性血管生成抑制剂 1 前体[H.sapiens]有点类似 | 1.1E-10 |
| 44 | M31627    | XBP1          | X-盒结合蛋白 1  | 1.1E-10 |
|    | AA156269  | EST           | 智人, 克隆 IMAGE:4794107, mRNA                         | 1.3E-10 |
| 46 | NM_006763 | BTG2          | BTG 家族, 成员 2                                       | 1.9E-10 |
| 47 | AW504052  | SEC15L        | SEC15(酿酒酵母)-样                                      | 2.1E-10 |
| 48 | NM_005400 | PRKCE         | 蛋白质激酶 C, $\epsilon$                                | 2.3E-10 |
| 49 | AI628151  | XBP1          | X-盒结合蛋白 1  | 2.7E-10 |
| 50 | AF043045  | FLNB          | 细丝蛋白 B, $\beta$ (肌动蛋白结合蛋白 278)                     | 3.5E-10 |
| 51 | U31383    | GNG10         | 鸟嘌呤核苷酸结合蛋白(G 蛋白), $\gamma 10$                      | 4.6E-10 |
| 52 | L10333    | RTN1          | reticulon 1  | 5.6E-10 |
| 53 | AK025099  | SIGIRR        | 单个 Ig IL-1R-相关分子                                   | 6.2E-10 |

[0303]

|    |           |          |  |         |
|----|-----------|----------|--|---------|
| 54 | AL039253  | LIV-1    | LIV-1 蛋白, 调节雌激素的                               | 7.4E-10 |
| 55 | AW949662  | KIAA0239 | KIAA0239 蛋白                                    | 8.0E-10 |
| 56 | D13629    | KTN1     | 移动结合蛋白 1 (驱动蛋白受体)                              | 1.5E-09 |
| 57 | NM_000165 | GJA1     | 间隙连接蛋白, $\alpha$ 1, 43kDa (间隙连接蛋白 43)          | 1.5E-09 |
| 58 | AA533079  | C1orf21  | 染色体 1 开放读码框 21                                 | 1.8E-09 |
| 59 | AF251056  | CAPS2    | calcyphosphine 2                               | 1.9E-09 |
| 60 | AF061016  | UGDH     | UDP-葡萄糖脱氢酶                                     | 2.0E-09 |
| 61 | U92544    | MAGED2   | 黑素瘤抗原, 家族 D, 2                                 | 2.1E-09 |
| 62 | BE617536  | RPL13A   | 核糖体蛋白 L13a                                     | 2.4E-09 |
| 63 | AK024102  | MYST1    | MYST 组蛋白乙酰转移酶 1                                | 2.5E-09 |
| 64 | BF212902  | EST      | 智人 mRNA; cDNA DKFZp564F053 (得自克隆 DKFZp564F053) | 2.8E-09 |
| 65 | AK025480  | FLJ21827 | 假拟蛋白 FLJ21827                                  | 3.0E-09 |
| 66 | AI376713  | EST      | EST, 与假拟蛋白 FLJ20378 [智人] [H.sapiens] 有点类似      | 3.6E-09 |
| 67 | AI028483  | EST      | EST  | 3.8E-09 |
| 68 | AK022249  | EST      | 智人 cDNA FLJ12187 fis, 克隆 MAMMA1000831          | 4.2E-09 |
| 69 | AI568527  | EST      | 智人 cDNA FLJ34849 fis, 克隆 NT2NE2011687          | 5.0E-09 |
| 70 | AL133074  | TP53INP1 | 肿瘤蛋白 p53 可诱导的核蛋白 1                             | 5.3E-09 |
| 71 | AF022116  | PRKAB1   | 蛋白质激酶, AMP-活化的, $\beta$ 1 非-催化的亚单位             | 6.1E-09 |
| 72 | AF007170  | C1orf34  | 染色体 1 开放读码框 34                                 | 9.7E-09 |
| 73 | AF042081  | SH3BGR1  | 结合 SH3 结构域的富含谷氨酸的蛋白质样                          | 1.2E-08 |
| 74 | AK027813  | MGC10744 | 假拟蛋白 MGC10744                                  | 1.4E-08 |
| 75 | M57609    | GLI3     | GLI-Kruppel 家族成员 GLI3 (Greig 头多指畸形综合征)         | 1.7E-08 |
| 76 | AL359600  | EST      | 智人 mRNA; cDNA DKFZp547C136 (得自克隆 DKFZp547C136) | 1.9E-08 |
| 77 | BQ006049  | TIMP1    | 金属蛋白酶 1 的组织抑制剂 (红细胞强化活性, 胶原酶抑制剂)               | 2.1E-08 |
| 78 | AF111849  | HELO1    | 酵母长链多不饱和脂肪酸延伸酶 2 的同系                           | 2.2E-08 |

[0304]

|     |           |              |  |         |
|-----|-----------|--------------|--|---------|
|     |           |              | 物  |         |
| 79  | AL157499  | RAB5EP       | Rab 接触蛋白-5   | 2.2E-08 |
| 80  | AK023199  | EST          | 智人 cDNA FLJ13137 fis, 克隆<br>NT2RP3003150                     | 2.5E-08 |
| 81  | J05176    | SERPINA<br>3 | 丝氨酸(或半胱氨酸)蛋白酶抑制剂, 分支<br>体 A ( $\alpha$ -1 抗蛋白酶, 抗胰蛋白酶), 成员 3 | 3.2E-08 |
| 82  | AA028101  | KIAA0303     | KIAA0303 蛋白  | 3.3E-08 |
| 83  | AI300588  | MAP2K4       | 促分裂原-活化的蛋白质激酶激酶 4  | 4.1E-08 |
| 84  | AA682861  | EST          | EST, 与假拟蛋白 FLJ20378 [智人]<br>[H.sapiens]中等类似                  | 4.6E-08 |
| 85  | M26393    | ACADS        | 酰基-辅酶 A 脱氢酶, C-2 至 C-3 短链                                    | 5.4E-08 |
| 86  | NM_001609 | ACADSB       | 酰基-辅酶 A 脱氢酶, 短的/分支的链   | 5.5E-08 |
| 87  | U91543    | CHD3         | 克罗莫结构域解旋酶 DNA 结合蛋白 3   | 5.7E-08 |
| 88  | AK023813  | FLJ10081     | 假拟蛋白 FLJ10081  | 6.0E-08 |
| 89  | BF111711  | FLJ20727     | 假拟蛋白 FLJ20727  | 7.0E-08 |
| 90  | AL049987  | EST          | 智人 mRNA; cDNA DKFZp564F112 (得自<br>克隆 DKFZp564F112)           | 7.2E-08 |
| 91  | AW081894  | EST          | EST  | 8.2E-08 |
| 92  | AK000350  | FLJ20343     | 假拟蛋白 FLJ20343  | 1.1E-07 |
| 93  | AA418493  | DPP7         | 二肽基肽酶 7  | 1.1E-07 |
| 94  | BE674061  | PIN4         | 蛋白质(肽基-脯氨酸顺/反异构酶) NIMA-<br>相互作用的, 4 (细小蛋白)                   | 1.2E-07 |
| 95  | AB011155  | DLG5         | discs, 大的(果蝇)同系物 5   | 1.2E-07 |
| 96  | L15203    | TFF3         | 三叶因子 3 (肠的)  | 1.4E-07 |
| 97  | NM_001552 | IGFBP4       | 胰岛素-样生长因子结合蛋白 4  | 1.4E-07 |
| 98  | M57230    | IL6ST        | 白介素 6 信号转导蛋白(gp130, 制癌蛋白<br>M 受体)                            | 1.5E-07 |
| 99  | N92706    | EST          | 智人 cDNA FLJ38461 fis, 克隆<br>FEBRA2020977                     | 1.5E-07 |
| 100 | M30704    | AREG         | 双调蛋白(神经鞘瘤-衍生的生长因子)   | 1.8E-07 |
| 101 | AB004066  | BHLHB2       | 含有碱性螺旋-环-螺旋结构域, B 类, 2                                       | 2.2E-07 |
| 102 | M15518    | PLAT         | 纤溶酶原激活物, 组织  | 2.3E-07 |
| 103 | BM697477  | ShrmL        | Shroom-相关蛋白  | 2.4E-07 |

[0305]

|     |           |           |  |         |
|-----|-----------|-----------|--|---------|
| 104 | R45979    | CELSR1    | 钙粘着蛋白, EGF LAG 七跨膜区 G-型受体 1 (火烈鸟同系物, 果蝇)         | 3.0E-07 |
| 105 | AL049365  | EST       | 智人 mRNA; cDNA DKFZp586A0618 (得自克隆 DKFZp586A0618) | 6.5E-07 |
| 106 | NM_003225 | TFF1      | 三叶因子 1 (乳腺癌中表达的雌激素-可诱导的序列)                       | 7.1E-07 |
| 107 | AI733356  | EST       | 智人 cDNA FLJ31746 fis, 克隆 NT2RI2007334            | 7.8E-07 |
| 108 | AF078853  | KIAA1243  | KIAA1243 蛋白                                      | 8.2E-07 |
| 109 | N30179    | PLAB      | 前列腺分化因子  | 1.0E-06 |
| 110 | BG026429  | SFRS2     | 剪接因子, 富含精氨酸/丝氨酸的 2                               | 2.4E-06 |
| 111 | AU149272  | EST       | EST  | 2.5E-06 |
| 112 | J03827    | NSEP1     | 核酸酶敏感元件结合蛋白 1                                    | 3.0E-06 |
| 113 | AJ276469  | C20orf35  | 染色体 20 开放读码框 35                                  | 3.4E-06 |
| 114 | AW295100  | LOC201562 | 假拟蛋白 LOC201562                                   | 3.9E-06 |
| 115 | J03817    | GSTM1     | 谷胱甘肽 S-转移酶 M1                                    | 4.8E-06 |
| 116 | AF288571  | LEF1      | 淋巴增强子-结合因子 1                                     | 5.1E-06 |
| 117 | AF069301  | PECI      | 过氧化物酶体 D3,D2-烯酰-CoA 异构酶                          | 5.3E-06 |
| 118 | AA621665  | EST       | EST  | 6.7E-06 |
| 119 | AI739486  | EST       | EST  | 8.0E-06 |
| 120 | X81438    | AMPH      | 双载蛋白(具有乳腺癌 128kDa 自身抗原的强直人综合征)                   | 8.7E-06 |
| 121 | U89606    | PDXK      | 吡哆醛(吡哆醇, 维生素 B6)激酶                               | 8.8E-06 |
| 122 | NM_017555 | EGLN2     | egl nine 同系物 2 (C. elegans)                      | 9.2E-06 |

[0306] 表 3 在 DCIS 和 IDC 中通常被上调的基因

| BRC 编号     | 登录号     | 标志        | 名称                                 |
|------------|---------|-----------|------------------------------------|
| 123        | D90041  | NAT1      | N-乙酰转移酶 1 (芳香胺 N-乙酰转移酶)            |
| [0307] 124 | M13755  | G1P2      | 干扰素, $\alpha$ -可诱导的蛋白质(克隆 IFI-15K) |
| 125        | D88308  | SLC27A2   | 溶质载体家族 27 (脂肪酸转运蛋白), 成员 2          |
| 126        | AW23506 | NM_004170 | SLC1A1 溶质载体家族 1 (神经元/上皮高亲和性谷       |

[0308]

|     |           |           |         |  |
|-----|-----------|-----------|---------|--|
|     | 1         |           |         | 氨酸转运蛋白, 系统 Xag), 成员 1  |
| 127 | K02215    |           | AGT     | 血管紧张肽原(丝氨酸(或半胱氨酸)蛋白酶抑制剂, 分支体 A ( $\alpha$ -1 抗蛋白酶, 抗胰蛋白酶), 成员 8) |
| 128 | AB032261  |           | SCD     | 酰基-CoA 去饱和酶( $\delta$ -9-去饱和酶)                                   |
| 129 | NM_000909 |           | NPY1R   | 神经肽 Y 受体 Y1  |
| 130 | AF017790  |           | HEC     | 在癌症中高表达, 富含 7 个赖氨酸的重复序列  |
| 131 | NM_007019 |           | UBE2C   | 遍在蛋白-缀合酶 E2C   |
| 132 | AF065388  |           | TSPAN-1 | tetraspan 1  |
| 133 | N70334    |           | DUSP10  | 双特异性磷酸酶 10   |
| 134 | AA621719  | NM_005496 | SMC4L1  | SMC4 染色体 4 结构维持-样 1 (酵母)   |
| 135 | AA676987  |           |         | EST  |
| 136 | AK001402  | NM_018131 | C10orf3 | 染色体 10 开放读码框 3   |
| 137 | AW949747  | NM_002051 | GATA3   | GATA 结合蛋白 3  |
| 138 | AK001472  | NM_018685 | ANLN    | 胞环蛋白, 肌动蛋白结合蛋白(scrap 同系物, 果蝇)                                    |
| 139 | AA789233  | NM_000088 | COL1A1  | 胶原, I 型, $\alpha$ 1  |
| 140 | AF070632  |           |         | 智人克隆 24405 mRNA 序列   |
| 141 | H04544    |           | NPY1R   | 神经肽 Y 受体 Y1  |
| 142 | AI015982  |           | CDCA1   | 细胞分裂周期相关 1   |
| 143 | NM_003979 |           | RAI3    | 诱导视黄酸的 3   |
| 144 | BF516445  | NM_053277 | CLIC6   | 氯化物细胞内通道 6   |
| 145 | AI361654  |           |         |  |
| 146 | AI077540  | NM_178530 |         | 智人 cDNA FLJ38379 fis, 克隆 FEBRA2002986                            |
| 147 | AI261804  |           |         | 智人 MSTP020 (MST020) mRNA, 完整的 cds                                |
| 148 | AK026559  |           | TPM3    | 原肌球蛋白 3  |
| 149 | J03473    |           | ADPRT   | ADP-核糖基转移酶(NAD <sup>+</sup> ; 聚(ADP-核糖)聚合酶)                      |

[0309]

|     |           |           |           |  |
|-----|-----------|-----------|-----------|--|
| 150 | NM_000187 |           | HGD       | 尿黑酸 1,2-双加氧酶(尿黑酸氧化酶)                           |
| 151 | L43964    |           | PSEN2     | 早老素 2 (老年性痴呆 4)                                |
| 152 | J05581    |           | MUC1      | 粘蛋白 1, 跨膜                                      |
| 153 | AA602499  | XM_379784 | GLCCI1    | 诱导糖皮质激素的转录物 1                                  |
| 154 | U37707    |           | MPP3      | 膜蛋白质, 棕榈酰化的 3 (MAGUK p55 亚族成员 3)               |
| 155 | AB030905  |           | CBX3      | chromobox 同系物 3 (HP1 $\gamma$ 同系物, 果蝇)         |
| 156 | AL138409  | NM_198278 |           | 智人 mRNA; cDNA DKFZp313L231 (得自克隆 DKFZp313L231) |
| 157 | AV756928  |           | SEC61G    | Sec61 $\gamma$                                 |
| 158 | AI205684  | NM_021979 | HSPA2     | 热休克 70kDa 蛋白 2                                 |
| 159 | BE739464  | NM_015161 | ARL6IP    | ADP-核糖基化因子-样 6 相互作用蛋白                          |
| 160 | AI081356  | NM_203463 | LOC253782 | 假拟蛋白 LOC253782                                 |
| 161 | AA167194  |           | LOC253782 | 假拟蛋白 LOC253782                                 |
| 162 | M90516    |           | GFPT1     | 谷氨酰胺-果糖-6-磷酸转氨酶 1                              |
| 163 | AL133074  | NM_033285 | TP53INP1  | 肿瘤蛋白质 p53 可诱导的核蛋白 1                            |
| 164 | AL137257  |           |           | 智人, 克隆 IMAGE:5296692, mRNA                     |
| 165 | AK025240  | NM_147128 | LOC223082 | LOC223082                                      |
| 166 | AJ007042  |           | WHSC1     | Wolf-Hirschhorn 综合征候选物 1                       |
| 167 | U42068    |           | GRP58     | 葡萄糖调节蛋白, 58kDa                                 |
| 168 | AJ132592  |           | ZNF281    | 锌指蛋白 281                                       |
| 169 | W93638    |           |           | EST  |
|     | AW977394  |           | C9orf12   | 染色体 9 开放读码框 12                                 |
| 171 | AI347925  | NM_001540 | HSPB1     | 热休克 27kDa 蛋白 1                                 |
| 172 | AK026587  |           | NET-6     | 跨膜 4 超家族成员 tetraspan NET-6                     |
| 173 | AI264621  |           | LASS2     | LAG1 确保存活力的同系物 2 (酿酒酵母)                        |
| 174 | AA767828  | XM_035527 | FLJ10980  | 假拟蛋白 FLJ10980                                  |
| 175 | AU142881  | NM_018184 | FLJ10702  | 假拟蛋白 FLJ10702                                  |

[0310] 表 4 在 DCIS 和 IDC 中通常被下调的基因

[0311]

| BRC<br>编号 | 登录号       | 标志        | 名称   |
|-----------|-----------|-----------|--|
| 176       | X52186    |           | ITGB4 整联蛋白, $\beta$ 4                                |
| 177       | NM_006297 |           | XRCC1 中国仓鼠细胞中的 X-射线修复互补缺陷的修复 1                       |
| 178       | X73460    |           | RPL3 核糖体蛋白 L3  |
| 179       | NM_001436 |           | FBL 核仁纤维蛋白   |
| 180       | X59373    |           | HOXD10 homeo 盒 D10                                   |
| 181       | J04208    |           | IMPDH2 IMP (肌苷一磷酸) 脱氢酶 2                             |
| 182       | L24203    |           | TRIM29 含三联基元的 29                                     |
| 183       | L10340    | NM_001958 | EEF1A2 真核翻译延伸因子 1 $\alpha$ 2                         |
| 184       | J04621    |           | SDC2 黏结蛋白聚糖 2 (硫酸乙酰肝素蛋白聚糖 1, 与细胞表面-相连的, 纤维蛋白聚糖)      |
| 185       | L08424    |           | ASCL1 缺刚毛-胸部刚毛复合体-样 1 (果蝇)                           |
| 186       | AI376713  |           | EST, 与假拟蛋白 FLJ20378 [智人] [H.sapiens] 有点类似            |
| 187       | AK026966  |           | EST 智人 cDNA: FLJ23313 fis, 克隆 HEP11919               |
| 188       | NM_001050 |           | SSTR2 促生长素抑制素受体 2                                    |
| 189       | AA632025  |           | EST EST  |
| 190       | N22918    | NM_144641 | FLJ32332 假拟蛋白 FLJ32332                               |
| 191       | AF272043  |           | ITM2C 膜内在蛋白质 2C                                      |
| 192       | M58459    |           | RPS4Y 核糖体蛋白 S4, Y-连锁的                                |
| 193       | AI133697  |           | EST 智人, 克隆 MGC:16362 IMAGE:3927795, mRNA, 完整的 cds    |
| 194       | AA780301  | NM_003793 | CTSF 组织蛋白酶 F   |
| 195       | M92843    |           | ZFP36 锌指蛋白 36, C3H 型, 同系物(小鼠)                        |
| 196       | AA570186  |           | EST 智人(人)HeLa 细胞克隆 CS0DK007YB08 的人全长 cDNA 5-PRIME 末端 |
| 197       | R56906    |           | EST EST  |
| 198       | AF208860  | NM_014452 | TNFRSF21 肿瘤坏死因子受体超家族, 成员 21                          |
| 199       | AK025216  |           | TAZ 具有 PDZ-结合基元的转录辅激活物(TAZ)                          |
| 200       | AA758394  |           | PTPN1 蛋白质酪氨酸磷酸酶, 非-受体型 1                             |

[0312]



|     |           |           |          |   |
|-----|-----------|-----------|----------|---|
| 201 | AA628530  | NM_016368 | ISYNA1   | 肌醇 1-磷酸合酶 A1  |
| 202 | AF161416  | NM_003749 | IRS2     | 胰岛素受体底物 2   |
| 203 | AL045916  |           | EST      | EST   |
| 204 | AW340972  |           | EST      | 智人 cDNA: FLJ22864 fis, 克隆 KAT02164                    |
| 205 | AI189414  |           | RNPC2    | 含 RNA-结合区(RNP1, RRM)的 2                               |
| 206 | AV705636  |           | EIF3S6IP | 真核翻译起始因子 3, 亚单位 6 相互作用蛋白                              |
| 207 | U28977    |           | CASP4    | 胱天蛋白酶 4, 与细胞凋亡有关的半胱氨酸蛋白酶                              |
| 208 | AV708528  | NM_018579 | MSCP     | 线粒体溶质载体蛋白   |
| 209 | AA022956  | NM_024667 | FLJ12750 | 假拟蛋白 FLJ12750   |
| 210 | AI928443  |           | EST      | 智人 cDNA FLJ38855 fis, 克隆 MESAN2010681                 |
| 211 | U14966    |           | RPL5     | 核糖体蛋白 L5  |
| 212 | AI857997  |           | TPBG     | 滋养层糖蛋白  |
| 213 | BF697545  |           | MGP      | 基质 Gla 蛋白   |
| 214 | AW575754  | NM_152309 | FLJ35564 | 假拟蛋白 FLJ35564   |
| 215 | AI352534  | NM_001753 | CAV1     | 窖蛋白 1, 胞膜窖蛋白, 22kDa                                   |
| 216 | NM_001985 |           | ETFB     | 电子转移黄素蛋白, $\beta$ 多肽                                  |
| 217 | AI743134  | NM_006216 | SERPINE2 | 丝氨酸(或半胱氨酸)蛋白酶抑制剂, 分支体 E (微管连接蛋白, 1 型纤溶酶原激活物抑制剂), 成员 2 |
| 218 | AW444709  | NM_001777 | CD47     | CD47 抗原(Rh-相关抗原, 与整联蛋白-相连的信号转导蛋白)                     |
| 219 | BF688910  | NM_001300 | COPEB    | 核心启动子元件结合蛋白   |
| 220 | AI818579  | NM_181847 | EST      | 智人, 克隆 IMAGE:3625286, mRNA, 部分 cds                    |
| 221 | S95936    |           | TF       | 运铁蛋白  |
| 222 | AF074393  |           | RPS6KA5  | 核糖体蛋白 S6 激酶, 90kDa, 多肽 5                              |
| 223 | NM_000591 |           | CD14     | CD14 抗原   |
| 224 | AK027181  | NM_031426 | IBA2     | 与离子钙结合的衔接分子 2   |
| 225 | X73079    |           | PIGR     | 聚合的免疫球蛋白受体  |
| 226 | NM_00134  |           | DAB2     | 残疾的同系物 2, 促分裂原-反应性磷蛋白                                 |

[0313]

|     |           |           |         |                                |
|-----|-----------|-----------|---------|--------------------------------|
|     | 3         |           |         | (果蝇)                           |
| 227 | M31452    |           | C4BPA   | 补体成分 4 结合蛋白, $\alpha$          |
| 228 | X07696    |           | KRT15   | 角蛋白 15                         |
| 229 | AF016004  |           | GPM6B   | 糖蛋白 M6B                        |
| 230 | NM_004078 |           | CSRP1   | 富含半胱氨酸和甘氨酸的蛋白质 1               |
| 231 | L36645    |           | EPHA4   | EphA4                          |
| 232 | D78011    |           | DPYS    | 二氢嘧啶酶                          |
| 233 | W60630    | NM_032801 | JAM3    | 连接黏附分子 3                       |
| 234 | AW956111  |           | D4S234E | 染色体 4(独特的) 234 表达序列上的 DNA 区段   |
| 235 | AF035752  |           | CAV2    | 窖蛋白 2                          |
| 236 | D37766    |           | LAMB3   | 层粘连蛋白, $\beta$ 3               |
| 237 | U66406    |           | EFNB3   | 肝配蛋白-B3                        |
| 238 | X52001    |           | EDN3    | 内皮缩血管肽 3                       |
| 239 | NM_000856 |           | GUCY1A3 | 鸟苷酸环化酶 1, 可溶的, $\alpha$ 3      |
| 240 | U60115    |           | FHL1    | 4 个和一半的 LIM 结构域 1              |
| 241 | D14520    | NM_001730 | KLF5    | Kruppel-样因子 5(肠的)              |
| 242 | M99487    |           | FOLH1   | 叶酸水解酶(前列腺-特异性膜抗原) 1            |
| 243 | U09873    |           | FSCN1   | 肌动蛋白束蛋白同系物 1, 肌动蛋白-束蛋白(紫色球海胆)  |
| 244 | AF017418  |           | MEIS2   | Meis1, 骨髓亲嗜性病毒整合位点 1 同系物 2(小鼠) |
| 245 | AF038540  | NM_206900 | RTN2    | reticulon 2                    |
| 246 | AF049884  | NM_021069 | ARGBP2  | Arg/Abl-相互作用蛋白 ArgBP2          |
| 247 | NM_001122 |           | ADFP    | 脂肪分化-相关蛋白                      |
| 248 | Y09926    |           | MASP2   | 甘露聚糖-结合凝集素丝氨酸蛋白酶 2             |
| 249 | M58297    |           | ZNF42   | 锌指蛋白 42(骨髓-特异性视黄酸-反应性的)        |
| 250 | AF035811  |           | PNUTL2  | 花生-样 2(果蝇)                     |
| 251 | L22214    |           | ADORA1  | 腺苷 A1 受体                       |
| 252 | AF177775  |           | CES1    | 羧酸酯酶 1(单核细胞/巨噬细胞丝氨酸酯酶 1)       |
| 253 | U07643    |           | LTF     | 乳运铁蛋白                          |

[0314]

|     |           |           |          |  |
|-----|-----------|-----------|----------|--|
| 254 | S76474    | NM_006180 | NTRK2    | 神经营养性酪氨酸激酶, 受体, 2 型                                |
| 255 | BE299605  | NM_012219 | MRAS     | 肌肉 RAS 致癌基因同系物                                     |
| 256 | NM_006225 |           | PLCD1    | 磷脂酶 C, $\delta 1$                                  |
| 257 | NM_005036 |           | PPARA    | 过氧化物酶体的增殖活化受体, $\alpha$                            |
| 258 | M22324    |           | ANPEP    | 丙氨酰(膜)氨基肽酶(氨肽酶 N, 氨肽酶 M, 微粒体氨肽酶, CD13, p150)       |
| 259 | BE877416  |           | TGFBR2   | 转化生长因子, $\beta$ 受体 II (70/80kDa)                   |
| 260 | BE561244  |           | RPL18A   | 核糖体蛋白 L18a   |
| 261 | AL048962  |           | EST      | 智人, 克隆 IMAGE:4243767, mRNA                         |
| 262 | L08895    |           | MEF2C    | MADS 盒转录增强因子 2, 多肽 C (肌细胞增强因子 2C)                  |
| 263 | U48707    |           | PPP1R1A  | 蛋白质磷酸酶 1, 调节性的(抑制剂)亚单位 1A                          |
| 264 | X56134    |           | RPLP2    | 核糖体蛋白, 大的 P2                                       |
| 265 | D84239    |           | FCGBP    | IgG 结合蛋白的 Fc 片段                                    |
| 266 | AK026181  |           | PHLDA1   | 血小板-白细胞 C 激酶底物同源性-样结构域, 家族 A, 成员 1                 |
| 267 | K01144    |           | CD74     | CD74 抗原(主要组织相容性复合物的不变多肽, II 类抗原-相关的)               |
| 268 | U25138    |           | KCNMB1   | 钾大电导率钙-活化通道, 亚族 M, $\beta$ 成员 1                    |
| 269 | X85337    | NM_053025 | MYLK     | 肌球蛋白, 轻链多肽激酶                                       |
| 270 | D83597    |           | LY64     | 淋巴细胞抗原 64 同系物, 放射性保护的 105kDa (小鼠)                  |
| 271 | NM_004024 |           | ATF3     | 转录激活因子 3   |
| 272 | BF126636  |           | SAA1     | 血清淀粉状蛋白 A1   |
| 273 | D13789    |           | MGAT3    | 甘露糖基( $\beta$ -1,4-)-糖蛋白 $\beta$ -1,4-N-乙酰葡萄糖氨基转移酶 |
| 274 | L41142    |           | STAT5A   | 信号转导及转录激活蛋白 5A                                     |
| 275 | AB040969  |           | KIAA1536 | KIAA1536 蛋白  |
| 276 | NM_002153 |           | HSD17B2  | 羟类固醇(17- $\beta$ )脱氢酶 2                            |
| 277 | AV646610  | NM_001546 | ID4      | DNA 结合的抑制剂 4, 显性失活的螺旋-环-                           |

[0315]

|     |           |           |              |  |
|-----|-----------|-----------|--------------|--|
|     |           |           |              | 螺旋蛋白   |
| 278 | X03663    |           | CSF1R        | 集落刺激因子 1 受体, 以前的 McDonough 猫肉瘤病毒(v-fms)致癌基因同系物   |
| 279 | U47025    |           | PYGB         | 磷酸化酶, 糖原; 脑  |
| 280 | M81349    |           | SAA4         | 血清淀粉状蛋白 A4, 组成型  |
| 281 | AI264201  | NM_000399 | EGR2         | 早期生长应答蛋白 2 (Krox-20 同系物, 果蝇)   |
| 282 | U18018    |           | ETV4         | Ets 变体基因 4 (E1A 增强子结合蛋白, E1AF)   |
| 283 | NM_004350 |           | RUNX3        | 小牛-相关转录因子 3  |
| 284 | BF337516  |           | CRYAB        | 晶体蛋白, $\alpha$ B   |
| 285 | AF027208  |           | PROML1       | prominin-样 1 (小鼠)  |
| 286 | D17408    |           | CNN1         | 钙调理蛋白 1, 碱性, 平滑肌   |
| 287 | NM_004010 |           | DMD          | 肌养蛋白(肌肉营养不良症, Duchenne 和 Becker 型)   |
| 288 | BF183952  |           | CSTA         | 半胱氨酸蛋白酶抑制剂 A (stefin A)  |
| 289 | M16445    |           | CD2          | CD2 抗原(p50), 绵羊红细胞受体   |
| 290 | AF055015  |           | EYA2         | 缺眼同系物 2 (果蝇)   |
| 291 | AI745624  |           | ELL2         | ELL-相关的 RNA 聚合酶 II, 延伸因子   |
| 292 | AK025329  |           | DKFZP566H073 | DKFZP566H073 蛋白  |
| 293 | BE745465  | NM_012427 | KLK5         | 激肽释放酶 5  |
| 294 | AK024578  | NM_031455 | DKFZP761F241 | 假拟蛋白 DKFZp761F241  |
| 295 | AI870306  | XM_380171 | IRX1         | Iroquois 同源异形盒蛋白 1   |
| 296 | H37853    | NM_022343 | C9orf19      | 染色体 9 开放读码框 19   |
| 297 | BF000047  |           | EST          | 智人全长插入物 cDNA 克隆 ZA79C08  |
| 298 | AF126780  |           | RetSDR2      | 视网膜短链脱氢酶/还原酶 2   |
| 299 | AI700341  |           | EST          | EST, 与假拟蛋白 FLJ20489 [智人] [H.sapiens]有点类似   |
| 300 | M87770    |           | FGFR2        | 成纤维细胞生长因子受体 2 (细菌-表达的激酶, 角质形成细胞生长因子受体, 颅骨面骨发育不全 1, Crouzon 综合征, Pfeiffer 综合征, Jackson-Weiss 综合征) |
| 301 | AA452368  | NM_144595 | FLJ30046     | 假拟蛋白 FLJ30046  |

[0316]

|     |           |           |           |  |
|-----|-----------|-----------|-----------|--|
| 302 | NM_021200 |           | PLEKHB1   | 含有血小板-白细胞 C 激酶底物同源性结构域, 家族 B (evectins) 成员 1     |
| 303 | AK026343  |           | hIAN2     | 人免疫相关的核苷酸 2                                      |
| 304 | AF251040  |           | C5orf6    | 染色体 5 开放读码框 6                                    |
| 305 | M87507    |           | CASP1     | 胱天蛋白酶 1, 与细胞凋亡-相关的半胱氨酸蛋白酶(白介素 1, $\beta$ , 转化酶)  |
| 306 | M97675    |           | ROR1      | 受体酪氨酸激酶-样孤独受体 1                                  |
| 307 | NM_020549 |           | CHAT      | 胆碱乙酰转移酶  |
| 308 | X00457    | NM_033554 | HLA-DPA1  | 主要组织相容性复合物, II 类, DP $\alpha$ 1                  |
| 309 | W72411    | NM_003722 | TP73L     | 肿瘤蛋白质 p73-样                                      |
| 310 | AI769569  |           | EST       | EST  |
| 311 | K02765    |           | C3        | 补体成分 3   |
| 312 | AW971490  |           | FLJ14906  | 假拟蛋白 FLJ14906                                    |
| 313 | AF077044  |           | RPAC2     | 可能是小鼠 RNA 聚合酶 1-3 的直向同源物 (16 kDa 亚单位)            |
| 314 | H70803    | NM_015278 | KIAA0790  | KIAA0790 蛋白                                      |
| 315 | AL050367  | XM_167709 | LOC221061 | 假拟蛋白 LOC221061                                   |
| 316 | AK001643  | NM_018215 | FLJ10781  | 假拟蛋白 FLJ10781                                    |
| 317 | AW182273  |           | EST       | 智人 cDNA FLJ31517 fis, 克隆 NT2RI2000007            |
| 318 | W67951    |           | EST       | 在染色体 6-受抑制的黑素瘤细胞中表达的人 S6 A-5 mRNA                |
| 319 | AL117605  |           | EST       | 智人 mRNA; cDNA DKFZp564N1063 (得自克隆 DKFZp564N1063) |
| 320 | AI376418  |           | EST       | 智人 cDNA FLJ35169 fis, 克隆 PLACE6012908            |
| 321 | AA683373  |           | EST       | EST  |
| 322 | AK022877  |           | EST       | 智人 cDNA FLJ12815 fis, 克隆 NT2RP2002546            |
| 323 | NM_002258 |           | KLRB1     | 杀伤细胞凝集素-样受体亚族 B, 成员 1                            |
| 324 | M69225    |           | BPAG1     | 大疱性类天疱疮抗原 1, 230/240kDa                          |
| 325 | AW299572  | NM_015461 | EHZF      | 早期生血锌指蛋白   |

[0317]

|     |           |           |              |   |
|-----|-----------|-----------|--------------|---|
| 326 | BE044467  | NM_005737 | ARL7         | ADP-核糖基化因子-样 7  |
| 327 | AA938297  | NM_017938 | FLJ20716     | 假拟蛋白 FLJ20716   |
| 328 | AA706316  | NM_033317 | ZD52F10      | 假拟基因 ZD52F10  |
| 329 | AI827230  | NM_153000 | APCDD1       | 腺瘤病结肠息肉病下调基因 1  |
| 330 | AK000251  |           | FLJ20244     | 假拟蛋白 FLJ20244   |
| 331 | N62352    | NM_020925 | KIAA1573     | KIAA1573 蛋白   |
| 332 | H53164    |           | ICSBP1       | 干扰素共有序列结合蛋白 1   |
| 333 | BE394824  |           | WFDC2        | WAP 四-二硫化物核心结构域 2   |
| 334 | AL117462  | NM_015481 | ZFP385       | 可能是小鼠锌指蛋白 385 的直向同源物  |
| 335 | NM_003186 |           | TAGLN        | 转凝蛋白  |
| 336 | U58514    |           | CHI3L2       | 几丁质酶 3-样 2  |
| 337 | AB026125  |           | ART-4        | ART-4 蛋白  |
| 338 | AL080059  | NM_033512 | KIAA1750     | KIAA1750 蛋白   |
| 339 | AA747005  |           | SDCCAG4<br>3 | 血清学限定的结肠癌抗原 43  |
| 340 | NM_005928 |           | MFG8         | 乳脂小球-EGF 因子 8 蛋白  |
| 341 | D62470    | NM_004796 | NRXN3        | 神经元表面蛋白 3   |
| 342 | N29574    |           | RAGD         | Rag D 蛋白  |
| 343 | K02276    |           | MYC          | v-myc 鸡白细胞增多症病毒致癌基因同系物(禽类)  |
| 344 | D78611    |           | MEST         | 中胚层特异性转录物同系物(小鼠)  |
| 345 | NM_022003 |           | FXYD6        | 含 FXYD 结构域的离子转运调节物 6  |
| 346 | BF508973  |           | RPL13        | 核糖体蛋白 L13   |
| 347 | NM_001615 |           | ACTG2        | 肌动蛋白, $\gamma$ 2, 平滑肌, 肠  |
| 348 | R41532    |           | EST          | EST, 与 POL2_MOUSE 逆转录病毒-相关 POL 多蛋白[含有: 逆转录酶; 内切核酸酶][M.musculus]有点类似 |
| 349 | AA142875  |           | EST          | EST   |
| 350 | U03688    |           | CYP1B1       | 细胞色素 P450, 家族 1, 亚族 B, 多肽 1   |
| 351 | W94363    |           | EST          | 智人全长插入物 cDNA 克隆 ZE12G01   |
| 352 | W44613    |           | HSJ001348    | 差异表达的 CO16 基因的 cDNA   |

[0318]

|     |           |           |               |  |
|-----|-----------|-----------|---------------|--|
| 353 | AL118812  |           | EST           | 智人 mRNA; cDNA DKFZp761G1111 (得自克隆 DKFZp761G1111) |
| 354 | D56064    |           | MAP2          | 微管-相关蛋白 2  |
| 355 | BF966838  | NM_172069 | KIAA2028      | 类似于 PH (血小板-白细胞 C 激酶底物同源性) 结构域                   |
| 356 | AI338625  | NM_014344 | FJX1          | 4 个被接在一起的盒 1 (果蝇)                                |
| 357 | AI263022  |           | EST           | EST  |
| 358 | AL050107  | NM_015472 | TAZ           | 具有 PDZ-结合基元的转录辅激活物(TAZ)                          |
| 359 | AI056364  | NM_033210 | FLJ14855      | 假拟蛋白 FLJ14855                                    |
| 360 | AI351898  | NM_032581 | DRCTNNB1A     | 被 Ctnnb1, a 下调                                   |
| 361 | AV700003  |           | ARL6IP2       | ADP-核糖基化-样因子 6 相互作用蛋白 2                          |
| 362 | NM_000700 |           | ANXA1         | 膜联蛋白 A1  |
| 363 | M81141    |           | HLA-DQB1      | 主要组织相容性复合物, II 类, DQ $\beta$ 1                   |
| 364 | AI598227  | NM_024911 | FLJ23091      | 假拟蛋白 FLJ23091                                    |
| 365 | BG034740  |           | ROPN1         | ropporin, Rho 结合蛋白相关蛋白 1                         |
| 366 | AB011175  |           | TBC1D4        | TBC1 结构域家族, 成员 4                                 |
| 367 | AK024449  |           | PP2135        | PP2135 蛋白  |
| 368 | AW978770  |           | DKFZP566A1524 | 假拟蛋白 DKFZp566A1524                               |
| 369 | AI821113  |           | EST           | 智人 cDNA FLJ36327 fis, 克隆 THYMU2005748            |
| 370 | AI057450  |           | SLC13A2       | 溶质载体家族 13 (依赖于钠的二羧酸转运蛋白), 成员 2                   |
| 371 | X86693    |           | SPARCL1       | SPARC-样 1 (mast9, hevin)                         |
| 372 | AI224952  | NM_173640 | FLJ40906      | 假拟蛋白 FLJ40906                                    |
| 373 | D13639    |           | CCND2         | 细胞周期蛋白 D2  |

[0319] 表 5 在由 DCIS 转变至 IDC 时表达增加的基因

[0320]

| BRC<br>编号 | 登录号    | 标志    | 名称            |
|-----------|--------|-------|---------------|
| 374       | U74612 | FOXMI | forkhead 盒 M1 |
| 375       | U63743 | KIF2C | 驱动蛋白家族成员 2C   |

[0321]



|     |               |               |                   |  |
|-----|---------------|---------------|-------------------|--|
| 376 | D88532        |               | PIK3R3            | 磷酸肌醇-3-激酶, 调节性亚单位, 多肽 3 (p55, $\gamma$ )               |
| 377 | NM_00553<br>2 |               | IFI27             | 干扰素, $\alpha$ -可诱导蛋白 27                                |
| 378 | D14657        |               | KIAA0101          | KIAA0101 基因产物  |
| 379 | AF030186      |               | GPC4              | 磷脂酰肌醇聚糖 4  |
| 380 | Z11566        |               | STMN1             | stathmin 1/癌蛋白 18                                      |
| 381 | U90914        | NM_0013<br>04 | CPD               | 羧肽酶 D  |
| 382 | NM_00253<br>4 |               | OAS1              | 2',5'-寡腺苷酸合成酶 1, 40/46kDa                              |
| 383 | S67310        |               | BF                | B-因子, 备解素  |
| 384 | AA192445      | NM_0201<br>82 | TMEPAI            | 跨膜, 前列腺雄激素诱导的 RNA                                      |
| 385 | AB003103      |               | PSMD12            | 蛋白酶体(前体, macropain) 26S 亚单位, 非-ATPase, 12              |
| 386 | BE878057      | NM_0307<br>96 | DKFZP564<br>K0822 | 假拟蛋白 DKFZp564K0822                                     |
| 387 | AB003698      |               | CDC7L1            | CDC7 细胞分裂周期 7-样 1 (酿酒酵母)                               |
| 388 | M91670        |               | E2-EPF            | 遍在蛋白载体蛋白   |
| 389 | AK023414      |               | FLJ13352          | 假拟蛋白 FLJ13352  |
| 390 | L09235        |               | ATP6V1A1          | ATPase, H <sup>+</sup> 转运, 溶酶体的 70kDa, V1 亚单位 A, 同种型 1 |
| 391 | AF007152      |               | ABHD3             | 含 abhydrolase 结构域的 3                                   |
| 392 | U33632        |               | KCNK1             | 钾通道, 亚族 K, 成员 1  |
| 393 | AA621719      | NM_0054<br>96 | SMC4L1            | SMC4 染色体 4 结构维持-样 1 (酵母)                               |
| 394 | AF176228      |               | DNMT3B            | DNA (胞嘧啶-5-)-甲基转移酶 3 $\beta$                           |
| 395 | H22566        | NM_0807<br>59 | DACH              | 德国猎犬同系物(果蝇)  |
| 396 | AI185804      | NM_2124<br>82 | FN1               | 纤连蛋白 1   |
| 397 | AI189477      | NM_0021<br>68 | IDH2              | 异柠檬酸脱氢酶 2 (NADP <sup>+</sup> ), 线粒体的                   |
| 398 | AA205444      |               | AP1S2             | 衔接子-相关蛋白复合物 1, 82 亚单位                                  |

[0322] 表 6 在由 DCIS 转变至 IDC 时表达降低的基因

[0323]

|     | 登录号           |               | 标志       | 名称   |
|-----|---------------|---------------|----------|--|
| 399 | AF070609      | NM_0041<br>72 | SLC1A3   | 溶质载体家族 1 (神经胶质高亲和性谷氨酸转运蛋白), 成员 3                 |
| 400 | U85267        |               | DSCR1    | Down 综合征临界区域基因 1                                 |
| 401 | NM_00539<br>7 |               | PODXL    | 足萼蛋白-样   |
| 402 | D13811        |               | AMT      | 氨基甲基转移酶(甘氨酸裂解系统蛋白 T)                             |
| 403 | X53586        |               | ITGA6    | 整联蛋白, $\alpha 6$                                 |
| 404 | L13288        |               | VIPR1    | 肠道血管活性肽受体 1                                      |
| 405 | M12125        |               | TPM2     | 原肌球蛋白 2 ( $\beta$ )                              |
| 406 | M65066        | NM_0027<br>35 | PRKAR1B  | 蛋白质激酶, cAMP-依赖型, 调节性的, I 型, $\beta$              |
| 407 | AJ001183      |               | SOX10    | SRY (性别决定区 Y)-盒 10                               |
| 408 | AW241712      |               | MXI1     | MAX 相互作用蛋白 1                                     |
| 409 | AL160111      |               | KIAA1649 | KIAA1649 蛋白                                      |
| 410 | X93920        |               | DUSP6    | 双特异性磷酸酶 6  |
| 411 | AF132734      | NM_0218<br>07 | SEC8     | 分泌蛋白 SEC8  |
| 412 | AI133467      |               |          | EST  |
| 413 | D88153        |               | HYA22    | HYA22 蛋白   |
| 414 | AF014404      |               | PTE1     | 过氧化物酶体酰基-CoA 硫酯酶                                 |
| 415 | BE907755      | NM_0133<br>99 | C16orf5  | 染色体 16 开放读码框 5                                   |
| 416 | AA135341      | NM_0210<br>78 | GCN5L2   | GCN5 全面控制氨基酸合成 5-样 2 (酵母)                        |
| 417 | AL110126      |               |          | 智人 mRNA; cDNA DKFZp564H1916 (得自克隆 DKFZp564H1916) |
| 418 | BE254330      | NM_0030<br>45 |          | 智人 mRNA; cDNA DKFZp564D016 (得自克隆 DKFZp564D016)   |
| 419 | BE264353      |               | RBP1     | 视黄醇结合蛋白 1, 细胞的                                   |
| 420 | W75991        |               |          | 智人, 克隆 IMAGE:4249217, mRNA                       |
| 421 | AF091434      |               | PDGFC    | 血小板衍生的生长因子 C                                     |

[0324]

|     |           |           |          |   |
|-----|-----------|-----------|----------|---|
| 422 | W67577    |           | CD74     | CD74 抗原(主要组织相容性复合物的不变多肽, II 类抗原-相关的)      |
| 423 | NM_002996 |           | CX3CL1   | 趋化因子(C-X3-C 基元)配体 1                       |
| 424 | AA024459  |           |          | EST                                       |
| 425 | NM_000163 |           | GHR      | 生长激素受体                                    |
| 426 | AA858162  | NM_032160 | NCAG1    | NCAG1                                     |
| 427 | BE327623  |           |          | EST, 与假拟蛋白 FLJ20234 [智人] [H.sapiens] 有点类似 |
| 428 | BE671156  |           | MAPRE2   | 微管-相关蛋白, RP/EB 家族, 成员 2                   |
| 429 | D12614    |           | LTA      | 淋巴毒素 $\alpha$ (TNF 超家族, 成员 1)             |
| 430 | L13720    |           | MGC5560  | 假拟蛋白 MGC5560                              |
| 431 | U15131    |           | ST5      | 抑制致瘤性 5                                   |
| 432 | Y00711    |           | LDHB     | 乳酸脱氢酶 B                                   |
| 433 | AI651212  |           |          | 智人 cDNA FLJ31125 fis, 克隆 IMR322000819     |
| 434 | M31159    |           | IGFBP3   | 胰岛素-样生长因子结合蛋白 3                           |
| 435 | NM_014447 |           | HSU52521 | arfaptin 1                                |
| 436 | AB011089  |           | TRIM2    | 含三联基元的 2                                  |
| 437 | BF969355  | NM_002612 | PDK4     | 丙酮酸脱氢酶激酶, 同工酶 4                           |
| 438 | AK025950  | XM_371114 | KIAA1695 | 假拟蛋白 FLJ22297                             |
| 439 | D86961    | NM_005779 | LHFPL2   | 脂肪瘤 HMGIC 融合配对物-样 2                       |
| 440 | AK025953  |           |          | 智人 cDNA: FLJ22300 fis, 克隆 HRC04759        |
| 441 | AJ223812  |           | CALD1    | 钙调结合蛋白 1                                  |
| 442 | R40594    |           |          | 智人 cDNA: FLJ22845 fis, 克隆 KAIA5195        |
| 443 | AF145713  |           | SCHIP1   | 施旺膜蛋白相互作用蛋白 1                             |
| 444 | AK024966  |           | FLJ21313 | 假拟蛋白 FLJ21313                             |
| 445 | NM_005596 |           | NFIB     | 核因子 I/B                                   |

[0325]

|     |           |           |         |                             |
|-----|-----------|-----------|---------|-----------------------------|
| 446 | NM_001613 |           | ACTA2   | 肌动蛋白, $\alpha 2$ , 平滑肌, 主动脉 |
| 447 | H03641    | XM_376328 | FAM13A1 | 具有序列相似性的家族 13, 成员 A1        |

[0326] 表 7 在 IDC 中通常被上调的基因

[0327]

| BRC 编号 | 登录号       | 标志        | 名称   |
|--------|-----------|-----------|--|
| 448    | X14420    |           | COL3A1 胶原, III 型, $\alpha 1$ (Ehlers-Danlos 综合征 IV 型, 常染色体显性)  |
| 449    | AF044588  |           | PRC1 胞质分裂的蛋白质调节物 1   |
|        | AF161499  |           | HSPC150 类似于遍在蛋白-缀合酶的 HSPC150 蛋白  |
| 451    | AA789233  | NM_000088 | COL1A1 胶原, I 型, $\alpha 1$   |
| 452    | U16306    |           | CSPG2 硫酸软骨素蛋白聚糖 2 (多能聚糖)   |
| 453    | NM_004425 |           | ECM1 细胞外基质蛋白 1   |
| 454    | NM_006855 |           | KDEL3 KDEL (Lys-Asp-Glu-Leu) 内质网蛋白保留受体 3   |
| 455    | AI972071  | NM_031966 | CCNB1 细胞周期蛋白 B1  |
| 456    | AF237709  | NM_018492 | TOPK T-LAK 细胞-来源的蛋白质激酶   |
| 457    | BE747327  |           | HIST1H1C 组蛋白 1, H1c  |
| 458    | J03464    |           | COL1A2 胶原, I 型, $\alpha 2$   |
| 459    | AI080640  | NM_006408 | AGR2 前梯度 2 同系物(Xenopus laevis)   |
| 460    | AA971042  |           | RHPN1 Rho 结合蛋白, Rho GTPase 结合蛋白 1  |
| 461    | AI419398  |           | MGC33662 假拟蛋白 MGC33662   |
| 462    | AI149552  | NM_004448 | EST, 与 ERB2_HUMAN 受体蛋白-酪氨酸激酶 erbB-2 前体(p185erbB2) (NEU 原癌基因) (C-erbB-2) (酪氨酸激酶-型细胞表面受体 HER2) (MLN 19) [H.sapiens] 中等类似 |
| 463    | D14874    |           | ADM 肾上腺髓质素   |
| 464    | X03674    | NM_000402 | G6PD 葡萄糖-6-磷酸脱氢酶   |
| 465    | NM_0023   |           | MAD2L1 MAD2 有丝分裂抑制缺陷-样 1 (酵母)  |

[0328]

|     |              |           |        |                                     |
|-----|--------------|-----------|--------|-------------------------------------|
|     | 58           |           |        |                                     |
| 466 | BF214508     |           | CYCS   | 细胞色素 c, 体细胞的                        |
| 467 | BG030536     | NM_001067 | TOP2A  | 拓扑异构酶(DNA) II $\alpha$ 170kDa       |
| 468 | X57766       |           | MMP11  | 基质金属蛋白酶 11 (溶基质素 3)                 |
| 469 | AA02990<br>0 | NM_015170 | SULF1  | 硫酸酯酶 1                              |
| 470 | AF053306     |           | BUB1B  | BUB1 出芽不受苯并咪唑抑制的 1 同系物 $\beta$ (酵母) |
| 471 | AF074002     |           | LGALS8 | 凝集素, 半乳糖苷-结合的, 可溶性的, 8 (半乳凝素 8)     |

[0329] 表 8 在 IDC 中通常被下调的基因

[0330]

| BRC<br>编号 | 登录号       | 标志            | 名称     |                                    |
|-----------|-----------|---------------|--------|------------------------------------|
| 472       | NM_004484 |               | GPC3   | 磷脂酰肌醇聚糖 3                          |
| 473       | NM_006219 |               | PIK3CB | 磷酸肌醇-3-激酶, 催化的, $\beta$ 多肽         |
| 474       | BE793000  |               | RBP1   | 视黄醇结合蛋白 1, 细胞的                     |
| 475       | AL117565  | NM_033<br>027 | AXUD1  | AXIN1 上调的 1                        |
| 476       | BF055342  |               | ZNF6   | 锌指蛋白 6 (CMPX1)                     |
| 477       | U03688    |               | CYP1B1 | 细胞色素 P450, 家族 1, 亚族 B, 多肽 1        |
| 478       | AF038193  | NM_004<br>311 |        | 智人, 克隆 IMAGE:3610040, mRNA         |
| 479       | X72760    | NM_002<br>292 | LAMB2  | 层粘连蛋白, $\beta$ 2 (层粘连蛋白 S)         |
| 480       | J03817    |               | GSTM1  | 谷胱甘肽 S-转移酶 M1                      |
| 481       | M69226    |               | MAOA   | 单胺氧化酶 A                            |
| 482       | BF690180  | NM_006<br>990 | WASF2  | WAS 蛋白家族, 成员 2                     |
| 483       | AL133600  |               | STAM2  | 信号转导衔接分子(SH3 结构域和 ITAM 基元) 2       |
| 484       | AF215981  |               | GPR2   | 与 G 蛋白-偶联的受体 2                     |
| 485       | BG149764  |               |        | 智人, 克隆 IMAGE:5286091, mRNA, 部分 cds |

[0331]

|     |           |               |               |  |
|-----|-----------|---------------|---------------|--|
| 486 | AF067800  |               | CLECSF6       | C-型(依赖于钙的, 碳水化合物-识别结构域)凝集素, 超家族成员 6              |
| 487 | AA713487  |               | PIK3R1        | 磷酸肌醇-3-激酶, 调节性亚单位, 多肽 1 (p85 $\alpha$ )          |
| 488 | AA828505  |               | FBXW7         | F-盒和 WD-40 结构域蛋白 7 (archipelago 同系物, 果蝇)         |
| 489 | AK021865  |               | CKIP-1        | CK2 相互作用蛋白 1; HQ0024c 蛋白                         |
| 490 | AK001605  |               | FLJ10743      | 假拟蛋白 FLJ10743                                    |
| 491 | AI041186  |               | HSPC182       | HSPC182 蛋白                                       |
| 492 | AA873363  | NM_144<br>650 | ADH8          | 醇脱氢酶 8   |
| 493 | NM_013409 |               | FST           | 促滤泡素抑制素  |
| 494 | AK000322  |               | FLJ20315      | 假拟蛋白 FLJ20315                                    |
| 495 | AB020637  | XM_290<br>546 | KIAA0830      | KIAA0830 蛋白                                      |
| 496 | AA872040  |               | INHBB         | 抑制素, $\beta$ B (激活蛋白 AB $\beta$ 多肽)              |
| 497 | NM_004430 |               | EGR3          | 早期生长应答蛋白 3                                       |
| 498 | D59989    |               |               | EST  |
| 499 | D78013    |               | DPYSL2        | 二氢嘧啶酶-样 2  |
| 500 | AI081821  |               |               | 智人 mRNA; cDNA DKFZp313M0417 (得自克隆 DKFZp313M0417) |
| 501 | AA309603  |               | KIAA1430      | KIAA1430 蛋白                                      |
| 502 | NM_004107 |               | FCGRT         | IgG 的 Fc 片段, 受体, 转运蛋白, $\alpha$                  |
| 503 | AW268719  |               |               | 智人 cDNA FLJ32438 fis, 克隆 SKMUS2001402            |
| 504 | BF446578  | NM_145<br>313 | LOC22100<br>2 | CG4853 基因产物                                      |
| 505 | BG054844  | NM_005<br>168 | ARHE          | ras 同系物基因家族, 成员 E                                |
| 506 | AF054987  |               | ALDOC         | 醛缩酶 C, 果糖-二磷酸                                    |
| 507 | AI052390  |               | FLJ20071      | dymeclin   |
| 508 | NM_004530 |               | MMP2          | 基质金属蛋白酶 2 (明胶酶 A, 72kDa 明胶酶, 72kDa IV 型胶原酶)      |
| 509 | AF054999  | NM_001<br>431 | EPB41L2       | 红细胞膜蛋白条带 4.1-样 2                                 |

[0332]

|     |          |               |          |                                       |
|-----|----------|---------------|----------|---------------------------------------|
| 510 | AU151591 | NM_182<br>964 | NAV2     | 神经元导航分子 2                             |
| 511 | AA447744 |               |          | EST                                   |
| 512 | R61253   |               | ST6GalII | $\beta$ -半乳糖苷 $\alpha$ -2,6-唾液酸转移酶 II |

[0333] 表 9 半 - 定量 RT-PCR 实验所用的引物序列

[0334]

| 登录号       | 标志            | 正向引物  | 反向引物   |
|-----------|---------------|---|--|
| A1261804  | EST           | 5' -CTGTTCTGGC<br>TTCGTTATGT TCT-3' (SEQ<br>ID NO :1) | 5' -AGAAAATACG<br>GTCCTCTTGT TGC-3' (SEQ<br>ID NO :2)  |
| AA205444  | AP1S2         | 5' -CACTGTAATG<br>CACGACATTT GA-3' (SEQ<br>ID NO :3)  | 5' -GTTACAGCTT<br>AGCACAAGGC ATC-3' (SEQ<br>ID NO :4)  |
| AA167194  | LOC25<br>3782 | 5' -ACCTCTGAGT<br>TTGATTTCCC AA-3' (SEQ ID<br>NO :5)  | 5' -CGAGGCTTGT<br>AACAATCTAC TGG-3' (SEQ<br>ID NO :6)  |
| AA676987  | EST           | 5' -GAAACTGTAC<br>GGGGGTAAA GAG-3' (SEQ<br>ID NO :7)  | 5' -CATCAATGTG<br>GTGAGTGACA TCT-3' (SEQ<br>ID NO :8)  |
| H22566    | DACH          | 5' -AAGCCCTTGG<br>AACAGAACAT ACT-3' (SEQ<br>ID NO :9) | 5' -CAGTAAACGT<br>GGTTCTCACA TTG-3' (SEQ<br>ID NO :10) |
| NM_018492 | TOPK          | 5' -AGACCCTAAAGATCGTC<br>CTTCTG-3' (SEQ ID NO :13)    | 5' -GTGTTTTAAGTCAGCATG<br>AGCAG-3' (SEQ ID NO :14)     |
| NM_002046 | GAPD          | 5' -CGACCACTTT<br>GTCAAGCTCA-3' (SEQ ID<br>NO :11)    | 5' -GGTTGAGCAC<br>AGGGTACTTT ATT-3' (SEQ<br>ID NO :12) |

[0335] 表 10 在单个病例的高度和低度分化型之间表达有所改变的基因列表

[0336]

| BR<br>C<br>编<br>号 | 登录号 | 标志 | 名称 | p-值 |
|-------------------|-----|----|----|-----|
|                   |     |    |    |     |

|     |           |           |              |                                       |         |
|-----|-----------|-----------|--------------|---------------------------------------|---------|
| 513 | AV729269  | XM_371074 | DKFZP564D166 | 推定的含锚蛋白重复序列的蛋白质                       | 3.1E-07 |
| 514 | AI246554  | NM_014222 | NDUFA8       | NADH脱氢酶(泛醌)1 $\alpha$ 亚复合物, 8, 19kDa  | 1.4E-06 |
| 515 | J04080    |           | C1S          | 补体成分 1, s 亚成分                         | 1.4E-05 |
| 516 | N93264    |           | EST          | 智人, 克隆 IMAGE:4908933, mRNA            | 1.4E-05 |
| 517 | NM_002318 |           | LOXL2        | 赖氨酰氧化酶-样 2                            | 1.6E-05 |
| 518 | J03464    |           | COL1A2       | 胶原, I 型, $\alpha$ 2                   | 2.4E-05 |
| 519 | U01184    | NM_002018 | FLII         | 不能飞的 I 同系物(果蝇)                        | 2.5E-05 |
| 520 | X63556    |           | FBN1         | 肌原纤蛋白 1 (马凡综合征)                       | 3.8E-05 |
| 521 | X78137    |           | PCBP1        | 聚(rC)结合蛋白 1                           | 4.6E-05 |
| 522 | AK021534  |           | EST          | 智人 cDNA FLJ11472 fis, 克隆 HEMBA1001711 | 6.3E-05 |
| 523 | AK024012  |           | NPD002       | NPD002 蛋白                             | 6.3E-05 |
| 524 | AI200892  |           | BIK          | BCL2-相互作用的杀伤者(诱导细胞凋亡)                 | 9.1E-05 |
| 525 | J03040    |           | SPARC        | 分泌的蛋白质, 酸性, 富含半胱氨酸(骨粘连蛋白)             | 9.3E-05 |
| 526 | AW970143  |           | C6orf49      | 染色体 6 开放读码框 49                        | 1.0E-04 |
| 527 | D62873    |           | EST          | 智人, 克隆 IMAGE:5288080, mRNA            | 1.2E-04 |
| 528 | D42041    |           | G2AN         | $\alpha$ 葡糖苷酶 II $\alpha$ 亚单位         | 1.2E-04 |
| 529 | AI376418  |           | EST          | 智人 cDNA FLJ35169 fis, 克隆 PLACE6012908 | 1.7E-04 |
| 530 | AK026744  | NM_024911 | FLJ23091     | 假拟蛋白 FLJ23091                         | 1.8E-04 |
| 531 | AF026292  |           | CCT7         | 含 TCP1 的陪伴蛋白, 亚单位 7 (eta)             | 2.0E-04 |
| 532 | Y10805    |           | HRMT1L2      | HMT1 hnRNP 甲基转移酶-样 2 (酿酒酵母)           | 2.1E-04 |
| 533 | L12350    |           | THBS2        | 血小板反应蛋白 2                             | 2.1E-04 |
| 534 | AK025706  |           | AMPD2        | 腺苷一磷酸脱氨酶 2 (同种型 L)                    | 2.4E-04 |
| 535 | BE618804  |           | PIG11        | p53-诱导蛋白                              | 2.5E-04 |
| 536 | AV713686  |           | RPS29        | 核糖体蛋白 S29                             | 2.8E-04 |
| 537 | M26481    |           | TACSTD1      | 肿瘤-相关的钙信号转导蛋白 1                       | 2.8E-04 |

[0337]



|        |     |           |           |          |   |         |
|--------|-----|-----------|-----------|----------|---|---------|
|        | 538 | D00099    |           | ATP1A1   | ATPase, Na <sup>+</sup> /K <sup>+</sup> 转运, $\alpha$ 1 多肽 | 2.9E-04 |
|        | 539 | AA946602  |           | ORMDL2   | ORM1-样 2 (酿酒酵母)   | 2.9E-04 |
|        | 540 | NM_001533 |           | HNRPL    | 核内不均一核糖核蛋白 L  | 3.9E-04 |
|        | 541 | BG107866  |           | SIVA     | CD27-结合(希鹞)蛋白   | 4.4E-04 |
|        | 542 | W72297    | NM_017866 | FLJ20533 | 假拟蛋白 FLJ20533   | 4.4E-04 |
|        | 543 | U76992    |           | HTATSF1  | HIV TAT 特异性因子 1   | 4.8E-04 |
|        | 544 | AA191454  | NM_198897 | FIBP     | 成纤维细胞生长因子(酸性)胞内结合蛋白                                       | 4.9E-04 |
|        | 545 | BE903483  |           | RPS20    | 核糖体蛋白 S20   | 5.4E-04 |
|        | 546 | AJ005282  |           | NPR2     | 利尿钠肽受体 B/鸟苷酸环化酶 B (心利尿钠肽受体 B)                             | 5.5E-04 |
|        | 547 | D86322    |           | CLGN     | calmegin  | 5.7E-04 |
|        | 548 | AA621665  |           | EST      | EST   | 5.8E-04 |
|        | 549 | M77349    |           | TGFBI    | 转化生长因子, $\beta$ -诱导的, 68kDa                               | 6.3E-04 |
|        | 550 | BE176466  |           | ZAP3     | ZAP3 蛋白   | 6.6E-04 |
| [0338] | 551 | AA776882  | NM_030795 | STMN4    | stathmin-样 4  | 7.1E-04 |
|        | 552 | AI261382  | NM_016334 | SH120    | 推定的与 G-蛋白偶联的受体  | 7.1E-04 |
|        | 553 | AB007618  |           | COX7A2L  | 细胞色素 c 氧化酶亚单位 VIIa 多肽 2 样                                 | 7.2E-04 |
|        | 554 | D21261    |           | TAGLN2   | 转凝蛋白 2  | 7.5E-04 |
|        | 555 | M68864    |           | LOC51035 | ORF   | 7.7E-04 |
|        | 556 | AB007836  |           | TGFB1I1  | 转化生长因子 $\beta$ 1 诱导的转录物 1                                 | 8.1E-04 |
|        | 557 | AA173339  |           | EST      | EST   | 8.4E-04 |
|        | 558 | D87810    |           | PMM1     | 磷酸甘露糖变位酶 1  | 8.4E-04 |
|        | 559 | M15798    | NM_183356 | ASNS     | 天冬酰胺合成酶   | 8.7E-04 |
|        | 560 | AW072418  |           | B7       | B7 蛋白   | 9.0E-04 |
|        | 561 | D38293    |           | AP3M2    | 衔接子-相关蛋白复合物 3, $\mu$ 2 亚单位                                | 9.5E-04 |
|        | 562 | NM_018950 |           | HLA-F    | 主要组织相容性复合物, I 类, F  | 1.0E-03 |

[0339]

|     |           |           |         |   |         |
|-----|-----------|-----------|---------|---|---------|
| 563 | NM_001219 |           | CALU    | calumenin   | 1.1E-03 |
| 564 | J04162    |           | FCGR3A  | IgG的Fc片段, 低亲和性 IIIa, (CD16)的受体                    | 1.1E-03 |
| 565 | U09873    |           | FSCN1   | 肌动蛋白束蛋白同系物1, 肌动蛋白束蛋白(紫色球海胆)                       | 1.1E-03 |
| 566 | N51082    | NM_080759 | DACH    | 德国猎犬同系物(果蝇)                                       | 1.3E-03 |
| 567 | NM_004199 |           | P4HA2   | 前胶原-脯氨酸, 2-酮戊二酸 4-双加氧酶(脯氨酸 4-羟化酶), $\alpha$ 多肽 II | 1.3E-03 |
| 568 | BE904196  |           | GNB1    | 鸟嘌呤核苷酸结合蛋白(G蛋白), $\beta$ 多肽 1                     | 1.3E-03 |
| 569 | L08895    |           | MEF2C   | MADS盒转录增强因子2, 多肽 C (肌细胞增强因子 2C)                   | 1.3E-03 |
| 570 | AK022670  | NM_016649 | C20orf6 | 染色体 20 开放读码框 6                                    | 1.3E-03 |
| 571 | AW157725  |           | POLR2F  | 聚合酶(RNA) II (DNA 介导的)多肽 F                         | 1.4E-03 |
| 572 | NM_004939 |           | DDX1    | DEAD/H (Asp-Glu-Ala-Asp/His)盒多肽 1                 | 1.4E-03 |
| 573 | X65463    | NM_021976 | RXRB    | 类视黄醇 X 受体, $\beta$                                | 1.5E-03 |
| 574 | Z68179    |           | LY6E    | 淋巴细胞抗原 6 复合物, 基因座 E                               | 1.5E-03 |
| 575 | BF976420  |           | SNRPF   | 核内小核糖核蛋白 F 多肽                                     | 1.5E-03 |
| 576 | D79986    |           | BTF     | Bcl-2-相关转录因子                                      | 1.5E-03 |
| 577 | AK001023  |           | NUBP2   | 核苷酸结合蛋白 2 (MinD 同系物, 大肠杆菌)                        | 1.6E-03 |
| 578 | BE065329  |           | EST     | EST   | 1.6E-03 |
| 579 | L34600    |           | MTIF2   | 线粒体翻译起始因子 2                                       | 1.7E-03 |
| 580 | D13630    |           | BZW1    | 碱性亮氨酸拉链结构域和 W2 结构域 I                              | 1.7E-03 |
| 581 | X15880    | NM_001848 | COL6A1  | 胶原, VI 型, $\alpha$ 1                              | 1.7E-03 |
| 582 | AB003723  |           | PIGQ    | 磷脂酰肌醇聚糖, Q 类                                      | 1.7E-03 |
| 583 | L36645    |           | EPHA4   | EphA4   | 1.7E-03 |
| 584 | BF974358  |           | RPS27   | 核糖体蛋白 S27 (metallopanstimulin                     | 1.8E-03 |

|        |          |           |           |  |         |
|--------|----------|-----------|-----------|--|---------|
|        |          |           |           | 1)                                       |         |
| 585    | AA747449 |           | HIP2      | 亨廷顿相互作用蛋白 2                              | 1.9E-03 |
| 586    | AA283813 |           | FLJ12150  | 假拟蛋白 FLJ12150                            | 2.0E-03 |
| 587    | L38995   | NM_003321 | TUFM      | Tu 翻译延伸因子, 线粒体的                          | 2.0E-03 |
| 588    | N67293   |           | EST       | 智人 cDNA FLJ11997 fis, 克隆<br>HEMBB1001458 | 2.1E-03 |
| 589    | AB014549 |           | KIAA0649  | KIAA0649 基因产物                            | 2.1E-03 |
| 590    | D38305   |           | TOB1      | ERBB2 的转导蛋白, 1                           | 2.2E-03 |
| 591    | L40391   | NM_006827 | TMP21     | 跨膜运输蛋白                                   | 2.2E-03 |
| 592    | H28960   |           | EST       | EST                                      | 2.2E-03 |
| 593    | U86753   |           | CDC5L     | CDC5 细胞分裂周期 5-样(S. pombe)                | 2.3E-03 |
| 594    | AI143226 |           | BLP1      | BBP-样蛋白 1                                | 2.3E-03 |
| 595    | M57730   |           | EFNA1     | 肝配蛋白-A1                                  | 2.3E-03 |
| 596    | AI928868 |           | UBR1      | 遍在蛋白蛋白质连接酶 E3 组分 n-识别蛋白 1                | 2.3E-03 |
| [0340] | 597      | AF077044  | RPAC2     | 可能是小鼠 RNA 聚合酶 1-3 的直向同源物(16 kDa 亚单位)     | 2.3E-03 |
|        | 598      | AF097431  | LEPRE1    | 富含亮氨酸脯氨酸的蛋白聚糖 (leprecan) 1               | 2.4E-03 |
|        | 599      | NM_004350 | RUNX3     | 小牛-相关转录因子 3                              | 2.4E-03 |
|        | 600      | AL162047  | NCOA4     | 核受体辅激活物 4                                | 2.5E-03 |
|        | 601      | BF915013  | EST       | 智人 cDNA FLJ37302 fis, 克隆<br>BRAMY2016009 | 2.5E-03 |
|        | 602      | Z37166    | BAT1      | HLA-B 相关转录物 1                            | 2.5E-03 |
|        | 603      | M81349    | SAA4      | 血清淀粉状蛋白 A4, 组成型                          | 2.6E-03 |
|        | 604      | AL137338  | NM_007214 | SEC63 蛋白                                 | 2.6E-03 |
|        | 605      | AI745624  | ELL2      | ELL-相关 RNA 聚合酶 II, 延伸因子                  | 2.6E-03 |
|        | 606      | BG167522  | HSPC016   | 假拟蛋白 HSPC016                             | 2.6E-03 |
|        | 607      | U58766    | TSTA3     | 组织特异性移植抗原 P35B                           | 2.7E-03 |
|        | 608      | J04474    | NM_000709 | 分支链酮酸脱氢酶 E1, $\alpha$ 多肽(槭糖尿病)           | 2.7E-03 |

[0341]

|     |           |           |           |  |         |
|-----|-----------|-----------|-----------|--|---------|
| 609 | H15977    | NM_021116 | EST       | 智人 cDNA FLJ30781 fis, 克隆 FEBRA2000874        | 2.8E-03 |
| 610 | AL049339  | NM_001304 | CPD       | 羧肽酶 D  | 2.8E-03 |
| 611 | AL133555  | NM_080821 | C20orf108 | 染色体 20 开放读码框 108                             | 2.9E-03 |
| 612 | AW662518  |           | FLJ10876  | 假拟蛋白 FLJ10876                                | 2.9E-03 |
| 613 | BE883507  | NM_003663 | CGGBP1    | CGG 三联体重复序列结合蛋白 1                            | 2.9E-03 |
| 614 | BE797472  |           | RPL17     | 核糖体蛋白 L17                                    | 3.0E-03 |
| 615 | U41371    |           | SF3B2     | 剪接因子 3b, 亚单位 2, 145kDa                       | 3.0E-03 |
| 616 | L39068    |           | DHPS      | N <sup>ε</sup> -(4-氨基丁基)赖氨酸合酶                | 3.1E-03 |
| 617 | NM_004517 |           | ILK       | 与整联蛋白-连接的激酶                                  | 3.1E-03 |
| 618 | U14972    |           | RPS10     | 核糖体蛋白 S10                                    | 3.2E-03 |
| 619 | U61500    |           | TMEM1     | 跨膜蛋白 1                                       | 3.3E-03 |
| 620 | NM_002719 |           | PPP2R5C   | 蛋白质磷酸酶 2, 调节性亚单位 B (B56), γ 同种型              | 3.3E-03 |
| 621 | AF053233  |           | VAMP8     | 小泡-相关膜蛋白 8 (内体小蛋白)                           | 3.3E-03 |
| 622 | NM_002822 | NM_198974 | PTK9      | PTK9 蛋白酪氨酸激酶 9                               | 3.3E-03 |
| 623 | U16996    |           | DUSP5     | 双特异性磷酸酶 5                                    | 3.3E-03 |
| 624 | AV705747  | NM_006276 | SFRS7     | 剪接因子, 富含精氨酸/丝氨酸的 7, 35kDa                    | 3.3E-03 |
| 625 | AF178984  |           | IER5      | 即早期反应蛋白 5                                    | 3.3E-03 |
| 626 | Z29093    |           | DDR1      | 网柄菌凝素结构域受体家族, 成员 1                           | 3.3E-03 |
| 627 | AB024536  |           | ISLR      | 含有富含亮氨酸的重复序列的免疫球蛋白超家族                        | 3.3E-03 |
| 628 | BF791601  |           | EMP2      | 上皮膜蛋白 2                                      | 3.3E-03 |
| 629 | AF061737  |           | SPC18     | 信号肽复合物(18kD)                                 | 3.3E-03 |
| 630 | AB002386  |           | EZH1      | Zeste 增强子同系物 1 (果蝇)                          | 3.5E-03 |
| 631 | AA634090  |           | EST       | 智人, 类似于核内不均一核糖核蛋白 A1, 克隆 IMAGE:2900557, mRNA | 3.5E-03 |
| 632 | AK023674  |           | FLJ13612  | 可能是神经元表达的钙结合蛋白的直向同源物                         | 3.6E-03 |

|        |           |           |           |  |         |
|--------|-----------|-----------|-----------|--|---------|
| 633    | D13626    |           | GPR105    | 与 G 蛋白-偶联的受体 105                                       | 3.7E-03 |
| 634    | AK026849  | XM_371844 | TSPYL     | TSPY-样   | 3.8E-03 |
| 635    | Y18643    |           | METTL1    | 甲基转移酶-样 1  | 3.9E-03 |
| 636    | AF176699  |           | FBXL4     | F-盒和富含亮氨酸重复序列的蛋白质 4                                    | 3.9E-03 |
| 637    | NM_003977 |           | AIP       | 芳烃受体相互作用蛋白   | 3.9E-03 |
| 638    | AK000498  |           | HARS      | 组氨酰-tRNA 合成酶   | 4.0E-03 |
| 639    | U05237    | NM_004459 | FALZ      | 胎儿 Alzheimer 抗原  | 4.0E-03 |
| 640    | BF696304  | NM_032832 | FLJ14735  | 假拟蛋白 FLJ14735  | 4.0E-03 |
| 641    | X14420    |           | COL3A1    | 胶原, III 型, $\alpha 1$ (Ehlers-Danlos 综合征 IV 型, 常染色体显性) | 4.1E-03 |
| 642    | BE796098  |           | NDUFS8    | NADH 脱氢酶(泛醌) Fe-S 蛋白 8, 23kDa (NADH-辅酶 Q 还原酶)          | 4.3E-03 |
| [0342] | 643       | X60221    | ATP5F1    | ATP 合酶, H <sup>+</sup> 转运的, 线粒体 F0 复合物, 亚单位 b, 同种型 1   | 4.4E-03 |
| 644    | AA135341  | NM_021078 | GCN5L2    | GCN5 全面控制氨基酸合成 5-样 2 (酵母)                              | 4.6E-03 |
| 645    | AF009368  |           | CREB3     | CAMP 应答元件结合蛋白 3 (luman)                                | 4.7E-03 |
| 646    | BF970013  |           | SPC12     | 信号肽酶 12kDa   | 4.7E-03 |
| 647    | W45522    |           | ATPIF1    | ATPase 抑制因子 1  | 4.7E-03 |
| 648    | AI733356  | NM_006306 | EST       | 智人 cDNA FLJ31746 fis, 克隆 NT2RI2007334                  | 4.8E-03 |
| 649    | AW117927  |           | EIF3S9    | 真核翻译起始因子 3, 亚单位 9 eta, 116kDa                          | 4.8E-03 |
| 650    | AF275798  | NM_012073 | CCT5      | 含 TCP1 的陪伴蛋白, 亚单位 5 ( $\epsilon$ )                     | 5.0E-03 |
| 651    | AI937126  |           | WTAP      | Wilms 肿瘤 1-相关蛋白  | 5.0E-03 |
| 652    | AK024891  | NM_203463 | LOC253782 | 假拟蛋白 LOC253782   | 5.1E-03 |
| 653    | D13629    |           | KTN1      | 移动结合蛋白 1 (驱动蛋白受体)                                      | 5.2E-03 |
| 654    | AI682994  |           | AHCYL1    | S-腺苷高半胱氨酸水解酶-样 1                                       | 5.3E-03 |

[0343]

|     |           |           |              |  |         |
|-----|-----------|-----------|--------------|--|---------|
| 655 | BF980325  | NM_005742 | ATP6V1C2     | ATPase, H <sup>+</sup> 转运的, 溶酶体的 42kDa, V1 亚单位 C 同种型 2 | 5.3E-03 |
| 656 | A1378996  | NM_005381 | NCL          | 核仁蛋白   | 5.3E-03 |
| 657 | D88153    |           | HYA22        | HYA22 蛋白   | 5.3E-03 |
| 658 | S67310    |           | BF           | B-因子, 备解素  | 5.4E-03 |
| 659 | AW438585  |           | EST          | 智人, 克隆 IMAGE:5273745, mRNA                             | 5.4E-03 |
| 660 | M12267    |           | OAT          | 鸟氨酸转氨酶(脉络膜视网膜回旋形萎缩)                                    | 5.5E-03 |
| 661 | AB001636  |           | DDX15        | DEAD/H (Asp-Glu-Ala-Asp/His) 盒多肽 15                    | 5.7E-03 |
| 662 | D13315    |           | GLO1         | 乙二醛酶 I   | 5.9E-03 |
| 663 | AF244931  |           | WDR10        | WD 重复结构域 10  | 5.9E-03 |
| 664 | AL050094  |           | IDH3B        | 异柠檬酸脱氢酶 3 (NAD <sup>+</sup> ) β                        | 6.0E-03 |
| 665 | AK022881  |           | KIAA1272     | KIAA1272 蛋白  | 6.0E-03 |
| 666 | AI720096  |           | RPL29        | 核糖体蛋白 L29  | 6.1E-03 |
| 667 | Y12781    |           | TBL1X        | 转导蛋白(β)-样 1X-连锁的                                       | 6.2E-03 |
| 668 | AI014538  | NM_138384 | LOC92170     | 假拟蛋白 BC004409  | 6.2E-03 |
| 669 | NM_020987 |           | ANK3         | 锚蛋白 3, 郎飞氏结(锚蛋白 G)                                     | 6.3E-03 |
| 670 | NM_004387 |           | NKX2-5       | NK2 转录因子相关的, 基因座 5 (果蝇)                                | 6.3E-03 |
| 671 | J03817    |           | GSTM1        | 谷胱苷肽 S-转移酶 M1  | 6.3E-03 |
| 672 | BF435769  |           | EST          | EST, 与假拟蛋白 FLJ20378 [智人] [H.sapiens] 有点类似              | 6.5E-03 |
| 673 | AL390147  |           | DKFZp547D065 | 假拟蛋白 DKFZp547D065                                      | 6.5E-03 |
| 674 | AA961412  | NM_003333 | UBA52        | 遍在蛋白 A-52 残基核糖体蛋白融合产物 1                                | 6.6E-03 |
| 675 | NM_002702 |           | POU6F1       | POU 结构域, 6 类, 转录因子 1                                   | 6.6E-03 |
| 676 | M58050    |           | MCP          | 膜辅因子蛋白(CD46, 滋养层-淋巴细胞交叉-反应性抗原)                         | 6.6E-03 |
| 677 | NM_00129  |           | CLNS1A       | 氯化物通道, 核苷酸-敏感的, 1A                                     | 6.7E-03 |

[0344]

|     |           |           |          |  |         |
|-----|-----------|-----------|----------|--|---------|
|     | 3         |           |          |  |         |
| 678 | BF213049  |           | COX7A2   | 细胞色素 c 氧化酶亚单位 VIIa 多肽 2<br>(肝脏)          | 6.7E-03 |
| 679 | AF236056  |           | GOLPH2   | 高尔基磷蛋白 2                                 | 6.7E-03 |
| 680 | U79285    | NM_021079 | NMT1     | N-豆蔻酰转移酶 1                               | 6.8E-03 |
| 681 | AB027196  |           | RNF10    | 环指蛋白 10                                  | 6.9E-03 |
| 682 | AA036952  |           | FLJ30973 | 假拟蛋白 FLJ30973                            | 7.0E-03 |
| 683 | AW732157  | NM_052963 | TOP1MT   | 线粒体拓扑异构酶 I                               | 7.1E-03 |
| 684 | AL049319  | NM_032804 | FLJ14547 | 假拟蛋白 FLJ14547                            | 7.3E-03 |
| 685 | BE613161  |           | EST      | 智人 cDNA FLJ37042 fis, 克隆<br>BRACE2011947 | 7.3E-03 |
| 686 | U28749    |           | HMGA2    | 高迁移率组 AT-hook 2                          | 7.3E-03 |
| 687 | BF793677  |           | MGC49942 | 假拟蛋白 MGC49942                            | 7.4E-03 |
| 688 | BG032216  | NM_017746 | FLJ20287 | 假拟蛋白 FLJ20287                            | 7.4E-03 |
| 689 | AL449244  |           | PP2447   | 假拟蛋白 PP2447                              | 7.5E-03 |
| 690 | AK024103  |           | EST      | 智人 cDNA FLJ14041 fis, 克隆<br>HEMBA1005780 | 7.5E-03 |
| 691 | U17838    |           | PRDM2    | 含 PR 结构域的 2, 具有 ZNF 结构域                  | 7.5E-03 |
| 692 | D86479    | NM_001129 | AEBP1    | AE 结合蛋白 1                                | 7.5E-03 |
| 693 | D50420    |           | NHP2L1   | NHP2 非-组蛋白染色体蛋白 2-样 1<br>(酿酒酵母)          | 7.5E-03 |
| 694 | D87258    |           | PRSS11   | 蛋白酶, 丝氨酸, 11 (IGF 结合的)                   | 7.5E-03 |
| 695 | BF434108  | NM_014187 | HSPC171  | HSPC171 蛋白                               | 7.6E-03 |
| 696 | NM_000705 |           | ATP4B    | ATPase, H+/K+交换的, $\beta$ 多肽             | 7.7E-03 |
| 697 | AF077599  |           | SBB103   | 假拟 SBBI03 蛋白                             | 7.7E-03 |
| 698 | NM_001530 |           | HIF1A    | 低氧-诱导因子 1, $\alpha$ 亚单位(碱性螺旋-环-螺旋转录因子)   | 7.8E-03 |
| 699 | AB023204  |           | EPB41L3  | 红细胞膜蛋白条带 4.1-样 3                         | 7.8E-03 |

[0345]

|     |           |           |            |                                  |         |
|-----|-----------|-----------|------------|----------------------------------|---------|
| 700 | AA253194  | NM_022121 | PIGPC1     | p53-诱导蛋白 PIGPC1                  | 7.9E-03 |
| 701 | BE502341  | NM_139177 | C17orf26   | 染色体 17 开放读码框 26                  | 7.9E-03 |
| 702 | AL050265  |           | TARDBP     | TAR DNA 结合蛋白                     | 8.0E-03 |
| 703 | AK001643  | NM_018215 | FLJ10781   | 假拟蛋白 FLJ10781                    | 8.3E-03 |
| 704 | BG179412  |           | COX7B      | 细胞色素 c 氧化酶亚单位 VIIb               | 8.6E-03 |
| 705 | X03212    |           | KRT7       | 角蛋白 7                            | 8.8E-03 |
| 706 | L07033    |           | HMGCL      | 3-羟甲基-3-甲基戊二酰-辅酶 A 裂合酶(羟甲基戊二酸尿症) | 9.0E-03 |
| 707 | M19383    |           | ANXA4      | 膜联蛋白 A4                          | 9.0E-03 |
| 708 | NM_001273 |           | CHD4       | 克罗莫结构域解旋酶 DNA 结合蛋白 4             | 9.1E-03 |
| 709 | NM_004461 |           | FARSL      | 苯丙氨酸-tRNA 合成酶-样                  | 9.1E-03 |
| 710 | AI192880  |           | CD44       | CD44 抗原(归巢功能和印度血型系)              | 9.1E-03 |
| 711 | AF038961  |           | MPDU1      | 甘露糖-P-长醇利用缺陷 1                   | 9.5E-03 |
| 712 | U67322    |           | C20orf18   | 染色体 20 开放读码框 18                  | 9.5E-03 |
| 713 | AA521017  |           | EST        | EST                              | 9.5E-03 |
| 714 | AA811043  | NM_003730 | RNASE6PL   | 核糖核酸酶 6 前体                       | 9.9E-03 |
| 715 | AA536113  |           | TMEPAI     | 跨膜, 前列腺雄激素的诱导 RNA                | 9.9E-03 |
| 716 | BF973104  |           | LOC2017255 | 假拟蛋白 LOC201725                   | 9.9E-03 |
| 717 | NM_000293 |           | PHKB       | 磷酸化酶激酶, $\beta$                  | 9.9E-03 |
| 718 | NM_000548 |           | TSC2       | 结节性硬化症 2                         | 1.0E-02 |

[0346] 表 11 在淋巴结 - 阳性和淋巴结 - 阴性肿瘤之间表达有所改变的基因列表

| [0347] | BRC 编号 | 登录号 | 标志 | 名称 | P-值 | + 或 - |
|--------|--------|-----|----|----|-----|-------|
|--------|--------|-----|----|----|-----|-------|

[0348]



|     |           |           |            |   |          |
|-----|-----------|-----------|------------|---|----------|
| 719 | BF686125  |           | UBA52      | 遍在蛋白 A-52 残基核糖体蛋白融合产物 1                       | 8.1E-09- |
| 720 | AA634090  |           |            | 智人, 类似于核内不均一核糖核蛋白 A1, 克隆 IMAGE:2900557, mRNA  | 1.4E-07- |
| 721 | L00692    |           | CEACAM3    | 癌胚抗原-相关的细胞粘附分子 3                              | 4.2E-07- |
| 722 | AW954403  | NM_004781 | VAMP3      | 小泡-相关膜蛋白 3 (小细胞小泡蛋白)                          | 2.2E-06+ |
| 723 | AA865619  |           | C21orf97   | 染色体 21 开放读码框 97                               | 2.6E-06- |
| 724 | W74502    | NM_032350 | MGC11257   | 假拟蛋白 MGC11257                                 | 2.4E-05+ |
| 725 | NM_002094 |           | GSPT1      | G1 至 S 相转换 1                                  | 2.7E-05+ |
| 726 | T55178    |           | KIAA1040   | KIAA1040 蛋白                                   | 3.2E-05- |
| 727 | L36983    |           | DNM2       | 发动蛋白 2  | 4.1E-05+ |
| 728 | Z21507    |           | EEF1D      | 真核翻译延伸因子 1 δ(鸟嘌呤核苷酸交换蛋白)                      | 5.2E-05- |
| 729 | AI581728  | NM_005507 | CFL1       | 肌动蛋白丝切蛋白 1 (非-肌肉)                             | 8.0E-05+ |
| 730 | NM_001293 |           | CLNS1A     | 氯化物通道, 核苷酸-敏感的, 1A                            | 9.0E-05+ |
| 731 | BF680847  |           | SEN2       | sentrin-特异性蛋白酶                                | 9.0E-05+ |
| 732 | AF100743  |           | NDUFS3     | NADH 脱氢酶(泛醌) Fe-S 蛋白 3, 30kDa (NADH-辅酶 Q 还原酶) | 9.8E-05+ |
| 733 | NM_004960 |           | FUS        | 融合蛋白, 得自 t(12;16)恶性脂肪肉瘤                       | 9.8E-05- |
| 734 | AK023975  | NM_015934 | NOP5/NOP58 | 核仁蛋白 NOP5/NOP58                               | 1.3E-04+ |
| 735 | AF083245  |           | PSMD13     | 蛋白酶体(前体, macropain) 26S 亚单位, 非-ATPase, 13     | 1.5E-04+ |
| 736 | AA129776  |           | SUOX       | 亚硫酸氧化酶  | 1.8E-04+ |
| 737 | U55766    | NM_007043 | HRB2       | HIV-1 rev 结合蛋白 2                              | 2.0E-04+ |
| 738 | BF526092  |           | LOC15446   | 假拟蛋白 BC003515                                 | 2.1E-04+ |

[0349]

|     |          |           |          |                               |  |  |  |  |          |  |  |
|-----|----------|-----------|----------|-------------------------------|--|--|--|--|----------|--|--|
|     |          |           | 7        |                               |  |  |  |  |          |  |  |
| 739 | BF677579 | XM_370754 | THTPA    | 硫酸素三磷酸酶                       |  |  |  |  | 2.3E-04+ |  |  |
| 740 | X98260   |           | ZRF1     | Zuotin 相关因子 1                 |  |  |  |  | 2.3E-04+ |  |  |
| 741 | BE440010 |           | LOC51255 | 假拟蛋白 LOC51255                 |  |  |  |  | 2.7E-04+ |  |  |
| 742 | AF007165 | NM_021008 | DEAF1    | 变形表皮自身调节因子 1 (果蝇)             |  |  |  |  | 2.7E-04+ |  |  |
| 743 | X78687   |           | NEU1     | 唾液酸酶 1 (溶酶体唾液酸酶)              |  |  |  |  | 3.0E-04+ |  |  |
| 744 | AW965200 |           |          | 智人, 克隆 IMAGE:5286019, mRNA    |  |  |  |  | 3.1E-04- |  |  |
| 745 | AK023240 |           | UGCGL1   | UDP-葡萄糖神经酰胺葡萄糖基转移酶-样 1        |  |  |  |  | 3.1E-04+ |  |  |
| 746 | M95712   |           | BRAF     | v-ras 鼠肉瘤病毒致癌基因同系物 B1         |  |  |  |  | 3.7E-04+ |  |  |
| 747 | L38995   | NM_003321 | TUFM     | Tu 翻译延伸因子, 线粒体的               |  |  |  |  | 3.9E-04+ |  |  |
| 748 | AW014268 |           | FLJ10726 | 假拟蛋白 FLJ10726                 |  |  |  |  | 4.2E-04+ |  |  |
| 749 | D49547   |           | DNAJB1   | DnaJ (Hsp40)同系物, 亚族 B, 成员 1   |  |  |  |  | 4.4E-04+ |  |  |
| 750 | BE466450 |           | AP4S1    | 衔接子-相关蛋白复合物 4, $\sigma 1$ 亚单位 |  |  |  |  | 4.5E-04+ |  |  |
| 751 | AB007944 |           | KIAA0475 | KIAA0475 基因产物                 |  |  |  |  | 4.9E-04- |  |  |
| 752 | AF034091 |           | MRPL40   | 线粒体核糖体蛋白 L40                  |  |  |  |  | 5.1E-04+ |  |  |

[0350] 表 12 组织临床信息

[0351]

| 编号        | 手术<br>年龄 | 经绝<br>状况 | T | N | M | 病程 | 组织类<br>型 | 淋巴细<br>胞浸润 | 血管侵<br>入 | ER | PgR |
|-----------|----------|----------|---|---|---|----|----------|------------|----------|----|-----|
| MMK010003 | 51       | pre      | 2 | 1 | 0 | 2  | a3       | 3          | 0        | +  | +   |
| MMK010004 | 47       | pre      | 2 | 1 | 0 | 2  | a1       | 0          | 0        | +  | +   |
| MMK010005 | 44       | pre      | 2 | 0 | 0 | 2  | a1       | 1          | 0        | +  | +   |
| MMK010013 | 45       | pre      | 2 | 1 | 0 | 2  | a1       | 1          | 0        | -  | -   |
| MMK010016 | 44       | pre      | 2 | 0 | 0 | 2  | a2       | 0          | 0        | -  | -   |
| MMK010025 | 46       | pre      | 2 | 0 | 0 | 2  | a1       | 0          | 0        | +  | +   |
| MMK010031 | 29       | pre      | 2 | 2 | 0 | 3  | a3       | 3          | 0        | -  | -   |

[0352]

|           |    |      |   |   |   |   |    |   |   |   |   |
|-----------|----|------|---|---|---|---|----|---|---|---|---|
| MMK010037 | 62 | post | 0 | 0 | 0 | 0 | Ia | 0 | 0 | + | + |
| MMK010042 | 47 | pre  | 2 | 1 | 0 | 2 | a3 | 1 | 2 | + | + |
| MMK010086 | 42 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |
| MMK010102 | 51 | pre  | 2 | 1 | 0 | 3 | a2 | 3 | 0 | + | + |
| MMK010110 | 39 | pre  | 2 | 0 | 0 | 2 | a1 | 2 | 0 | - | - |
| MMK010129 | 52 | pre  | 2 | 2 | 0 | 3 | a1 | 2 | 0 | - | - |
| MMK010135 | 41 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |
| MMK010138 | 38 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |
| MMK010145 | 51 | pre  | 2 | 1 | 0 | 2 | a3 | 0 | 0 | + | + |
| MMK010147 | 49 | pre  | 2 | 1 | 0 | 2 | a1 | 1 | 0 | + | + |
| MMK010149 | 35 | pre  | 2 | 0 | 0 | 2 | a3 | 1 | 0 | - | - |
| MMK010175 | 38 | pre  | 2 | 0 | 0 | 2 | a3 | 0 | 0 | + | + |
| MMK010178 | 51 | pre  | 0 | 0 | 0 | 0 | Ia | 0 | 0 | + | + |
| MMK010207 | 40 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |
| MMK010214 | 42 | pre  | 2 | 1 | 0 | 2 | a1 | 0 | 0 | - | - |
| MMK010247 | 48 | pre  | 2 | 1 | 0 | 2 | a2 | 3 | 0 | - | - |
| MMK010252 | 52 | pre  | 2 | 1 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010255 | 47 | pre  | 2 | 0 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010302 | 46 | pre  | 2 | 1 | 0 | 2 | a2 | 2 | 1 | - | - |
| MMK010304 | 48 | pre  | 2 | 1 | 0 | 2 | a3 | 1 | 0 | + | + |
| MMK010326 | 53 | post | 0 | 0 | 0 | 0 | Ia | 0 | 0 | - | - |
| MMK010327 | 43 | pre  | 2 | 1 | 0 | 2 | a1 | 1 | 1 | + | + |
| MMK010341 | 42 | pre  | 2 | 1 | 0 | 2 | a1 | 2 | 0 | + | + |
| MMK010370 | 46 | pre  | 2 | 1 | 0 | 2 | a3 | 2 | 0 | + | + |
| MMK010397 | 38 | pre  | 2 | 1 | 0 | 2 | a3 | 3 | 2 | + | + |
| MMK010411 | 46 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |
| MMK010431 | 50 | pre  | 2 | 0 | 0 | 2 | a3 | 0 | 0 | - | - |
| MMK010435 | 49 | pre  | 2 | 1 | 0 | 2 | a3 | 0 | 0 | + | + |
| MMK010453 | 49 | pre  | 2 | 1 | 0 | 2 | a3 | 3 | 0 | + | + |
| MMK010471 | 42 | pre  | 2 | 1 | 0 | 2 | a1 | 3 | 0 | - | - |
| MMK010473 | 40 | pre  | 2 | 1 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010478 | 38 | pre  | 2 | 2 | 0 | 3 | a2 | 0 | 0 | + | + |
| MMK010491 | 46 | pre  | 2 | 0 | 0 | 2 | a3 | 1 | 0 | + | + |
| MMK010497 | 44 | pre  | 0 | 0 | 0 | 0 | Ia | 0 | 0 | - | + |
| MMK010500 | 45 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |

[0353]

|           |    |      |   |   |   |   |    |   |   |   |   |
|-----------|----|------|---|---|---|---|----|---|---|---|---|
| MMK010502 | 51 | pre  | 2 | 0 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010508 | 51 | pre  | 2 | 1 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010521 | 21 | pre  | 2 | 0 | 0 | 2 | a1 | 1 | 1 | - | - |
| MMK010552 | 49 | pre  | 2 | 0 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010554 | 51 | pre  | 2 | 0 | 0 | 2 | a3 | 2 | 0 | + | + |
| MMK010571 | 45 | pre  | 2 | 1 | 1 | 4 | a3 | 3 | 0 | + | + |
| MMK010591 | 40 | pre  | 0 | 0 | 0 | 0 | la | 0 | 0 | - | + |
| MMK010613 | 37 | pre  | 0 | 0 | 0 | 0 | la | 0 | 0 | - | + |
| MMK010623 | 39 | pre  | 2 | 1 | 0 | 2 | a1 | 3 | 0 | + | + |
| MMK010624 | 39 | pre  | 2 | 1 | 0 | 2 | a1 | 3 | 0 | + | + |
| MMK010626 | 48 | pre  | 2 | 0 | 0 | 2 | a1 | 1 | 1 | - | - |
| MMK010631 | 41 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |
| MMK010640 | 35 | pre  | 0 | 0 | 0 | 0 | la | 0 | 0 | + | + |
| MMK010644 | 47 | pre  | 2 | 2 | 0 | 2 | a3 | 3 | 0 | + | + |
| MMK010646 | 37 | pre  | 2 | 1 | 0 | 2 | a3 | 1 | 0 | + | + |
| MMK010660 | 46 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | - | - |
| MMK010671 | 45 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | - | - |
| MMK010679 | 68 | post | 0 | 0 | 0 | 0 | la | 0 | 0 | + | + |
| MMK010680 | 58 | post | 0 | 0 | 0 | 0 | la | 0 | 0 | - | + |
| MMK010709 | 33 | pre  | 2 | 0 | 0 | 2 | a3 | 0 | 2 | - | - |
| MMK010711 | 51 | pre  | 0 | 0 | 0 | 0 | la | 0 | 0 | - | + |
| MMK010724 | 40 | pre  | 2 | 1 | 0 | 2 | a3 | 3 | 2 | + | + |
| MMK010744 | 41 | pre  | 0 | 0 | 0 | 0 | la | 0 | 0 | + | + |
| MMK010758 | 40 | pre  | 2 | 1 | 0 | 2 | a1 | 0 | 1 | + | + |
| MMK010760 | 42 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |
| MMK010762 | 50 | pre  | 2 | 1 | 0 | 2 | a3 | 3 | 1 | + | + |
| MMK010769 | 33 | pre  | 2 | 0 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010772 | 45 | pre  | 2 | 1 | 0 | 2 | a3 | 2 | 0 | - | - |
| MMK010779 | 46 | pre  | 2 | 1 | 0 | 2 | a2 | 0 | 1 | - | - |
| MMK010780 | 31 | pre  | 2 | 0 | 0 | 2 | a2 | 0 | 0 | - | - |
| MMK010781 | 44 | pre  | 2 | 0 | 0 | 2 | a3 | 0 | 2 | + | + |
| MMK010794 | 52 | pre  | 2 | 1 | 0 | 2 | a3 | 2 | 1 | + | + |
| MMK010818 | 51 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 2 | + | + |
| MMK010835 | 42 | pre  | 0 | 0 | 0 | 0 | la | 0 | 0 | + | + |
| MMK010846 | 47 | pre  | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |

[0354]

|           |    |     |   |   |   |   |    |   |   |   |   |
|-----------|----|-----|---|---|---|---|----|---|---|---|---|
| MMK010858 | 42 | pre | 2 | 1 | 0 | 2 | a3 | 2 | 3 | + | + |
| MMK010864 | 52 | pre | 2 | 1 | 0 | 2 | a1 | 0 | 1 | - | - |
| MMK010869 | 45 | pre | 2 | 0 | 0 | 2 | a1 | 0 | 1 | - | - |
| MMK010903 | 47 | pre | 2 | 0 | 0 | 2 | a1 | 0 | 0 | + | + |

[0355] 表 13

[0356]

|       |  |                  |
|-------|--|------------------|
| Si1-F | 5' -CACCGAACGATATAAAGCCAGCCTTCAAGAGAGGC<br>TGGCTTTATATCGTTC-3' | SEQ ID<br>NO. 23 |
| Si1-R | 5' -AAAAGAACGATATAAAGCCAGCCTCTCTGAAGG<br>CTGGCTTTATATCGTTC-3'  | SEQ ID<br>NO. 24 |
| Si1-靶 | 5' -GAACGATATAAAGCCAGCC-3'                                     | SEQ ID<br>NO. 25 |
| Si3-F | 5' -CACCTGGATGAATCATACCAGATTCAAGAGATCT<br>GGTATGATTATCCAG-3'   | SEQ ID<br>NO. 26 |
| Si3-R | 5' -AAAAGTGGATGAATCATACCAGATCTCTGAATCT<br>GGTATGATTATCCAG-3'   | SEQ ID<br>NO. 27 |
| Si3-靶 | 5' -CTGGATGAATCATACCAGA-3'                                     | SEQ ID<br>NO. 28 |
| Si4-F | 5' -CACCGTGTGGCTTGCCTAAATAATTCAAGAGATTA<br>TTACGCAAGCCACAC-3'  | SEQ ID<br>NO. 29 |
| Si4-R | 5' -AAAAGTGTGGCTTGCCTAAATAATCTCTGAATTATT<br>ACGCAAGCCACAC-3'   | SEQ ID<br>NO. 30 |
| Si4-靶 | 5' -GTGTGGCTTGCCTAAATAA-3'                                     | SEQ ID<br>NO. 31 |

[0357] 工业实用性

[0358] 通过联合使用激光 - 捕获解剖和全基因组 cDNA 微阵列获得的本文所述乳腺癌基因表达分析鉴定出特异性的可用作癌症预防和治疗之靶标的基因。根据这些差异表达基因亚群的表达, 本发明提供了可用于鉴定或检测乳腺癌的分子诊断标记。

[0359] 本文所述的方法也可用于鉴定其它可用于预防、诊断和治疗乳腺癌的分子靶标。本文报道的数据会使人们加深对乳腺癌的了解, 有利于开发新的诊断策略, 并能提供鉴定治疗药物和预防剂的分子靶标所用的线索。所述信息能使人们更加深刻地理解乳腺肿瘤的发生, 并能提供指示剂, 用于开发新的诊断、治疗和最终预防乳腺癌的策略。

[0360] 本文提及的所有专利、专利申请和出版物皆全文列入本文作为参考。

[0361] 另外, 尽管根据本发明的具体实施方案详细描述了本发明, 但应理解上述说明书实际上是例举和解释性的, 目的仅是为了阐明本发明及其优选实施方案。通过常规的实验, 本发明的多种变化和修饰对于本领域技术人员而言是显而易见的, 并不背离本发明的精神和范围。因此, 本发明不受上述说明书的限制, 而受下述权利要求书及其等同物的限制。

## 序列表

<110> 肿瘤疗法科学股份有限公司 (ONCOTHERAPY SCIENCE, INC.)  
 东京大学 (THE UNIVERSITY OF TOKYO)

<120> 诊断乳腺癌的方法

<130> ONC-A0306P2

<150> US 60/505, 571

<151> 2003-09-24

<160> 31

<170> PatentIn version 3.1

<210> 1

<211> 23

<212> DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400> 1

ctgttctggc ttcgttatgt tct

23

<210> 2

<211> 23

<212> DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400> 2

agaaaatacg gtcctcttgt tgc

23

<210> 3

<211>22

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>3

cactgtaatg cacgacattt ga

22

<210>4

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>4

gttacagctt agcacaaggc atc

23

<210>5

<211>22

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>5

acctctgagt ttgatttccc aa

22

<210>6

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列



<400>6  
cgaggcttgt aacaatctac tgg 23

<210>7  
<211>23  
<212>DNA  
<213> 人工的

<220>  
<223> 人工合成的引物序列

<400>7  
gaaactgtac gggggttaaa gag 23

<210>8  
<211>23  
<212>DNA  
<213> 人工的

<220>  
<223> 人工合成的引物序列

<400>8  
catcaatgtg gtgagtgaca tct 23

<210>9  
<211>23  
<212>DNA  
<213> 人工的

<220>  
<223> 人工合成的引物序列

<400>9  
aagcccttgg aacagaacat act 23

<210>10  
<211>23  
<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>10

cagtaaacgt ggttctcaca ttg

23

<210>11

<211>20

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>11

cgaccacttt gtcaagctca

20

<210>12

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>12

ggttgagcac agggctacttt att

23

<210>13

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>13

agaccctaaa gatcgtcctt ctg

23

<210>14

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的引物序列

<400>14

gtgttttaag tcagcatgag cag

23

<210>15

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

<400>15

tcccgcgcgc tttgtaggat tcgttcaaga gacgaatcct acaaagcgcg c

51

<210>16

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

<400>16

aaaagcgcgc tttgtaggat tcgtctcttg aacgaatcct acaaagcgcg c

51

<210>17

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

<400>17

tccccgtacg cggataactt cgattcaaga gatcgaagta ttccgcgtac g

51

<210>18

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

<400>18

aaaacgtacg cggataactt cgatctcttg aatcgaagta ttccgcgtac g

51

<210>19

<211>21

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物序列

<400>19

atggaaatcc catcaccatc t

21

<210>20

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物序列

<400>20

ggttgagcac agggactttt att

23

<210>21

<211>20

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物序列

<400>21

gccttcatca tccaaacatt

20

<210>22

<211>20

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物序列

<400>22

ggcaaatatg tctgccttgt

20

<210>23

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

<400>23

caccgaacga tataaagcca gccttcaaga gaggctggct ttatategtt c

51

<210>24

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

|   |    |
|---|----|
| <400>24   |    |
| aaaagaacga tataaagcca gcctctcttg aagctggct ttatategtt c | 51 |
| <210>25   |    |
| <211>19   |    |
| <212>DNA  |    |
| <213> 人工的   |    |
| <220>   |    |
| <223> 人工合成的 siRNA 靶序列                                   |    |
| <400>25   |    |
| gaacgatata aagccagcc                                    | 19 |
| <210>26   |    |
| <211>51   |    |
| <212>DNA  |    |
| <213> 人工的   |    |
| <220>   |    |
| <223> 人工合成的 siRNA 寡核苷酸序列                                |    |
| <400>26   |    |
| cacctggat gaatcatacc agattcaaga gatctggat gattcatcca g  | 51 |
| <210>27   |    |
| <211>51   |    |
| <212>DNA  |    |
| <213> 人工的   |    |
| <220>   |    |
| <223> 人工合成的 siRNA 寡核苷酸序列                                |    |
| <400>27   |    |
| aaaactggat gaatcatacc agatctcttg aatctggat gattcatcca g | 51 |
| <210>28   |    |
| <211>19   |    |

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 靶序列

<400>28

ctggatgaat cataccaga

19

<210>29

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

<400>29

caccgtgtgg cttgcgtaaa taattcaaga gattatttac gcaagccaca c

51

<210>30

<211>51

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 寡核苷酸序列

<400>30

aaaagtgtgg cttgcgtaaa taatctcttg aattatttac gcaagccaca c

51

<210>31

<211>19

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 siRNA 靶序列

<400>31

gtgtggcttg cgtaaataa

19



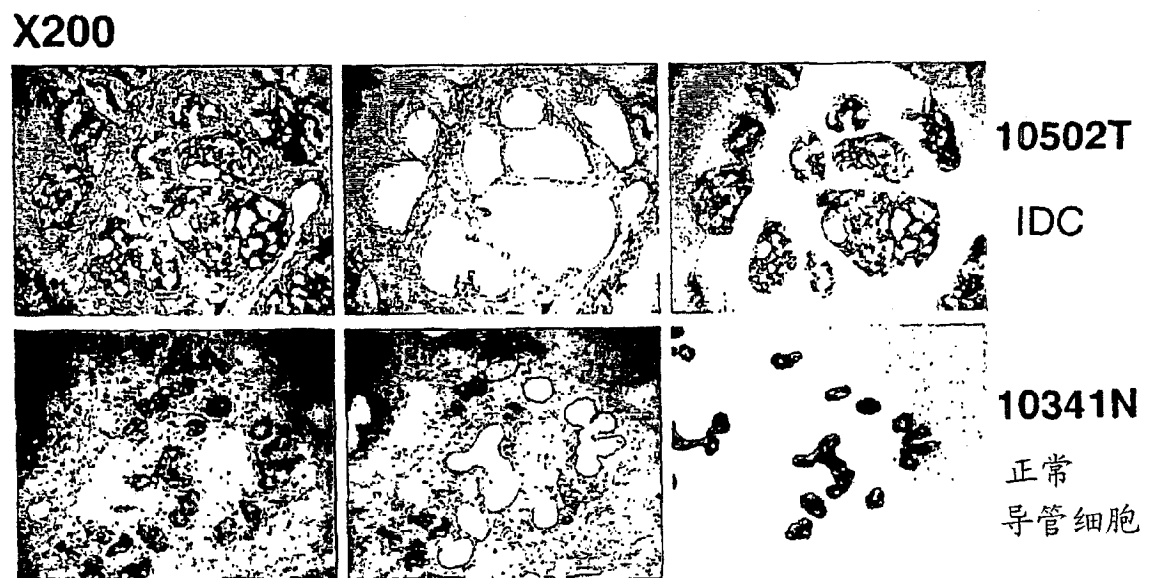
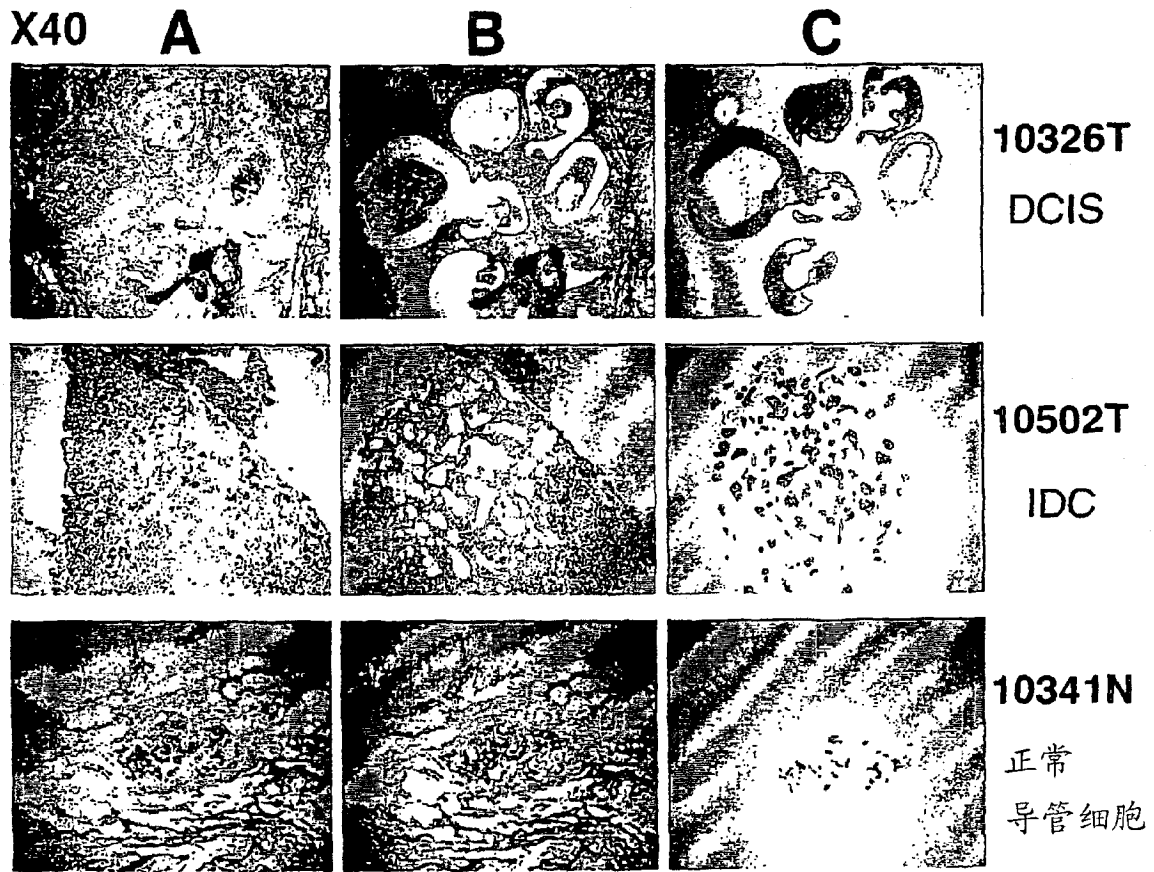


图 1

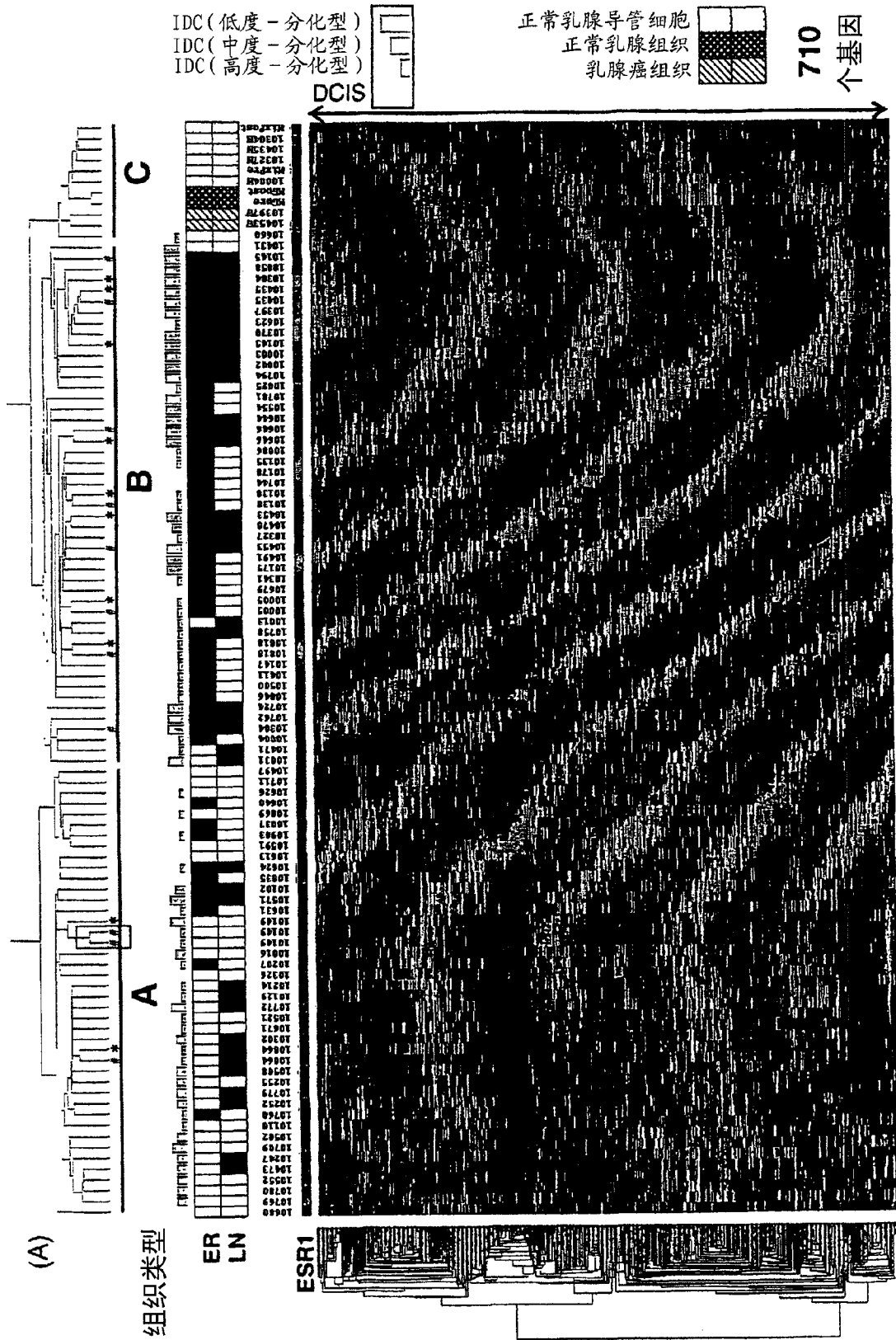


图 2a

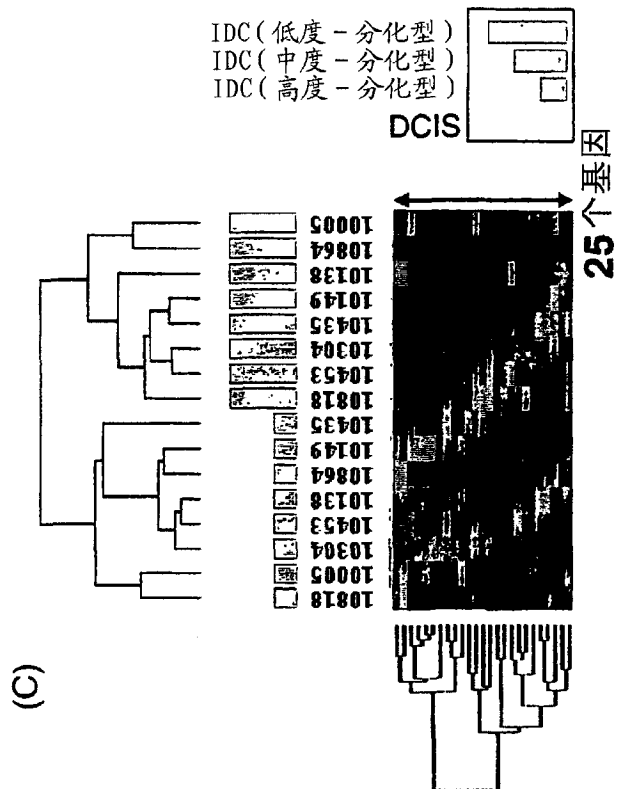


图 2c

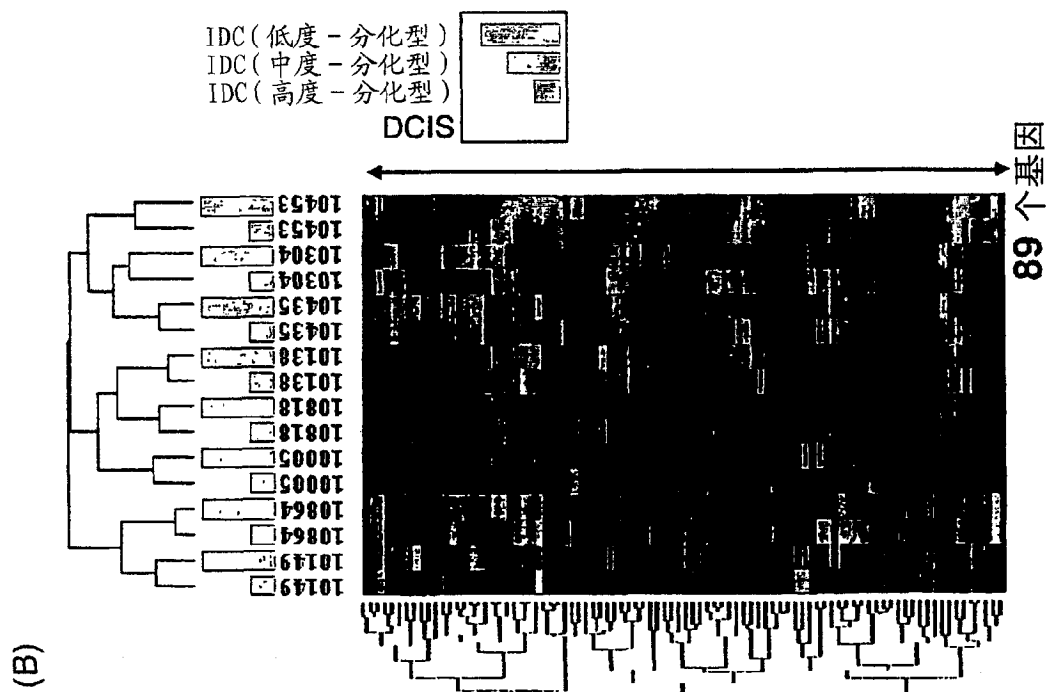


图 2b

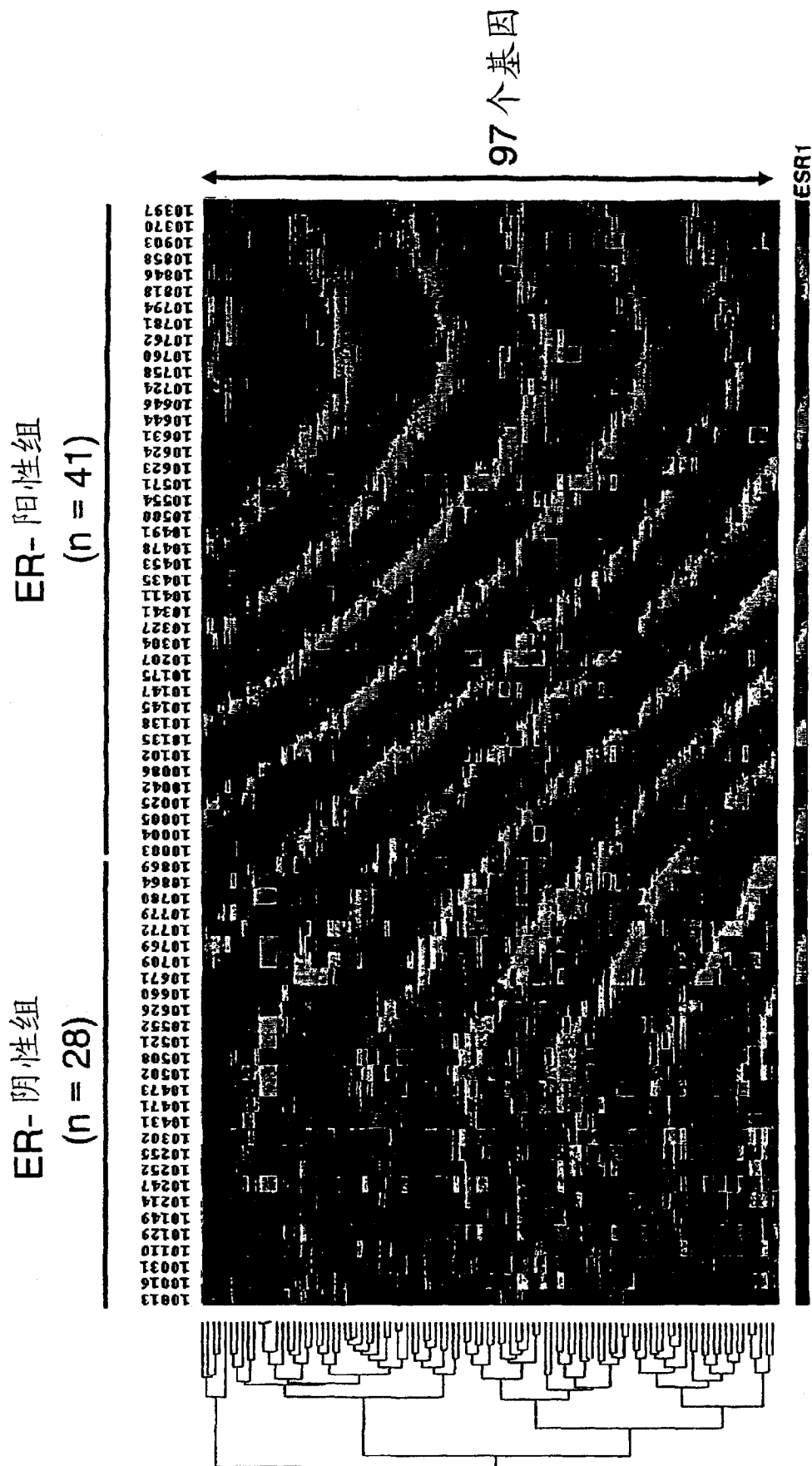
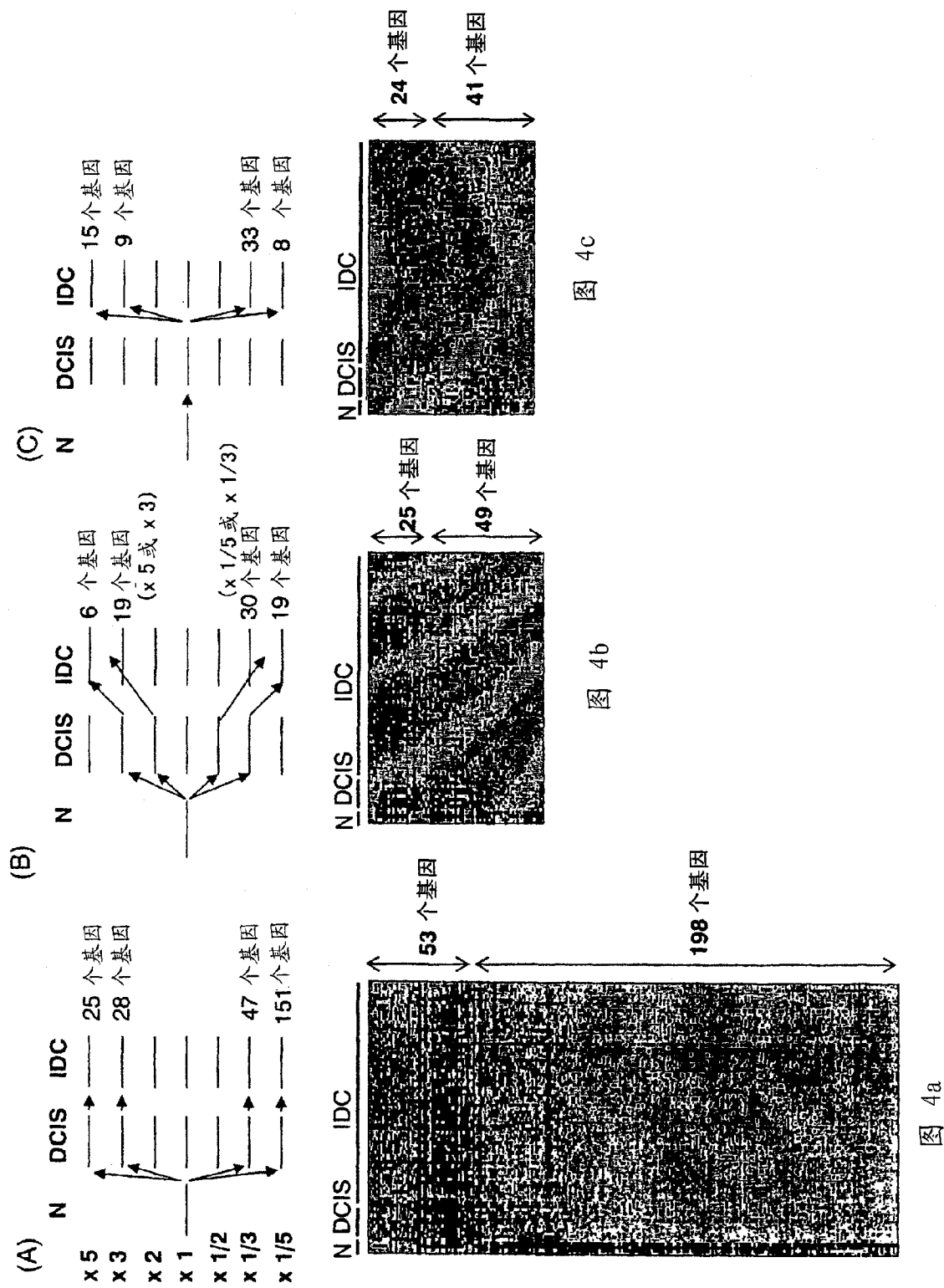


图 3



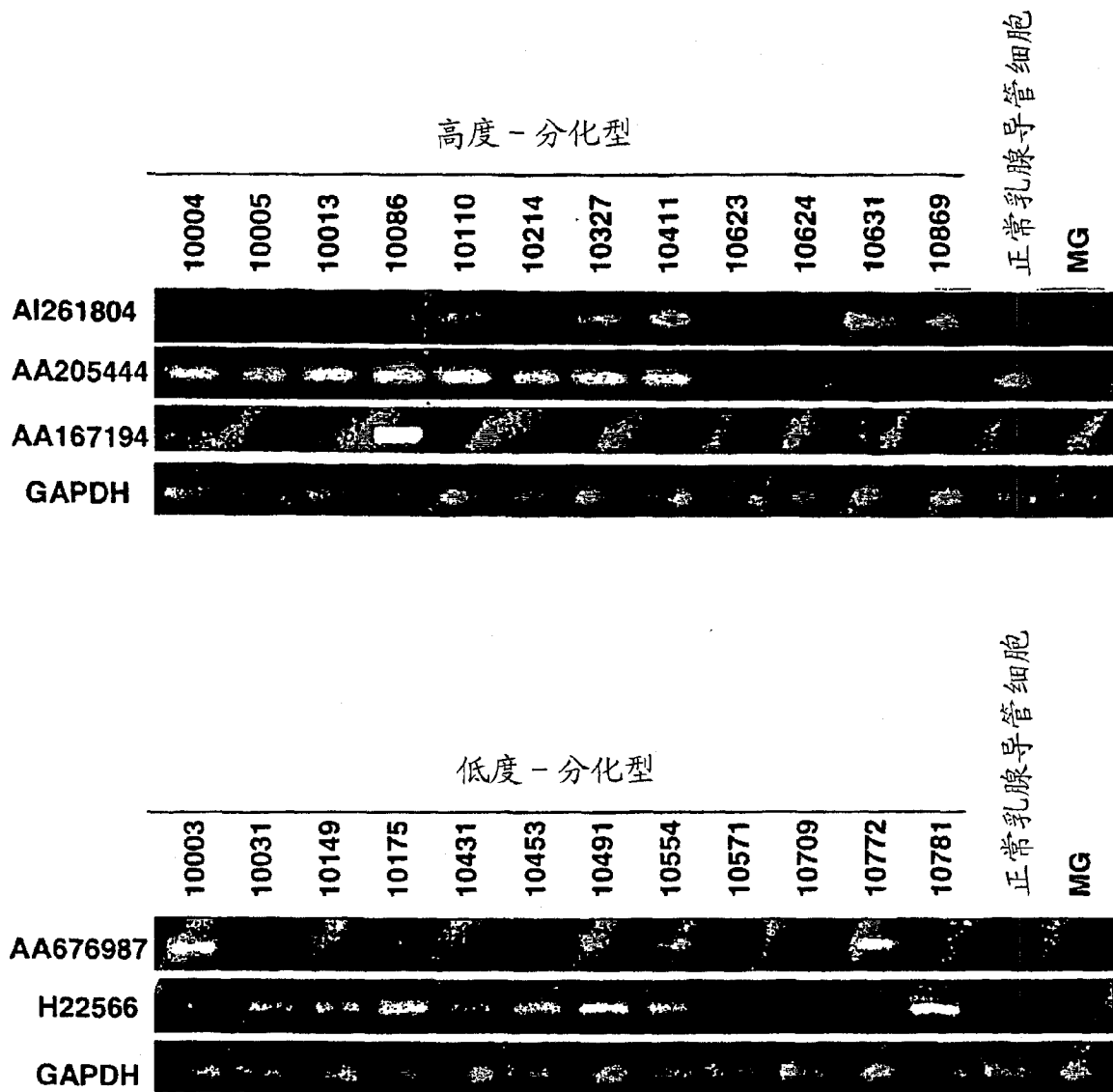


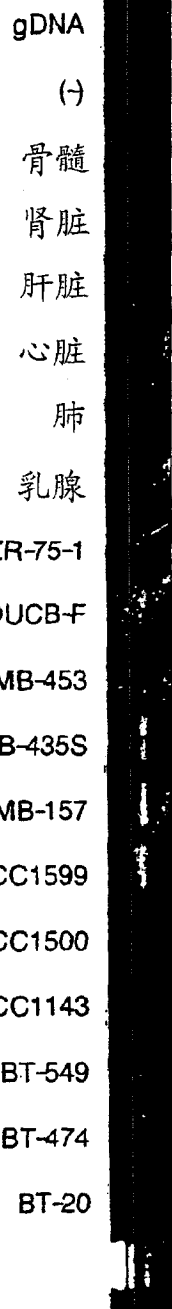
图 5



图 6a



图 6b



a

b

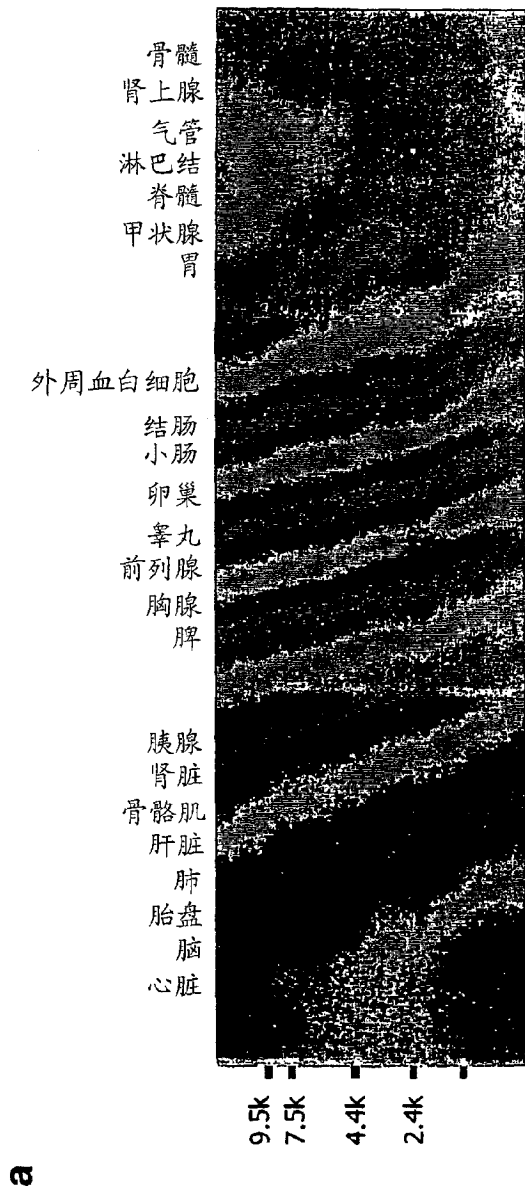


图 7a

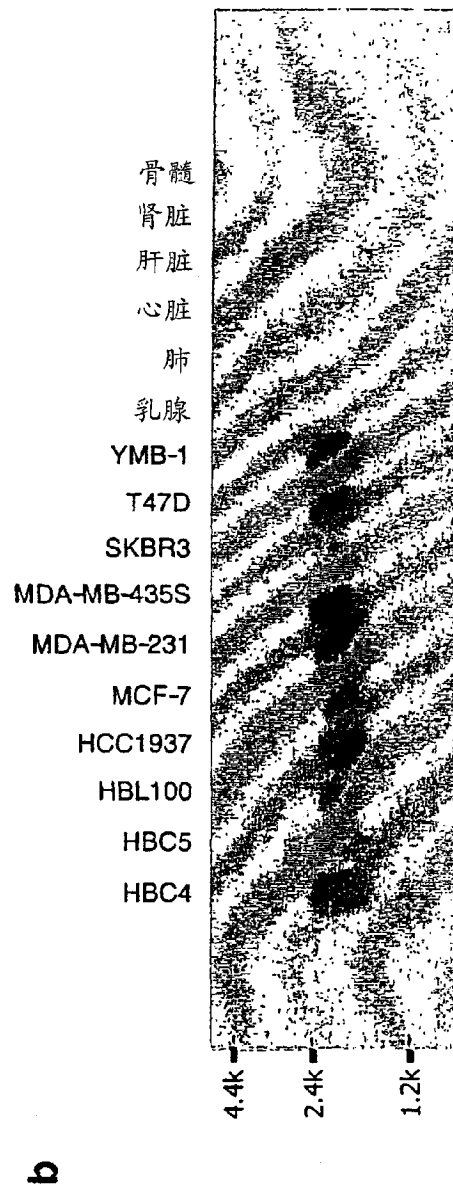


图 7b



a

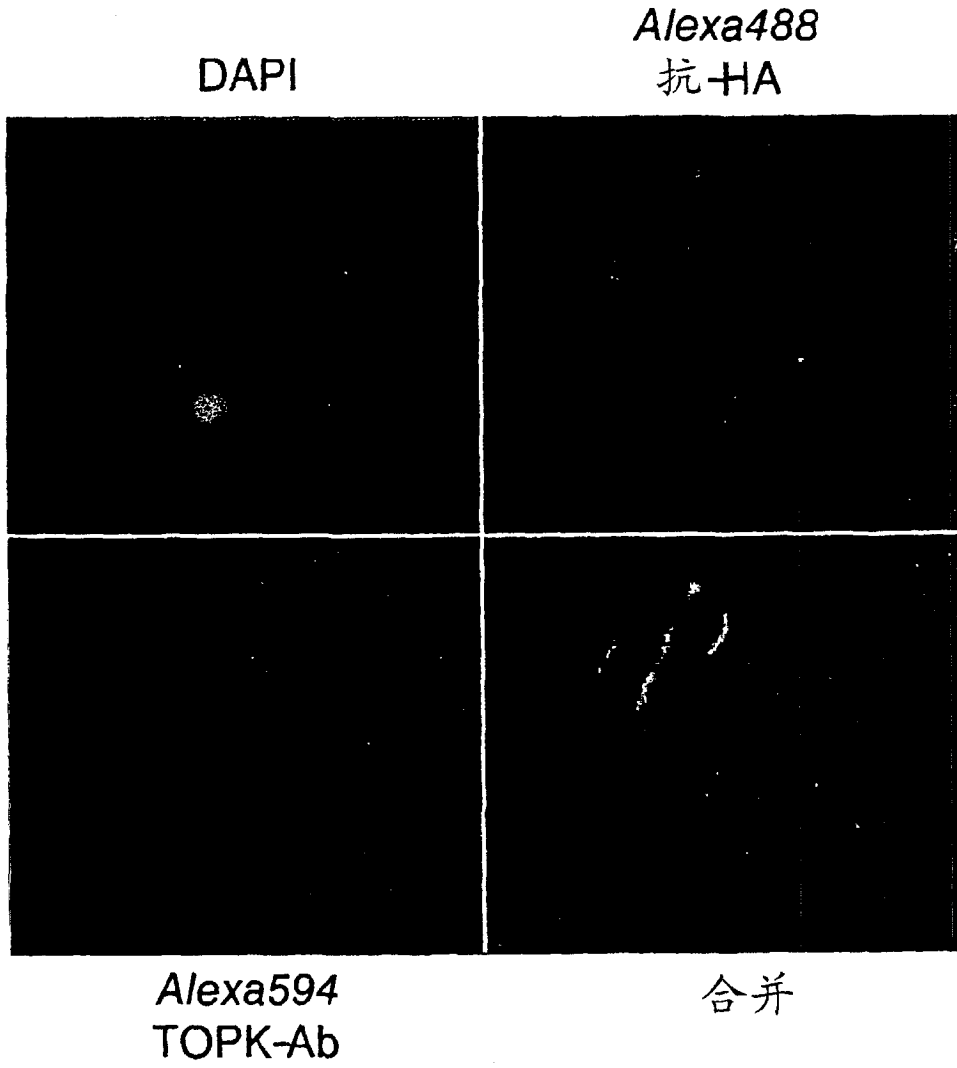


图 8a

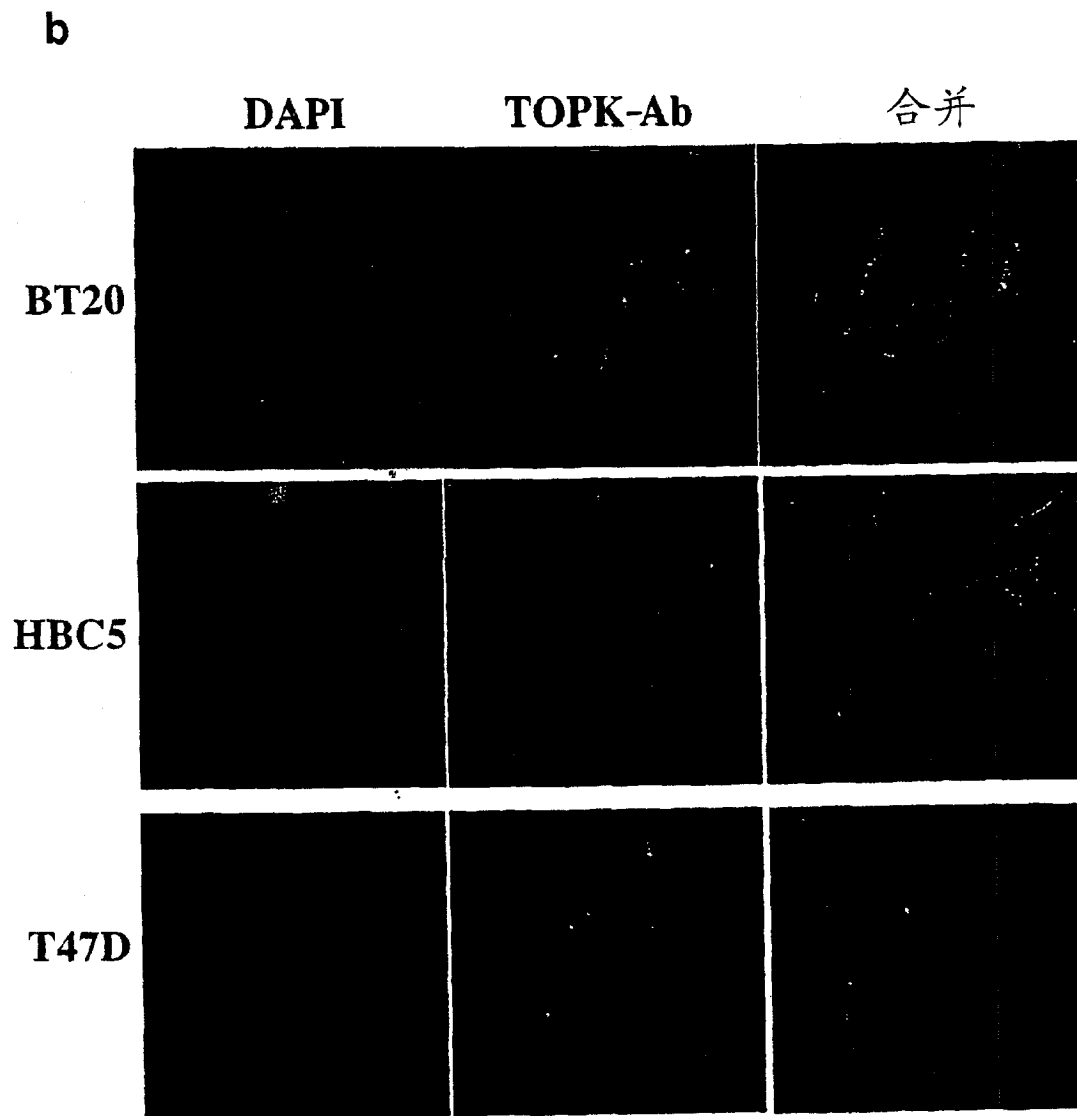


图 8b

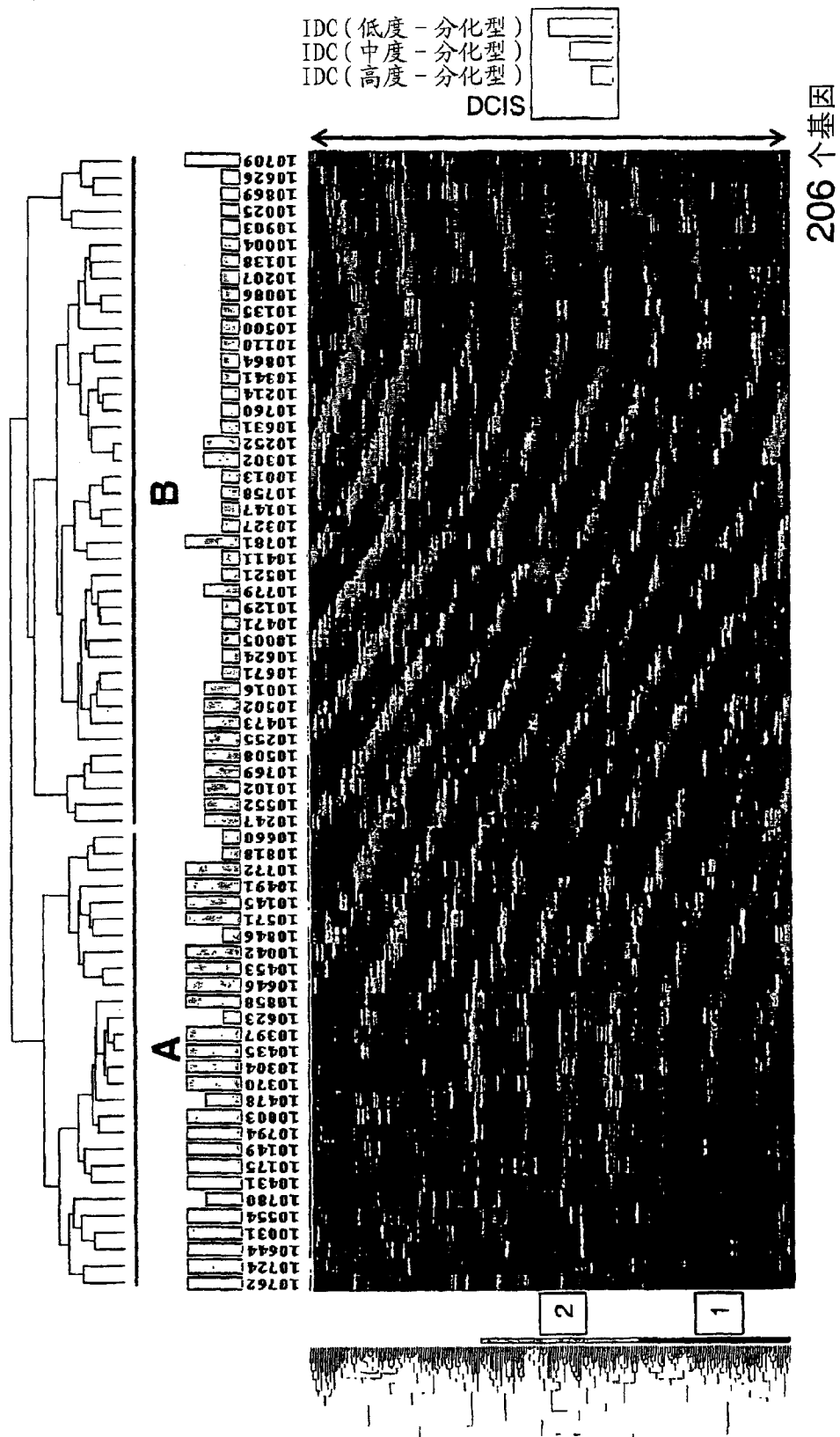
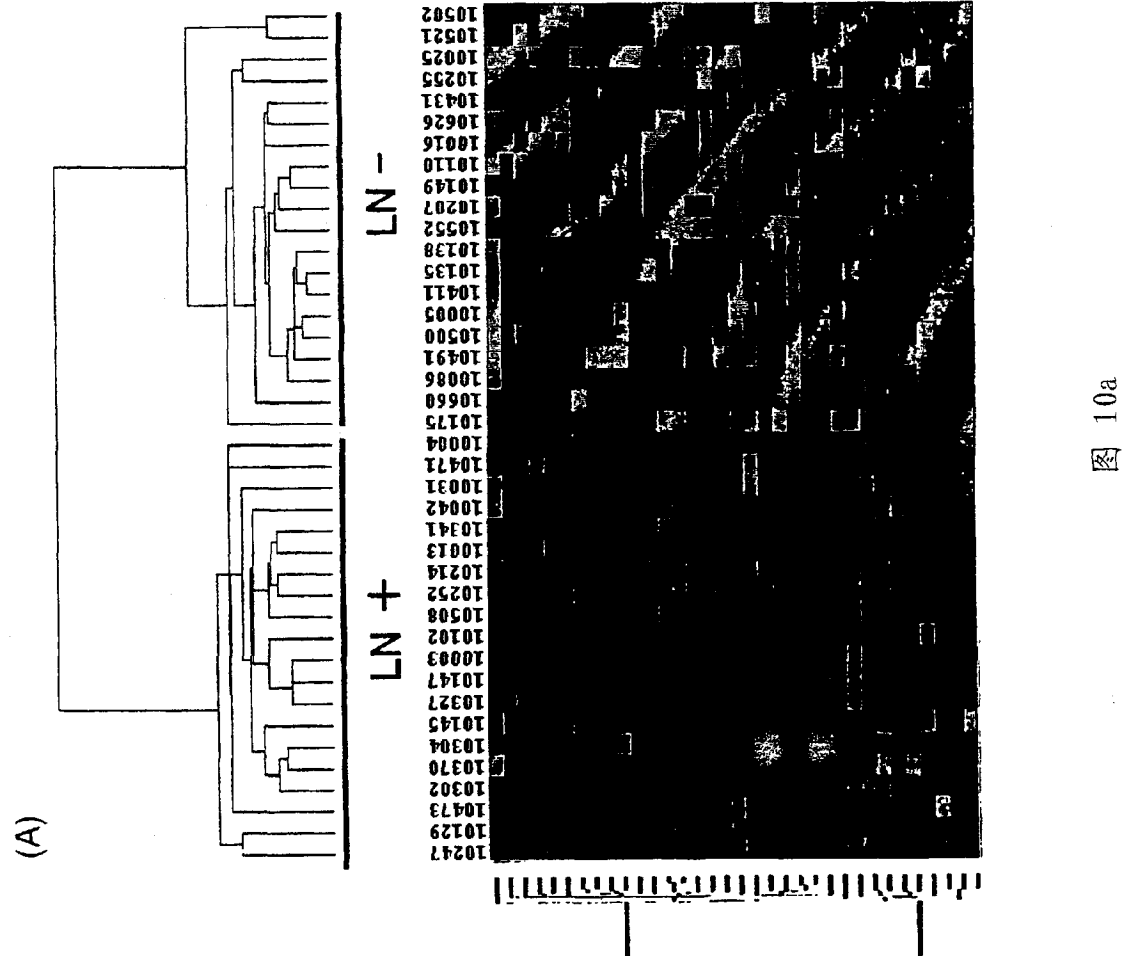
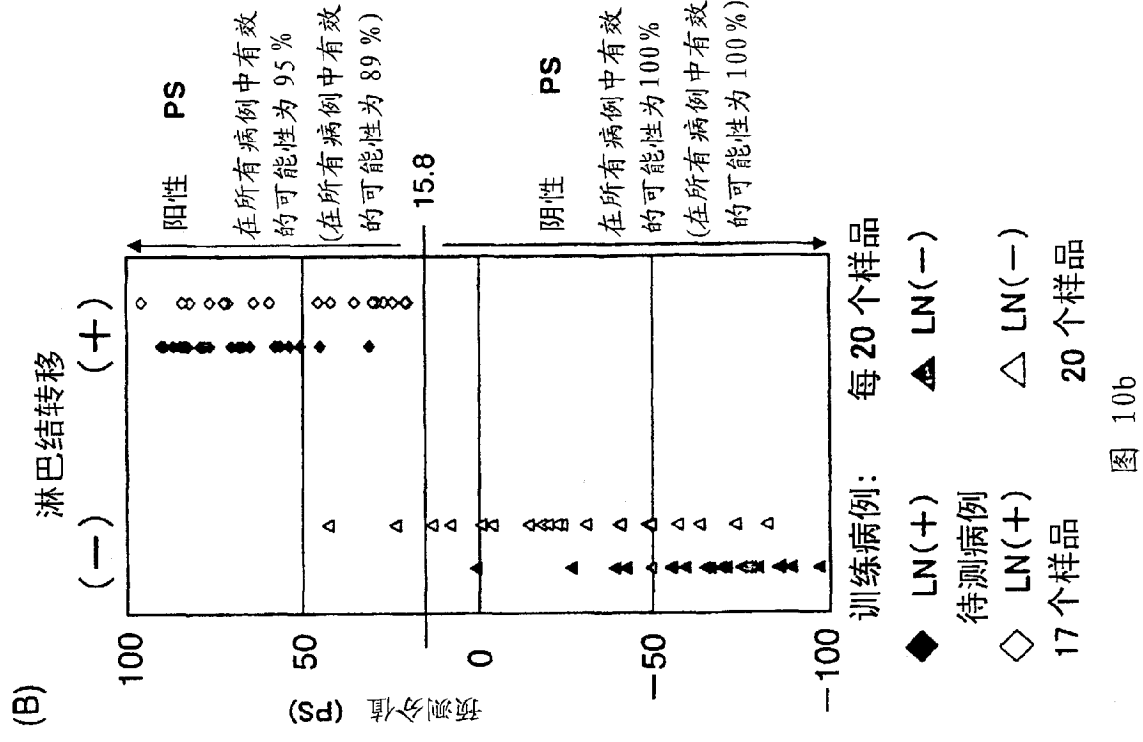


图 9



(C)

|      |    |       |     |    |
|------|----|-------|-----|----|
| 所有病例 |    | 淋巴结状态 |     | 总计 |
|      |    | (+)   | (-) |    |
| PS   | 阳性 | 37    | 2   | 39 |
|      | 阴性 | 0     | 38  | 38 |
| 总计   |    | 37    | 40  | 77 |

P<0.001  
(检测独立性的 $\chi^2$ 检验)

|      |    |       |     |    |
|------|----|-------|-----|----|
| 待测病例 |    | 淋巴结状态 |     | 总计 |
|      |    | (+)   | (-) |    |
| PS   | 阳性 | 17    | 2   | 19 |
|      | 阴性 | 0     | 18  | 18 |
| 总计   |    | 17    | 20  | 37 |

P<0.001  
(检测独立性的 $\chi^2$ 检验)

图 10c