



(12) 发明专利

(10) 授权公告号 CN 1898563 B

(45) 授权公告日 2011. 11. 23

(21) 申请号 200480034709. 8  
 (22) 申请日 2004. 08. 10  
 (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/JP2004/011741 2004. 08. 10  
 (87) PCT申请的公布数据  
 W02005/029067 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

(56) 对比文件  
 US 6365358 B1, 2002. 04. 02, 全文.  
 CN 1211279 A, 1999. 03. 17, 全文.  
 WO 02/04514 A, 2002. 01. 17, 全文.  
 ORLOWSKI ROBERT Z, et al. The role of the ubiquitination-proteasome pathway in breast cancer: applying drugs that affect the ubiquitin-proteasome pathway to the therapy of breast cancer. 《BREAST CANCER RESEARCH》. 2002, 第 5 卷 (第 1 期), 1-7.  
 LIPKOWITZ S, et al. The role of the ubiquitination-proteasome pathway in breast cancer. Ubiquitin mediated degradation of growth factor receptors in the pathogenesis and treatment of cancer. 《BREAST CANCER RESEARCH》. 2002, 第 5 卷 (第 1 期), 8-15.

审查员 王晓媛

(51) Int. Cl.  
 C12Q 1/68 (2006. 01)  
 G01N 33/48 (2006. 01)

权利要求书 1 页 说明书 41 页  
 序列表 18 页 附图 16 页

(54) 发明名称  
 诊断乳腺癌的方法

(57) 摘要

本文描述了检测、诊断和治疗乳腺癌 (BRC) 的组合物和方法。本发明具体描述了三种与 BRC 相关的基因, 本文将其称为 A5657, B9769 和 C7965, 与正常细胞相比, 所述基因在 BRC 细胞中被上调。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。

CN 1898563 B

1. 筛选治疗或预防乳腺癌的化合物的方法,所述方法包括下述步骤:
  - a) 使受试化合物与由 SEQ ID NO :1 定义的多核苷酸编码的多肽接触;
  - b) 检测多肽和受试化合物之间的结合活性;和
  - c) 选择能结合多肽的受试化合物。
2. 筛选治疗或预防乳腺癌的化合物的方法,所述方法包括下述步骤:
  - a) 使候选化合物与表达标记基因的细胞接触,其中标记基因由 SEQ IDNO :1 定义;和
  - b) 选择与对照相比能降低所述一种或多种标记基因的表达水平的候选化合物。
3. 权利要求 2 的方法,其中所述细胞含有乳腺癌细胞。
4. 筛选治疗或预防乳腺癌的化合物的方法,所述方法包括下述步骤:
  - a) 使受试化合物与由 SEQ ID NO :1 定义的多核苷酸编码的多肽接触;
  - b) 检测步骤 (a) 的多肽的生物活性;和
  - c) 选择与缺乏受试化合物时检测的所述多肽的生物活性相比,能抑制所述多肽生物活性的受试化合物。
5. 筛选治疗或预防乳腺癌的化合物的方法,所述方法包括下述步骤:
  - 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转录调节区的控制之下表达的报道基因,其中标记基因由 SEQ ID NO :1 定义;
  - b) 测定所述报道基因的表达水平或活性;和
  - c) 选择与对照相比能降低所述报道基因的表达水平或活性的候选化合物。
6. 权利要求 1 的方法,其中所述乳腺癌是 IDC。
7. 权利要求 2 的方法,其中所述乳腺癌是 IDC。
8. 权利要求 6 或 7 的方法,其中所述细胞含有 IDC 细胞。
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. 针对由 SEQ ID NO :1 定义的多核苷酸的小的干扰 RNA 在制备用于治疗或预防乳腺癌的组合物中的用途,其中所述 siRNA 含有有义链,所述有义链含有 SEQ ID NO :28 或 29 的核苷酸序列作为靶序列。

## 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)。

### 技术领域

[0002] 本发明涉及诊断乳腺癌的方法。

[0003] 发明背景

[0004] 乳腺癌是一种复杂的疾病,其特征在于大量基因中存在多个遗传的和渐成的改变 (Katherine N. N., Richard W. and Barbara L. W. Breast cancer genetics: What we know and what we need. *Nat Med*, 7(5): 552-556, 2001.)。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,包括非典型性导管增生、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。尽管乳腺癌发生的阶段类似于其它组织中的癌症,但推动乳腺癌进程的精确分子机制仍是未知的。无论如何,有一点是显而易见的,即导致原发性乳腺癌发生、发展和转移的分子因素将会成为有价值的靶标,可用于开发较好的乳腺癌早期诊断、治疗和预防的手段。

[0005] 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/或治疗乳腺癌的新策略的主要焦点。

[0006] 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; Johannes B., Esther Z. and Axel U. Molecular targets for breast cancer therapy and prevention. *Nat Med*, 7(5): 548-552, 2001.)。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 DNA microarray 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 cDNA microarray. *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 cDNA microarray. *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 DNA microarray 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. )。

[0007] 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 cDNA microarrays. *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. )。

[0008] 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。

[0009] 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。

[0010] 经证实, CD8<sup>+</sup> 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-ESO-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。

[0011] 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 Cancer*81 :387-94(1999))。

[0012] 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 <sup>51</sup>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 Antigens*47 :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 USA93 :4102-7(1996))。

[0013] 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, 我们使用展现了 23, 040 个基因的 cDNA 微阵列对纯化的乳腺癌细胞群中的基因表达模式进行了大规模分析。更具体地, 联合使用 cDNA 微阵列和激光束显微解剖, 检查了 81 个乳腺肿瘤精确的基因组范围内的表达分布图, 所述肿瘤包括 12 个原位导管癌 (DCIS) 和 69 个浸润性导管癌 (IDC)。在鉴定出的上调基因中, 下述 3 个基因因为在乳腺癌细胞中的显著过表达而被选中: A5657 (以前称之为 AF161499); B9769 (以前称之为 AA156269) 和 C7965 (以前称之为 AW977394)。本文详细讨论的发现暗示着这些基因在肿瘤细胞生长增殖中起关键作用, 因此是开发抗癌药物的理想靶标。

[0014] 发明概述

[0015] 本发明涉及以下 3 个基因, 在所检测的乳腺癌病例中, 所述基因分别在 49 个病例中的 38 例, 73 个病例中的 30 例和 49 个病例中的 28 例中被显著上调: A5657 (SEQ ID NO :1), 编码类似于遍在蛋白缀合酶的 HSPC150 蛋白 (SEQ ID NO :2); B9769 (SEQ ID NO :3), 编码假拟蛋白 BC016861 (SEQ ID NO :4); 和 C7965 (SEQ ID NO :5), 编码 LOC90557 蛋白 (SEQ ID NO :6)。随后的半定量 RT-PCR 和 Northern 印迹分析证实: 与正常人组织, 包括乳腺导管细胞和正常乳腺相比, 临床乳腺癌样品和乳腺癌细胞系中的 A5657, B9769 和 C7965 被显著过表达。特别是 B9769, 它在 Er  $\alpha$  阳性乳腺癌细胞系中高表达。免疫细胞化学染色表明: 在乳腺癌细胞系 T47D 细胞中, 外源性 A5657, B9769 和 C7965 分别定位于细胞质和 / 或核器, 细胞质和细胞质器。特别是在 COS7 和 T47D 细胞的中间丝状网络中观察到外源性 B9769。另外, 通过免疫沉淀结合试验证实 A5657 蛋白与遍在蛋白相互作用, 暗示 A5657 蛋白潜在具有 E2 遍在蛋白 - 酶活性。分别用小干扰 RNA (siRNA) 处理乳腺癌细胞能有效抑制 A5657, B9769 和 C7965 的表达, 并抑制乳腺癌细胞 / 肿瘤生长。另外还发现当在 NIH3T3 细胞中瞬时过表达时, 所有这 3 个基因都能在 MTT 试验中显著促进细胞增殖, 暗示它们在细胞生长增殖中起关键作用。这些发现表明 A5657, B9769 和 C7965 的过表达可能参与乳腺肿瘤发生, 并且可能成为特异性治疗乳腺癌患者的理想策略。

[0016] 因此, 本发明包括发现这 3 个基因的表达与乳腺癌 (BRC) 显著相关。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- 相关基因是一个或多个选自 A5657 (SEQ ID NO :1), B9769 (SEQ ID NO :3) 和 C7965 (SEQ ID NO :5) 的基因。表达水平的改变, 例如与基因的正常对照水平相比, BRC- 相关基因的表达水平提高表示受试者患有 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的一个或多个选自 A5657, B9769 和 C7965 的 BRC- 相关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。

[0020] 或者,将样品中两个或多个 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,当基因表达水平与对照水平相比增加至少 10%,至少 25%,至少 50%或更高时,可认为基因表达是“经改变的”。或者,如果与对照水平相比增加至少 1 倍,至少 2 倍,至少 5 倍或更多倍时,可认为基因表达是“经改变的”。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基因表达。

[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,其中包含两个或多个选自 A5657, B9769 和 C7965 的 BRC- 相关基因的基因表达水平。

[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活性来鉴定药剂的方法,所述药剂能抑制 BRC- 相关基因,例如选自 A5657, B9769 和 C7965 的 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 BRC- 相关基因或其基因产物的正常对照水平或活性相比,所述基因或其基因产物的表达水平或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,例如降低选自 A5657, B9769 和 C7965 的 BRC- 相关基因的表达。

[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 BRC- 相关基因 A5657, B9769 和 C7965 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短的干扰 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 BRC- 相关基因 A5657, B9769 和 C7965 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 BRC- 相关基因 A5657, B9769 和 C7965 的 BRC 核酸的表达。

[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 BRC- 相关基因 A5657, B9769 和 C7965 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）生成、细胞毒 T 淋巴细胞诱导或抗体产生即可测定免疫细胞刺激。

[0028] 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。

[0029] 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。

[0030] 附图简述

[0031] 图 1 显示了半 - 定量 RT-PCR 的结果。图中显示出 (a)A5657, (b)B9769 和 (c) C7965 在乳腺癌患者的肿瘤细胞 (3T, 31T, 149T, 175T, 431T, 453T, 491T, 554T, 571T, 709T, 772T 和 781T), 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) 和正常人组织中的表达水平。Pre 表示正常乳腺导管细胞, MG 表示乳腺, LUN 表示肺, LIV 表示肝, HEA 表示心脏, KID 表示肾脏, BM 表示骨髓。

[0032] 图 2 显示出在多种人组织 (上面的一组)、乳腺癌细胞系和正常人的重要器官 (下面的一组) 中, (a)A5657, (b)B9769 和 (c)C7965 转录物的 Northern 印迹分析结果。

[0033] 图 3 显示出外源性 (a)A5657, (b)B9769 (上面的一组, 低密度的细胞; 下面的一组, 高密度的细胞) 和 (d)C7965 以及对照模拟物 (Mock) 在 T47D 乳腺癌细胞中的亚细胞定位。(c) 部分表示外源性 B9769 与其它细胞骨架蛋白相比的亚细胞定位。

[0034] 图 4 显示出被设计用于降低乳腺癌细胞中的 A5657 表达的小 - 干扰 RNA (siRNA) 的生长 - 抑制作用。(a) 部分表示半 - 定量 RT-PCR 的结果, 该结果表明: 将 siRNA 导入 T47D 细胞之后, 在含有新霉素的选择性培养基中培养, 在 28 天的培养物中, 乳腺癌细胞 (T47D) 中 A5657 的内源性表达受到抑制。将 GAPDH 用作内部对照。(b) 部分表示 MTT 试验的结果, 该结果表明通过击倒 T47D 细胞中的 A5657, 集落数目有所减少。(c) 部分表示集落 - 形成试验的结果, 该结果表明通过击倒 T47D 细胞中的 A5657, 集落数目有所减少。

[0035] 图 5 显示出被设计用于降低乳腺癌细胞中的 B9769 表达的小 - 干扰 RNA (siRNA) 的生长 - 抑制作用。(a) 部分表示半 - 定量 RT-PCR 的结果, 该结果表明: 将 siRNA 导入 T47D 细胞之后, 在含有新霉素的选择性培养基中培养, 在 28 天的培养物中, 乳腺癌细胞 (T47D) 中 B9769 的内源性表达受到抑制。将 GAPDH 用作内部对照。(b) 部分表示 MTT 试验的结果, 该结果表明通过击倒 T47D 细胞中的 B9769, 集落数目有所减少。(c) 部分表示集落 - 形成试验的结果, 该结果表明通过击倒 T47D 细胞中的 B9769, 集落数目有所减少。

[0036] 图 6 显示出被设计用于降低乳腺癌细胞中的 C7965 表达的小 - 干扰 RNA (siRNA) 的生长 - 抑制作用。(a) 部分表示半 - 定量 RT-PCR 的结果, 该结果表明: 将 siRNA 导入 T47D 细胞之后, 在含有新霉素的选择性培养基中培养, 在 28 天的培养物中, 乳腺癌细胞 (T47D) 中 C7965 的内源性表达受到抑制。将 GAPDH 用作内部对照。(b) 部分表示 MTT 试验的结果, 该结果表明通过击倒 T47D 细胞中的 C7965, 集落数目有所减少。(c) 部分表示集落 - 形成试验的结果, 该结果表明通过击倒 T47D 细胞中的 C7965, 集落数目有所减少。



[0037] 图 7 显示出 A5657, B9769 和 C7965 的过表达可增加细胞增殖的速率。将 NIH3T3 细胞铺于 6 孔培养板, 并分别用所示的 A5657, B9769 和 C7965 表达载体瞬时转染。在第 1, 2, 4 和 6 天进行 MTT 试验以监测表示细胞增殖的 O.D. 值。柱表示平均值  $\pm$  S.E. ( $n = 2$ )。

[0038] 图 8 表示 A5657 蛋白的遍在蛋白化。(A) 部分表示体内遍在蛋白化试验的结果。对细胞裂解物直接进行抗 -Flag 印迹 (左) 或先用抗 -HA 抗体免疫沉淀再进行抗 -Flag 免疫印迹 (右)。(B) 部分表示对细胞裂解物直接进行抗 -HA 印迹 (左) 或先用抗 -FLAG 抗体免疫沉淀再进行抗 -HA 免疫印迹 (右)。遍在蛋白与 FLAG-A5657 偶联。Western 印迹上多出的条带表示遍在蛋白化的 A5657。

[0039] 详细描述

[0040] 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。

[0041] 乳腺癌细胞一般以实体肿块的形式存在, 其具有高度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。

[0042] 考虑到这些问题, 本发明人通过激光束显微解剖法 (LMM) 制备纯化的乳腺癌细胞群和正常的乳腺上皮导管细胞群。使用展示了 23, 040 个基因的 cDNA 微阵列分析得自 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 (数据未显示) 鉴定为在 BRC 细胞中一般会上调的基因, 将 288 个基因 (数据未显示) 鉴定为在 BRC 细胞中一般会下调的基因。选择具有检测患者血清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。本发明具体涉及以下发现, 即在正常乳腺上皮细胞和 BRC 患者的癌之间, 3 个核酸, 也即 A5657, B9769 和 C7965 的表达模式有所改变。A5657 基因 (SEQ ID NO :1) 构成新序列并编码类似于遍在蛋白缀合酶的 HSPC150 蛋白 (SEQ ID NO :2) (Genbank 登录号 NM\_014176)。与 DCIS 细胞和正常乳腺上皮细胞相比, IDC 细胞中的 A5657 被上调。B9769 基因 (SEQ ID NO :3) 构成新序列并编码假拟蛋白 (SEQ ID NO :4) (Genbank 登录号 NM\_138770)。与正常乳腺上皮细胞相比, IDC 细胞中的 B9769 被上调。C7965 基因 (SEQ ID NO :5) 及其编码的蛋白质 (SEQ ID NO :6) 构成被称为 LOC90557 的已知序列。与正常乳腺上皮细胞相比, DCIS 细胞和 IDC 细胞中的 C7965 被上调。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 A5657 基因能用作区分 IDC 和 DCIS 的诊断标记以及诊断用的 BRC 基因靶, 可改变该基因的表达以治疗或减轻 IDC 症状。

[0043] 本文将这些在 BRC 患者中表达水平受调制 (即增加) 的基因统称为“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 (例如选自 A5657, B9769 和 C7965 的 BRC- 相关基因)。

[0044] 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。

[0045] 本发明包括检测 (例如测定) 选自 A5657, B9769 和 C7965 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。

[0046] 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)。

[0047] 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。

[0048] 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 1.0,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”。

[0049] 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。

[0050] 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。

[0051] 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。

[0052] 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。

[0053] 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、马或牛。

[0054] 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,或者可以按常规方法鉴定。

**[0055] 新的核苷酸、多肽、载体和宿主细胞**

**[0056]** 本发明提供了新的人基因 A5657 和 B9769,与相应的非癌乳腺上皮相比,在大多数乳腺癌中,所述基因的表达显著提高。分离的 A5657 基因包含如 SEQ ID NO :1 所述的多核苷酸序列,即含有 928 个核苷酸的 cDNA 序列。A5657 具体由 7 个外显子组成,并位于在基因组中跨越约 10.3kb 的染色体 1q32.1 上。cDNA 转录物最终编码 197 个氨基酸的多肽,即类似于遍在蛋白缀合酶的 HSPC150 蛋白。分离的 B9769 基因包含如 SEQ ID NO :3 所述的多核苷酸,即含有 1472 个核苷酸的 cDNA 序列。B9769 具体由 8 个外显子组成,并位于在基因组中跨越约 5.7kb 的染色体 2q21.2 上。ORF 起始于外显子 1,终止于外显子 8。cDNA 转录物最终编码 378 个氨基酸的多肽。

**[0057]** 除了包含 SEQ ID NO :1 和 3 所述多核苷酸序列的新型人基因 A5657 和 B9769 外,本发明还包括其简并序列和突变体,条件是它们编码包含 SEQ ID NO :2 和 4 所示氨基酸序列的 A5657 或 B9769 蛋白或其功能等同物。与 A5657 或 B9769 功能等同的多肽的例子包括例如源自其它生物体的与人 A5657 或 B9769 蛋白相对应的同源性蛋白质以及所述人蛋白质的突变体。

**[0058]** 本发明还包括与本发明人鉴定的多肽功能等同的多肽和编码所述功能等同多肽的多核苷酸。在本发明的上下文中,术语“功能等同”指的是该多肽保留了 A5657 或 B9769 蛋白(氨基酸序列分别示于 SEQ ID NO :2 和 4)的特征性生物学显著活性。例如,A5657 和 B9769 基因的特征在于:与正常细胞相比,BRC 细胞中的这两个基因被特异性过表达。与 DCIS 细胞相比,IDC 细胞中的 A5657 基因被过表达。另外,本文阐明它们的过表达可促进细胞增殖。因此,在本发明的上下文中,A5657 或 B9769 蛋白的功能等同物应具有类似于野生型蛋白质的细胞增殖活性。细胞增殖是可以用本领域的常规检测法和技术(例如下文实施例中描述的 MTT 试验)测定的参数。另外,本文阐明 A5657 蛋白具有 E2 遍在蛋白酶活性并能与遍在蛋白结合。因此,在本发明的上下文中,A5657 蛋白的功能等同物应具有类似于野生型蛋白质的遍在蛋白酶和/或结合活性。

**[0059]** 因此,本发明还希望包括本文所公开序列的某些突变体或变体。例如,本发明包括编码 SEQ ID NO :2 或 4 的蛋白质的多核苷酸,所述蛋白质中的一个或多个氨基酸可被取代、缺失、插入和/或添加,只要所得蛋白质能保留野生型蛋白质的生物学显著活性即可。在优选实施方案中,与正常细胞相比,BRC 细胞中由本发明的多核苷酸编码的功能等同蛋白质类似地被过表达。本发明还包括在严紧条件下与 A5657 或 B9769 DNA(分别包含 SEQ ID NO :1 和 3 的核苷酸序列)杂交的多核苷酸,只要所得多核苷酸编码与 A5657 或 B9769 蛋白功能等同的蛋白质即可。通过本领域技术人员众所周知的方法,包括本文所述的方法(参见实施例一节)可测定生物学显著活性。

**[0060]** 本发明提供了分离的多核苷酸,其编码本文所述的多肽或其片段。优选分离的多肽由与 SEQ ID NO :1 或 3 所示核苷酸序列至少约 60% 相同的核苷酸序列编码。更优选分离的核酸分子与 SEQ ID NO :1 或 3 所示核苷酸序列的同源性至少约为 65%,70%,75%,80%,85%,90%,91%,92%,93%,94%,95%,96%,97%,98%,99% 或更高。

**[0061]** 为了测定两个核酸的同源性百分比,以获得最佳比较为目的来比对序列。两个序列之间的同源性百分比是这两个序列共有的相同位置数目的函数(即同源性百分比=相

同位置数目 / 位置 (例如重叠位置) 总数乘以 100)。可使用如本文所述的常规技术, 在允许或不允许缺口存在的情况下测定两个序列之间的同一性百分比。在计算同一性百分比时, 一般计数精确的匹配。

[0062] 可使用任何常规的数学算法, 例如 Karlin 和 Altschul 的 BLAST 算法 (S. Karlin and S. F. Altschul, Proc. Natl. Acad. Sci. USA. 1990, 87 :2264-2268 ;S. Karlin and S. F. Altschul, Proc. Natl. Acad. Sci. USA. 1993, 90 :5873-5877) 测定两个序列之间的同一性百分比。BLAST 算法被编入 Altschul 等的 BLASTN 和 BLASTX 程序 (S. F. Altschul et al., J. Mol. Biol. 1990, 215 :403)。当根据 BLASTN 分析核苷酸序列时, 适当的参数包括例如分值 = 100, 字长 = 12。另一方面, 通过 BLASTX 分析氨基酸序列的适当参数包括例如分值 = 50, 字长 = 3。为了获得以比较为目的的带缺口比对, 可按 Altschul et al. (1997) Nucleic Acids Res. 25 :3389 所述利用带缺口的 BLAST。或者使用 PSI-Blast 进行叠代检索, 以检测分子之间的亲疏关系。当利用 BLAST, 带缺口的 BLAST 和 PSI-Blast 程序时, 优选使用各个程序 (例如 XBLAST 和 NBLAST) 的缺省参数。然而, 本领域技术人员可以容易地调整参数以适应具体的目的。

[0063] 另一个可用于比较序列的数学算法的例子是 Myers and Miller (1988) CABIOS4 : 11-17 中的算法。所述算法被编入 ALIGN 程序 (2.0 版), 该程序是 GCG 序列比对软件包的一部分。当利用 ALIGN 程序比较氨基酸序列时, 可使用 PAM120 优势残基表 (weight residue table), 使缺口长度罚分为 12, 缺口罚分为 4。

[0064] 与给定蛋白质功能等同的多肽的制备方法是本领域技术人员众所周知的, 并且包括在蛋白质中导入突变的常规方法。例如, 本领域技术人员通过定点诱变 (Hashimoto-Gotoh et al., Gene 152 :271-275 (1995) ;Zoller and Smith, Methods Enzymol 100 :468-500 (1983) ;Kramer et al., Nucleic Acids Res. 12 :9441-9456 (1984) ;Kramer and Fritz, Methods Enzymol 154 :350-367 (1987) ;Kunkel, Proc Natl Acad Sci USA 82 :488-492 (1985) ;Kunkel, Methods Enzymol 85 :2763-2766 (1988)) 在 A5657 或 B9769 蛋白的氨基酸序列中导入适当的突变, 即可制备出与所述的任一种蛋白质功能等同的多肽。已知经突变或修饰的蛋白质能保留原始的生物活性, 其中所述蛋白质具有因在某个氨基酸序列中取代、缺失、插入和 / 或添加一个或多个氨基酸残基而被修饰的氨基酸序列 (Mark et al., Proc Natl Acad Sci USA 81 :5662-5666 (1984) ;Zoller and Smith, Nucleic Acids Res 10 :6487-6500 (1982) ;Dalbadié-McFarland et al., Proc Natl Acad Sci USA 79 :6409-6413 (1982))。氨基酸突变也可自然发生。因此, 本发明包括具有 A5657 或 B9769 蛋白的氨基酸序列的蛋白质, 其中一个或多个氨基酸已经突变, 条件是所得突变多肽与野生型 A5657 或 B9769 蛋白的功能等同。

[0065] 对可突变的氨基酸数目没有特别的限制, 只要能维持生物学显著活性即可。一般可突变多达约 50 个氨基酸, 优选突变约 30 个氨基酸, 更优选突变约 10 个氨基酸, 甚至更优选突变约 3 至 5 个氨基酸。类似地, 对突变位点没有特别的限制, 只要突变不会破坏生物学显著活性即可。

[0066] 可在一个或多个预测的、优选为非必需的氨基酸残基处进行氨基酸取代。“非必需的”氨基酸残基是可改变自蛋白质的野生型序列 (例如 SEQ IDNO :2 和 4 所示的序列), 却不会改变生物活性的残基, 而“必需的”氨基酸残基是生物活性所需的残基。优选将待突变

的氨基酸残基取代为不同的氨基酸,从而使氨基酸侧链的特性是保守的(已知为保守的氨基酸取代过程)。本文中的术语“保守的氨基酸取代”指的是氨基酸残基被具有类似化学侧链的氨基酸残基所取代。本领域已详细定义过具有类似侧链的氨基酸残基组。氨基酸组的例子包括具有以下侧链的氨基酸:碱性侧链(例如赖氨酸,精氨酸,组氨酸)、酸性侧链(例如天冬氨酸,谷氨酸)、不带电的极性侧链(例如甘氨酸,天冬酰胺,谷氨酰胺,丝氨酸,苏氨酸,酪氨酸,半胱氨酸)、非极性侧链(例如丙氨酸,缬氨酸,亮氨酸,异亮氨酸,脯氨酸,苯丙氨酸,甲硫氨酸,色氨酸)、 $\beta$ -分支侧链(例如苏氨酸,缬氨酸,异亮氨酸)和芳香族侧链(例如酪氨酸,苯丙氨酸,色氨酸,组氨酸)。氨基酸分组的其它例子包括具有下述链特性、特征和/或官能团或共有特性的侧链:疏水氨基酸(A, I, L, M, F, P, W, Y, V),亲水氨基酸(R, D, N, C, E, Q, G, H, K, S, T),脂肪族侧链(G, A, V, L, I, P);含羟基侧链(S, T, Y);含硫原子侧链(C, M);含羧酸和酰胺侧链(D, N, E, Q);含碱基侧链(R, K, H)和含芳香族侧链(H, F, Y, W)。需说明的是括号中的字母表示氨基酸的单字母代码。对可突变的氨基酸数目没有特别的限制,只要能维持活性即可。一般可突变多达约50个氨基酸,优选突变约30个氨基酸,更优选突变约10个氨基酸,甚至更优选突变约3个氨基酸。类似地,对突变位点没有特别的限制,只要突变不会破坏活性即可。或者,突变氨基酸的数目一般相当于总氨基酸的30%或更少,或20%或更少,或10%或更少,优选5%或更少,或3%或更少,更优选2%或更少或1%或更少。

[0067] 在A5657或B9769蛋白的氨基酸序列中添加一个或多个氨基酸残基所得多肽的例子是含有A5657或B9769蛋白的融合蛋白。融合蛋白是A5657或B9769蛋白与其它肽或蛋白质的融合蛋白,本发明包括所述融合蛋白。通过本领域技术人员众所周知的技术即可制备融合蛋白,例如将编码本发明的A5657或B9769蛋白的DNA与编码其它肽或蛋白质的DNA连接,使得读码框匹配,将融合DNA插入表达载体并在宿主中表达所述融合DNA。对于与本发明的蛋白质融合的肽或蛋白质没有限制。

[0068] 可用于与本发明的蛋白质融合的已知肽包括例如FLAG(Hopp et al., *Biotechnology*6:1204-1210(1988))、含有6个His(组氨酸)残基的6×His、10×His、流感病毒凝集素(HA)、人c-myc片段、VSP-GP片段、p18HIV片段、T7-标记、HSV-标记、E-标记、SV40T抗原片段、lck标记、 $\alpha$ -微管蛋白片段、B-标记、C蛋白片段等。可与本发明的蛋白质融合的蛋白质包括例如GST(谷胱甘肽-S-转移酶)、流感病毒凝集素(HA)、免疫球蛋白恒定区、 $\beta$ -半乳糖苷酶、MBP(麦芽糖-结合蛋白)等。

[0069] 通过将编码上述融合肽或蛋白质的商购DNA与编码本发明多肽的DNA融合,并表达所制备的融合DNA即可制备出融合蛋白。

[0070] 另一种本领域已知的分离功能等同多肽的方法包括常规的杂交技术(Sambrook et al., *Molecular Cloning*2nd ed. 9.47-9.58, Cold Spring Harbor Lab. Press(1989))。本领域技术人员能容易地分离出与编码人A5657或B9769蛋白的DNA序列(即SEQ ID NO:1或3)的全部或部分具有高度同源性的DNA,并且能从分离的DNA中分离出人A5657或B9769蛋白的功能等同多肽。本发明的多肽包括由能与编码A5657或B9769蛋白的DNA序列的全部或部分杂交的DNA编码、并且与A5657或B9769蛋白功能等同的多肽。所述多肽包括对应于人源蛋白质的哺乳动物同系物(例如由猴、大鼠、兔和牛基因编码的多肽)。

[0071] 本领域技术人员可以常规选择分离DNA所用的杂交条件,所述DNA编码与A5657

或 B9769 蛋白功能等同的多肽。例如,可通过于 68°C 使用“Rapid-hyb 缓冲液”(Amersham LIFE SCIENCE) 进行预杂交达 30 分钟或更长时间,加入经标记的探针并于 68°C 加热 1 小时或更长时间来进行杂交。例如,可在低严谨条件下进行下述洗涤步骤。低严谨条件是例如 42°C, 2×SSC, 0.1% SDS, 或优选为 50°C, 2×SSC, 0.1% SDS。更优选使用高度严谨条件。高度严谨条件是例如室温下用 2×SSC, 0.01% SDS 洗涤 3 次, 每次 20 分钟, 然后于 37°C 用 1×SSC, 0.1% SDS 洗涤 3 次, 每次 20 分钟, 再于 50°C 用 1×SSC, 0.1% SDS 洗涤 2 次, 每次 20 分钟。然而, 几个诸如温度和盐浓度的因素可影响杂交的严谨度, 本领域技术人员可适当选择这些因素以获得所需的严谨度。

[0072] 可以使用基因扩增法, 例如聚合酶链反应 (PCR) 法替代杂交来分离 DNA, 所述 DNA 编码与 A5657 或 B9769 蛋白功能等同的多肽, 所述方法使用基于编码蛋白质的 DNA 序列信息 (SEQ ID NO :1 或 3) 所合成的引物。

[0073] 由通过上述杂交技术或基因扩增技术分离的 DNA 编码的、与 A5657 或 B9769 蛋白功能等同的多肽一般与人 A5657 或 B9769 蛋白的氨基酸序列具有高度同源性。本文中的“高度同源性”一般指的是同源性为 40% 或更高, 优选为 60% 或更高, 更优选为 80% 或更高, 甚至更优选为 85%, 90% 或 95% 或更高。根据“Wilbur and Lipman, Proc Natl Acad Sci USA80 :726-30 (1983)”中的算法即可测定多肽的同源性。

[0074] 本发明多肽的氨基酸序列、分子量、等电点、糖链的存在或缺乏、或形式可以有所不同, 这取决于用于产生所述多肽的细胞或宿主或所使用的纯化方法。无论如何, 只要其功能等同于本发明的 A5657 或 B9769 蛋白, 就落入本发明的范围。

[0075] 通过本领域技术人员众所周知的方法可将本发明的多肽制备成重组蛋白质或天然蛋白质。通过下述方法即可制备重组蛋白质, 即将编码本发明多肽的 DNA (例如含有核苷酸序列 SEQ ID NO :1 或 3 的 DNA) 插入适当的表达载体, 将载体导入适当的宿主细胞, 获得提取物, 通过对提取物进行层析来纯化多肽, 所述层析包括例如离子交换层析、反相层析、凝胶过滤、或利用固定有抗本发明蛋白质之抗体的层析柱的亲亲和层析, 或一种以上所述柱的组合。另外, 当在宿主细胞 (例如动物细胞和大肠杆菌) 中将本发明的多肽表达成与谷胱甘肽-S-转移酶的融合蛋白或添加有多个组氨酸的重组蛋白质时, 可使用谷胱甘肽柱或镍柱来纯化所表达的重组蛋白质。或者, 当将本发明的多肽表达成被 c-myc, 多个组氨酸或 FLAG 标记的蛋白质时, 可分别使用抗 c-myc, His 或 FLAG 的抗体来检测和纯化所述蛋白质。

[0076] 纯化融合蛋白之后, 必要时也可以通过用凝血酶或 Xa 因子切割来切除非目的多肽区域。通过本领域技术人员已知的方法, 例如通过使结合有能与 A5657 或 B9769 蛋白结合的下述抗体的亲和柱与表达本发明多肽的组织或细胞提取物接触, 即可分离出天然蛋白质。抗体可以是多克隆抗体或单克隆抗体。

[0077] 本发明还包括本发明多肽的部分肽。部分肽具有本发明多肽所特有的氨基酸序列, 并由至少 7 个氨基酸, 优选为 8 个或更多个氨基酸, 更优选为 9 个或更多个氨基酸组成。部分肽可用于例如制备抗本发明多肽的抗体, 筛选与本发明多肽结合的化合物, 以及筛选本发明多肽的促进剂或抑制剂。

[0078] 通过基因工程, 通过已知的肽合成法或通过用适当的肽酶消化本发明的多肽, 即可产生本发明的部分肽。例如, 可以使用固相合成或液相合成来进行肽合成。

[0079] 另外, 本发明提供了编码本发明多肽的多核苷酸。本发明的多核苷酸的例子包括

含有 SEQ ID NO :1 或 3 的核苷酸序列的 DNA。本发明的多核苷酸可用于体内或体外产生如上所述的本发明多肽,或者可用于因编码本发明蛋白质的基因中的基因异常所致疾病的基因治疗。可以使用任何形式的本发明的多核苷酸,只要它能编码本发明的多肽即可,其包括 mRNA, RNA, cDNA, 基因组 DNA 和化学合成的多核苷酸。本发明的多核苷酸包括含有给定核苷酸序列及其简并序列的 DNA, 只要所得 DNA 能编码本发明的多肽即可。

[0080] 优选本发明的多核苷酸是分离的。本文所用的“分离的多核苷酸”是一种取自其原有的环境(例如如果是天然的,即为其天然环境),通过人为手段改变自其天然状态的多核苷酸,其结构与任何天然核酸结构或横跨三个以上基因的天然基因组核酸的任何片段的结构都不相同。因此,该术语包括例如(a)一种 DNA,其片段具有天然基因组 DNA 分子的部分序列,该 DNA 不含有在基因组 DNA 中天然邻接该核酸的编码序列(即位于该核酸 5' 和 3' 末端的序列),并且该 DNA 与在其天然存在的生物体的基因组中邻接该分子部分的任一编码序列都不邻接;(b)以所得分子不同于任何天然载体或基因组 DNA 的方式掺入原核或真核载体或基因组 DNA 的核酸;(c)分离的分子,如 cDNA、基因组片段、聚合酶链反应(PCR)产生的片段或限制性片段;和(d)重组核苷酸序列,其为杂合基因,即编码融合蛋白的基因的一部分。该定义特别排除的是不同 DNA 分子、转染细胞或细胞克隆的随机的未鉴定混合物中存在的核酸,例如 DNA 文库,如 cDNA 或基因组 DNA 文库中出现的核酸。

[0081] 可通过本领域技术人员众所周知的方法分离本发明的多核苷酸。所述方法包括例如上文讨论的杂交技术(E. M. Southern, J. Mol. Biol. 1975, 98 :503-517)以及聚合酶链反应(PCR)技术(R. K. Saiki et al., Science 1985, 230 :1350-1354; R. K. Saiki et al., Science 1988, 239 :487-491)。更具体地,本领域技术人员使用 SEQ ID NO :1 和 3 所示的多核苷酸或其部分作为探针,或使用能与 SEQ ID NO :1 和 3 所示的多核苷酸特异性杂交的寡核苷酸作为引物,一般可从其它动物(如人)中分离出与 SEQ ID NO :1 和 3 所示的多核苷酸高度同源的多核苷酸。另外,本发明的多核苷酸还包括可通过杂交技术或 PCR 技术进行分离,并能与 SEQ ID NO :1 和 3 所示的多核苷酸杂交的多核苷酸。

[0082] 优选在严紧条件下进行分离上述多核苷酸的杂交反应。用允许在含有一些错配的核酸序列之间形成杂交复合物的缓冲液进行杂交。在高严紧度下,仅当核酸分子几乎完全互补时杂交复合物才能保持稳定。多个因素决定了杂交的严紧度,包括 cDNA 的 G+C 含量、盐浓度和温度。例如,通过降低盐浓度或升高杂交温度即可增加严紧度。杂交和洗涤的温度条件对严紧度影响较大,可使用解链温度( $T_m$ )对其进行调节。 $T_m$  随着杂交碱基对中组成核苷酸的比例以及杂交溶液组分(盐、甲酰胺和十二烷基硫酸钠的浓度)的变化而变化。在一些基于膜的杂交所用的溶液中添加有机溶剂,如甲酰胺可使反应在较低的温度下进行。因此,考虑到相关参数,本领域技术人员可选择适当的条件以获得基于适当严紧度的经验或实验方法。

[0083] 严紧杂交条件的例子包括含有 6M 尿素, 0.4% SDS 和  $0.5\times$ SSC 的条件,以及具有与该条件等同的严紧度的条件。当在具有较高严紧度的条件(例如 6M 尿素, 0.4% SDS 和  $0.1\times$ SSC)下进行杂交时,有望分离出具有较高同源性的多核苷酸。在如上所述的较高严紧度条件下分离的多核苷酸有望编码在氨基酸水平上与 SEQ ID NO :2 和 4 所示氨基酸序列具有较高同源性的多肽。上述的“高同源性”指的是在整个氨基酸序列中同一性至少为 40% 或更高,优选为 60% 或更高,更优选为 80% 或更高,甚至更优选为 85%, 90% 或 95% 或

更高。

[0084] 本发明还包括 SEQ ID NO :1 和 3 所示多核苷酸的等位基因变体。等位基因是两种或多种可替换的基因形式中的一种,它占据具体染色体或连锁结构中的相同基因座,但在一个或多个突变位点处,它与所述基因座的其它等位基因有所不同。因此,等位基因变体在正常的基因序列中含有改变。完全基因测序经常能鉴定出给定基因的多个等位基因变体(有时是数百个)。等位基因变体与原始的多核苷酸具有高百分比的同一性,每一百个碱基约有例如 3 个碱基有所不同。等位基因变体一般编码基本上相同的蛋白质,一般不会影响所得的表型。通过例如上述的杂交技术,可以常规获得具体多核苷酸的等位基因变体。

[0085] 等位基因变体的例子包括遗传多态性。多态性包括个体中存在的 DNA 序列的差异。对遗传连锁分析而言,在 1% 以上的人口出现的基因变异被认为是有用的多态性。某些多态性(例如限制性片段长度多态性(RFLP))导致个体中由具体的酶,如限制性酶切割的 DNA 片段大小发生变化。导致 RFLP 的多态性序列可用作物理图谱和遗传连锁图上的标记。另一例常见的遗传多态性是“单核苷酸多态性”(SNP)。SNP 因取代、插入或缺失而在单个碱基中含有改变。改变可以是保守的(用嘌呤替代嘌呤)或是非-保守的(将嘌呤变成嘧啶),可以导致或不导致编码氨基酸的改变。所述改变可使个体易感具体的疾病。

[0086] 通过本领域技术人员已知的方法即可制备本发明的多核苷酸。例如,可通过下述方法制备本发明的多核苷酸:由表达本发明多肽的细胞制备 cDNA 文库,并使用本发明 DNA(如 SEQ ID NO :1 或 3)的部分序列作为探针进行杂交。例如,通过 Sambrook et al., Molecular Cloning, Cold Spring Harbor Laboratory Press(1989)中所述的方法即可制备 cDNA 文库;或者也可以使用商购的 cDNA 文库。也可通过下述方法制备 cDNA 文库:从表达本发明多肽的细胞中提取 RNA,基于本发明 DNA 的序列(如 SEQ ID NO :1 或 3)合成寡 DNA,使用寡 DNA 作为引物进行 PCR,并扩增编码本发明蛋白质的 cDNA。

[0087] 另外,通过对所得 cDNA 的核苷酸进行测序,即可常规测定由 cDNA 编码的翻译区,并能容易地获得本发明多肽的氨基酸序列。另外,通过使用所得 cDNA 或其部分作为探针筛选基因组 DNA 文库,即可分离出基因组 DNA。

[0088] 使用逆转录酶,用所得 mRNA 合成 cDNA。可使用商购的试剂盒,如 AMV 逆转录酶第一链 cDNA 合成试剂盒(Seikagaku Kogyo)合成 cDNA。或者,可以根据 5' -RACE 法(Frohman et al., Proc Natl Acad Sci USA85 :8998-9002(1988); Belyavsky et al., Nucleic Acids Res17 :2919-2932(1989))合成和扩增 cDNA,所述方法使用本文所述的引物、5' -AmplifINDER RACE 试剂盒(Clontech)和聚合酶链反应(PCR)。

[0089] 由 PCR 产物制备所需 DNA 片段,并与载体 DNA 连接。使用重组载体转化大肠杆菌等,由选定菌落制备所需重组载体。通过常规方法,如双脱氧核苷酸链终止法证实所需 DNA 的核苷酸序列。

[0090] 考虑到用于表达的宿主中的密码子使用频率,将本发明多核苷酸的核苷酸序列设计成能够更有效地表达(Grantham et al., Nucleic Acids Res9 :43-74(1981))。可通过商购试剂盒或常规方法改变本发明多核苷酸的序列。例如,可通过用限制性酶消化、插入合成的寡核苷酸或适当的多核苷酸片段、添加接头或插入起始密码子(ATG)和/或终止密码子(TAA, TGA 或 TAG)来改变序列。

[0091] 本发明还提供了载体,其中插入了本发明的多核苷酸。本发明的载体可用于在宿



主细胞中维持本发明的多核苷酸,尤其是 DNA,从而表达本发明的多肽,或者可用于施用本发明的多核苷酸以进行基因治疗。

[0092] 当宿主细胞为大肠杆菌,并且在大肠杆菌(如 JM109, DH5 $\alpha$ , HB101 或 XL1Blue)中大量扩增和制备载体时,载体应具有欲在大肠杆菌中扩增的 "ori" 和用于选择转化的大肠杆菌的标记基因(例如用药物进行选择性的药物-抗性基因,如氨苄青霉素、四环素、卡那霉素、氯霉素等)。例如,可使用 M13-系列载体、pUC-系列载体、pBR322、pBluescript、pCR-Script 等。另外,也可使用 pGEM-T、pDIRECT 和 pT7 亚克隆和提取 cDNA 以及上述载体。当使用载体制备本发明的蛋白质时,表达载体特别有用。例如,欲在大肠杆菌中表达的载体应具有欲在大肠杆菌中扩增的上述特征。当将大肠杆菌,如 JM109, DH5 $\alpha$ , HB101 或 XL1Blue 用作宿主细胞时,载体应具有能在大肠杆菌中有效表达所需基因的启动子,例如 lacZ 启动子(Ward et al., Nature341 :544-546(1989);FASEB J6 :2422-2427(1992))、araB 启动子(Better et al., Science240 :1041-1043(1988))或 T7 启动子等。在此方面,可使用例如 pGEX-5X-1(Pharmacia)、“QIAexpress 系统”(Qiagen)、pEGFP 和 pET(此时,宿主优选为表达 T7RNA 聚合酶的 BL21)来替代上述载体。另外,载体也可含有多肽分泌所用的信号序列。介导多肽分泌至大肠杆菌周质的信号序列的例子为 pelB 信号序列(Lei et al., J Bacteriol169 :4379(1987))。将载体导入靶宿主细胞的方法包括例如氯化钙法和电穿孔法。

[0093] 除了大肠杆菌外,也可使用下述载体制备本发明的多肽,例如得自哺乳动物的表达载体(例如 pcDNA3(Invitrogen)和 pEGF-BOS(Nucleic Acids Res18(17) :5322(1990)), pEF, pCDM8),得自昆虫细胞的表达载体(例如“Bac-to-BAC 杆状病毒表达系统”(GIBCO BRL), pBacPAK8),得自植物的表达载体(例如 pMH1, pMH2),得自动物病毒的表达载体(例如 pHSV, pMV, pAdexLcw),得自逆转录病毒的表达载体(例如 pZIpneo),得自酵母的表达载体(例如“毕赤氏酵母表达试剂盒”(Invitrogen), pNV11, SP-Q01)以及得自枯草芽孢杆菌的表达载体(如 pPL608, pKTH50)。

[0094] 为了在动物细胞,如 CHO, COS 或 NIH3T3 细胞中表达载体,载体应具有在所述细胞中进行表达所必需的启动子,例如 SV40 启动子(Mulligan et al., Nature277 :108(1979))、MMLV-LTR 启动子、EF1 $\alpha$  启动子(Mizushima et al., Nucleic Acids Res18 :5322(1990))、CMV 启动子等,优选还具有选择转化子所用的标记基因(例如用药物进行选择性的药物抗性基因(如新霉素、G418))。具有这些特征的已知载体的例子包括例如 pMAM, pDR2, pBK-RSV, pBK-CMV, pOPRSV 和 pOP13。

[0095] 另外,可使用一些方法在稳定表达基因的同时在细胞内扩增基因的拷贝数。例如,可将含有互补 DHFR 基因的载体(如 pCHO 1)导入核酸合成途径缺失的 CHO 细胞,然后通过氨甲蝶呤(MTX)进行扩增。另外,当瞬时表达基因时,可以使用这样一种方法,其中含有 SV40 复制起点的载体(pcD 等)被转化至染色体上含有表达 SV40T 抗原的基因的 COS 细胞。

[0096] 按上述方法获得的本发明多肽可分离自宿主细胞的内部或外部(如培养基),并被纯化成基本上纯的均质多肽。本文所用的与给定多肽有关的术语“基本上纯”指的是多肽基本上不含污染物,如其它生物大分子,培养基(如果是重组生产的话)或化合物前体(如果是化学合成的话)。基本上纯的多肽是指干重至少约 75%,优选至少约 80%,更优选至少约 85,90,95 或 99%(例如至少 80,85,95 或 99%)纯。通过任何适当的标准常规方法皆

可测定纯度,所述方法包括例如柱层析、聚丙烯酰胺凝胶电泳或 HPLC 分析。

[0097] 例如,可以适当选择柱层析、过滤、超滤、盐沉淀、溶剂沉淀、溶剂提取、蒸馏、免疫沉淀、SDS-聚丙烯酰胺凝胶电泳、等电点电泳、透析和重结晶,并将它们组合用于分离和纯化多肽。

[0098] 层析的例子包括例如亲和层析、离子交换层析、疏水层析、凝胶过滤、反相层析、吸附层析等(Strategies for Protein Purification and Characterization :A Laboratory Course Manual.Ed.Daniel R.Marshak et al., Cold Spring Harbor Laboratory Press(1996))。通过液相层析,如 HPLC 和 FPLC 即可进行这些层析。因此,本发明提供了通过上述方法制备的高度纯化的多肽。

[0099] 通过在纯化前或纯化后用适当的蛋白质修饰酶处理本发明的多肽,即可对其进行任意修饰或部分缺失。有用的蛋白质修饰酶包括但不限于胰蛋白酶、胰凝乳蛋白酶、赖氨酰内肽酶、蛋白激酶、葡糖苷酶等。

[0100] 抗体 :

[0101] 本发明还提供了与本发明的多肽结合的抗体。本发明的抗体可以任何形式被使用,如单克隆或多克隆抗体,并且包括通过用本发明的多肽免疫动物,如兔获得的抗血清、所有类别的多克隆和单克隆抗体、人抗体和通过基因重组产生的人源化抗体。

[0102] 可用作抗原以获得抗体的本发明多肽可得自任何动物,但优选得自哺乳动物,如人、小鼠或大鼠,更优选得自人。人源多肽可得自本文公开的核苷酸或氨基酸序列。

[0103] 根据本发明,用作免疫抗原的多肽可以是完整蛋白质或蛋白质的部分肽。部分肽可含有例如本发明多肽的氨基(N)-末端或羧基(C)-末端片段。

[0104] 可用抗原免疫任何哺乳动物,但优选考虑与细胞融合所用亲代细胞的相容性。一般使用啮齿目、兔形目(Lagomorpha)和灵长目动物。啮齿目动物包括但不限于例如小鼠、大鼠和仓鼠。兔形目动物包括但不限于例如兔。灵长目动物包括但不限于例如 Catarrhini 下目的猴(old world 猴),如食蟹猴(Macaca fascicularis),恒河猴(rhesus monkey),sacred 狒狒和黑猩猩(chimpanzees)。

[0105] 用抗原免疫动物的方法是本领域已知的。腹膜内注射或皮下注射抗原是免疫哺乳动物的标准方法。更具体地,可稀释抗原,将其悬浮于适当量的磷酸缓冲盐水(PBS)、生理盐水等中。必要时,可将抗原悬浮液与适当量的标准佐剂,如弗氏完全佐剂混合,制成乳剂,然后施用给哺乳动物。优选随后按照 4 至 21 天的间隔数次施用与适当量的弗氏不完全佐剂混合的抗原。适当的载体也可用于免疫。按上文所述进行免疫之后,通过标准方法检查血清中所需抗体量的增加。

[0106] 通过从经检查后确定血清中所需抗体量有所增加的经免疫哺乳动物体内收集血液,并通过使用任何常规方法从血液中分离出血清,即可制备抗本发明多肽的多克隆抗体。多克隆抗体包括含有多克隆抗体的血清以及可以从血清中分离含有多克隆抗体的组分。可以使用例如与本发明多肽偶联的亲合柱获得仅识别本发明多肽的组分,并使用蛋白 A 或蛋白 G 柱进一步纯化该组分,从而由所述组分制备免疫球蛋白 G 或 M。

[0107] 为了制备单克隆抗体,从用抗原免疫的哺乳动物体内收集免疫细胞,按上文所述检查血清中水平有所增加的所需抗体,并进行细胞融合。用于细胞融合的免疫细胞优选得自脾脏。其它可与上述免疫细胞融合的优选亲代细胞包括例如哺乳动物的骨髓瘤细胞,更

优选为具有用药物选择融合细胞时所需的获得型特性的骨髓瘤细胞。

[0108] 可根据已知方法,例如 Milstein 等的方法 (Galfre and Milstein, *Methods Enzymol* 173 :3-46 (1981)) 融合上述免疫细胞和骨髓瘤细胞。

[0109] 通过在标准的选择培养基,如 HAT 培养基(含有次黄嘌呤、氨基蝶呤和胸苷的培养基)中培养即可选择出通过细胞融合获得的杂交瘤。一般在 HAT 培养基中维持细胞培养物数天至数周,维持的时间应足以使除所需杂交瘤以外的所有其它细胞(非融合细胞)死亡。然后进行标准的有限稀释以筛选和克隆能产生所需抗体的杂交瘤细胞。

[0110] 除了上文所述的用抗原免疫非人动物以制备杂交瘤的方法外,也可在体外用多肽、表达多肽的细胞或其裂解物免疫人淋巴细胞,例如被 EB 病毒感染的人淋巴细胞。然后,使经免疫的淋巴细胞与得自人的能够无限分裂的骨髓瘤细胞,如 U266 融合,从而产生能生成所需人抗体的杂交瘤,所述抗体能结合所述多肽(未审公开的日本专利申请 (JP-A) Sho63-17688)。

[0111] 随后,将所得杂交瘤移植到小鼠腹腔中并抽取腹水。可通过例如硫酸铵沉淀、蛋白 A 或蛋白 G 柱、DEAE 离子交换层析或偶联有本发明多肽的亲和柱来纯化所获得的单克隆抗体。本发明的抗体不仅可用于纯化和检测本发明的多肽,还可用作本发明多肽的候选激动剂和拮抗剂。另外,所述抗体还可用于与本发明多肽相关之疾病的抗体治疗。当给人体施用所得抗体时(抗体治疗),优选人抗体或人源化抗体以降低免疫原性。

[0112] 例如,可以用选自多肽、表达多肽的细胞或其裂解物的抗原免疫携有人抗体基因所有组成成分的转基因动物。然后从动物体内收集产生抗体的细胞并与骨髓瘤细胞融合,从而获得杂交瘤,从中可制备出抗所述多肽的人抗体(参见 W092-03918, W093-2227, W094-02602, W094-25585, W096-33735 和 W096-34096)。

[0113] 或者,可通过癌基因无限增殖化能产生抗体的免疫细胞,如经免疫的淋巴细胞,并将其用于制备单克隆抗体。

[0114] 也可使用基因工程技术重组制备如此获得的单克隆抗体(参见例如 Borrebaeck and Larrick, *Therapeutic Monoclonal Antibodies*, published in the United Kingdom by MacMillan Publishers LTD (1990))。例如,可从能产生抗体的免疫细胞,如杂交瘤或经免疫的淋巴细胞中克隆编码抗体的 DNA,将其插入适当载体,并导入宿主细胞以制备重组抗体。本发明还提供了按上文所述制备的重组抗体。

[0115] 另外,本发明的抗体可以是抗体片段或经修饰的抗体,只要它们能结合本发明的一种或多种多肽即可。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))。

[0116] 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。

[0117] 或者,本发明的抗体可以是非人抗体可变区与人抗体恒定区的嵌合抗体,或者是含有非人抗体互补决定区(CDR)、人抗体框架区(FR)和恒定区的人源化抗体。使用已知技术即可制备出所述抗体。

[0118] 可将按上文所述获得的抗体纯化至均质。例如,可根据用于普通蛋白质的分离和纯化方法进行抗体的分离和纯化。例如,通过适当选择和组合使用柱层析,如亲和层析、过滤、超滤、盐析、透析、SDS聚丙烯酰胺凝胶电泳、等电聚焦等可以分离抗体(Antibodies: A Laboratory Manual.Ed Harlow and David Lane, Cold Spring Harbor Laboratory, (1988)),但并不局限于此。蛋白A柱和蛋白G柱可用作亲和柱。可以使用的蛋白A柱的例子包括例如HyperD, POROS和Sepharose F.F.(Pharmacia)。

[0119] 除亲和层析以外的层析例包括例如:离子交换层析、疏水层析、凝胶过滤、反相层析、吸附层析等(Strategies for Protein Purification and Characterization: A Laboratory Course Manual.Ed Daniel R.Marshak et al., Cold Spring Harbor Laboratory Press(1996))。可通过液相层析,如HPLC和FPLC进行层析过程。

[0120] 例如,可以使用吸光度测定、酶联免疫吸附试验(ELISA)、酶免疫测定(EIA)、放射免疫测定(RIA)和/或免疫荧光来测定本发明抗体的抗原结合活性。在ELISA中,将本发明的抗体固定在培养板上,将本发明的多肽上样于板上,然后上样含有所需抗体的样品,如产生抗体的细胞的培养物上清液或纯化的抗体。然后上样能识别第一抗体并被酶,如碱性磷酸酶标记的第二抗体,并保温培养板。接着,在洗涤之后,在培养板中加入酶底物,如对硝基苯磷酸,测定吸光度以评价样品的抗原结合活性。多肽的片段,如C-末端或N-末端片段可用作抗原来评价抗体的结合活性。可使用BIAcore(Pharmacia)评价本发明抗体的活性。

#### [0121] 诊断乳腺癌

[0122] 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。

[0123] 测定受试细胞或生物样品中一种或多种BRC-相关基因,如A5657, B9769和C7965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,受试群体中一种或多种BRC-相关基因A5657, B9769和C7965表达的增加表示受试者患有BRC或有发展成为BRC的风险。

[0124] 与正常对照水平相比,受试群体中一种或多种BRC-相关基因表达水平的增加表示受试者患有BRC或有发展成为BRC的风险。例如,一系列BRC-相关基因(例如A5657, B9769和C7965)中一种、两种或三种基因的增加表示受试者患有BRC或有发展成为BRC的风险。

[0125] 鉴定抑制BRC-相关基因表达的药剂:

[0126] 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。与正常对照水平相比 (或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- 相关基因的抑制剂, 并可用于抑制 BRC。

[0127] 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。

[0128] 评价在受试者中治疗 BRC 的效力:

[0129] 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。

[0130] 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, 而受试细胞群体和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。

[0131] 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。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。

[0132] 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。

[0133] 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。

[0134] 选择适用于具体个体的 BRC 治疗剂:

[0135] 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。

[0136] 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,然后测定一种或多种 BRC- 相关基因 A5657, B9769 和 C7965 的表达。

[0137] 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,测定受试样品的基因表达模式,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。

[0138] 相对于含有 BRC 的参照细胞群体而言,一种或多种 BRC- 相关基因 A5657, B9769 和 C7965 表达的降低表示该药剂具有治疗潜力。

[0139] 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。

[0140] 鉴定治疗剂的筛选试验:

[0141] 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一种或多种 BRC- 相关基因 A5657, B9769 和 C7965 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。

[0142] 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),并测定细胞中一种或多种 BRC- 相关基因 A5657, B9769 和 C7965 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。

[0143] 能抑制过表达或上调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。

[0144] 在另一个实施方案中,本发明提供了筛选候选药剂的方法,所述药剂是治疗 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,即可鉴定出能用作治疗 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:

[0145] a) 使受试化合物与由 A5657, B9769 和 C7965 编码的多肽接触;

[0146] b) 检测多肽和受试化合物之间的结合活性;和

[0147] c) 选择能结合多肽的受试化合物。

[0148] 或者,本发明的筛选方法可包括下述步骤:

[0149] 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 A5657, B9769 和 C7965;和

[0150] b) 选择与对照相比能降低一种或多种标记基因的表达水平的候选化合物。

[0151] 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。

[0152] 或者,本发明的筛选方法可包括下述步骤:

[0153] a) 使受试化合物与由选自 BRC- 相关基因 A5657, B9769 和 C7965 的多核苷酸编码的多肽接触;

[0154] b) 检测步骤 (a) 的多肽的生物活性 ;和

[0155] c) 选择与缺乏受试化合物时检测的生物活性相比能抑制多肽生物活性的化合物。

[0156] 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。

[0157] 在本发明筛选治疗细胞增殖性疾病所用化合物的方法的另一个实施方案中,所述方法利用本发明多肽的生物活性作为指标。由于本发明的 A5657, B9769 和 C7965 蛋白具有促进细胞增殖的活性,使用该活性作为指标即可筛选出能抑制其中一种蛋白质的该活性的化合物。该筛选方法包括步骤:(a) 使受试化合物与选自 A5657, B9769 和 C7965 的多肽接触 ;(b) 检测步骤 (a) 的多肽的生物活性 ;和 (c) 选择与缺乏受试化合物时检测的生物活性相比能抑制多肽生物活性的化合物。在本方法中,优选在活细胞中表达多肽,以细胞增殖为指标检测生物活性。另外,优选活细胞是被选自 A5657, B9769 和 C7965 的多核苷酸转染的宿主细胞。例如,NIH3T3 可用作宿主细胞。

[0158] 任何多肽都可用于筛选,只要它包含 A5657, B9769 和 C7965 蛋白的生物活性即可。所述生物活性包括人 A5657, B9769 和 C7965 蛋白的细胞增殖活性。例如,可使用人 A5657, B9769 和 C7965 蛋白,也可使用与这些蛋白质功能等同的多肽。所述多肽可被细胞内源性或外源地表达。

[0159] 或者,本发明的筛选方法可包括下述步骤:

[0160] 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道基因,其中一种或多种标记基因选自 BRC- 相关基因 A5657, B9769 和 C7965 ;

[0161] b) 测定所述报道基因的表达水平或活性 ;和

[0162] c) 选择与对照相比能降低所述报道基因的表达水平或活性的候选化合物。

[0163] 适当的报道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。

[0164] 可以使用任何受试化合物,例如细胞提取物、细胞培养物上清液、发酵微生物的产物、海洋生物提取物、植物提取物、纯化的或粗的蛋白质、肽、非 - 肽化合物、合成的小分子化合物和天然化合物。

[0165] 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所编码蛋白质的活性,并可用于治疗或预防乳腺癌。

[0166] 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。

[0167] 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。

[0168] 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 adenostrix* 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。

[0169] 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 Polysorbate80 (TM) 和 HCO-50 一起使用。

[0170] 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。

[0171] 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。

[0172] 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。

[0173] 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。

[0174] 评估患有乳腺癌的受试者的预后情况:

[0175] 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 BRC- 相关基因 A5657, B9769 和 C7965 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。

[0176] 例如,与正常对照相比一种或多种 BRC- 相关基因 A5657, B9769 和 C7965 表达的增加表示预后不佳。相反,与正常对照相比一种或多种 BRC- 相关基因 A5657, B9769 和 C7965 的表达类似表示受试者的预后较好。优选通过比较选自 BRC- 相关基因 A5657, B9769 和 C7965 的基因的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。



[0177] 试剂盒：

[0178] 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。

[0179] 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。

[0180] 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 BRC- 相关基因 A5657, B9769 和 C7965 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两种或多种由 BRC- 相关基因 A5657, B9769 和 C7965 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。

[0181] 微阵列和核酸混合物 (pluralities)：

[0182] 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 BRC- 相关基因 A5657, B9769 和 C7965 表示的核酸序列。通过检测与微阵列结合的核酸，即可鉴定出两种或多种由 BRC- 相关基因 A5657, B9769 和 C7965 表示的核酸的表达水平。

[0183] 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 BRC- 相关基因 A5657, B9769 和 C7965 表示的核酸。在多个实施方案中，核酸混合物包括两种或多种由 A5657, B9769 和 C7965 所列的 BRC- 相关基因表示的核酸。

[0184] 抑制乳腺癌的方法：

[0185] 本发明还提供了通过降低一种或多种上调的 BRC- 相关基因 A5657, B9769 和 C7965 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一种或多种 BRC- 相关基因 A5657, B9769 和 C7965 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。

[0186] 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。

[0187] 反义核酸：

[0188] 如上所述，可使用对应于 BRC- 相关基因 A5657, B9769 和 C7965 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 BRC- 相关基因 A5657, B9769 和 C7965 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 BRC- 相关基因 A5657, B9769 和 C7965 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 BRC- 相关基因 A5657, B9769 和 C7965 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同源性的至少为 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同源性。

[0189] 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 A5657, B9769 和 C7965 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。

[0190] 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)。

[0191] 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。

[0192] 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。

[0193] 可根据患者的身体状况适当调整本发明的反义核酸的剂量，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。

[0194] 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。

[0195] 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 BRC- 相关基因 A5657, B9769 或 C7965 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。

[0196] 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。

[0197] 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 BRC- 相关基因 A5657, B9769 和 C7965 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。

[0198] 与靶 mRNA 杂交的 A5657, B9769 或 C7965 的 siRNA 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, 能降低或抑制分别由 A5657, B9769 和 C7965 基因编码的 A5657, B9769 和 C7965 多肽的产生。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 A5657, B9769 和 C7965 siRNA 的核酸序列包括作为靶序列的 SEQ ID NO :28, 29, 30, 31, 32, 33 和 34 的核苷酸序列。另外, 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。

[0199] 可以能结合 mRNA 转录物的形式将 A5657, B9769 或 C7965 的 siRNA 直接导入细胞。或者, 可用载体承载编码 A5657, B9769 或 C7965 的 siRNA 的 DNA。

[0200] 例如, 通过以允许两条链表达 (通过转录 DNA 分子) 的方式, 将 A5657, B9769 或 C7965 靶序列克隆至表达载体中, 即可产生载体, 所述载体具有侧翼于 A5657, B9769 或 C7965 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 A5657, B9769 或 C7965 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 A5657, B9769 或 C7965 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 A5657, B9769 或 C7965 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 A5657, B9769 或 C7965 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。

[0201] 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 SEQ ID NO :28, 29, 30, 31, 32, 33 和 34 的核苷酸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.)。

[0202] 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。

[0203] 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。

[0204] 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.

[0205] 例如,下文显示了本发明优选的具有发夹环结构的 siRNA。在下述结构中,环序列可选自 CCC, UUCG, CCACC, CCACACC 和 UUCAAGA GA。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:

[0206] 对 A5657-siRNA 而言:

[0207] caucgcaacuguguugacc-[b]-ggucaacacaguugcgaug(针对 SEQ ID NO:28 的靶序列);

[0208] ugccagacaguggacagag-[b]-cucuguccacugucuggca(针对 SEQ ID NO:29 的靶序列)。

[0209] 对 B9769-siRNA 而言:

[0210] gccugcaguuccugcagca-[b]-ugcugcaggaacugcaggc(针对 SEQ ID NO:30 的靶序列);

[0211] gcuuccagucugucaaguc-[b]-gacuugacagacuggaagc(针对 SEQ ID NO:31 的靶序列);

[0212] agcagaggccucuaaugca-[b]-ugcauuagaggccucugcu(针对 SEQ ID NO:32 的靶序列)。

[0213] 对 C7965-siRNA 而言: acugcuccucucagcuucc-[b]-ggaagcugagaggagcagu(针对 SEQ ID NO:33 的靶序列); 和 guacgcuuacuggcauca-[b]-uugaugccaguaagcguac(针对 SEQ ID NO:34 的靶序列)。

[0214] 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。

[0215] 选择 siRNA 靶位点:

[0216] 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。

[0217] 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。

[0218] 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。

[0219] 侧翼于 A5657, B9769 或 C7965 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 A5657, B9769 或 C7965 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 pol III 转录单位或人 H1RNA 启动子。为了将载体导入细胞,可使用转染-增强剂。FuGENE (Roche Diagnostics), Lipofectamin 2000 (Invitrogen), Oligofectamin (Invitrogen) 和 Nucleofactor (Wako Pure Chemical) 可用作转染-增强剂。

[0220] 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。

[0221] 抗体：

[0222] 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（如 A5657, B9769 或 C7965）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。

[0223] 上文提及了抗体，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（即上调标记基因如 A5657, B9769 或 C7965 的基因产物）或与之密切相关的抗原相互作用（即结合）。如上所述，在本发明的上下文中，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，通过使用酶，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))。

[0224] 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。

[0225] 或者，抗体可包括嵌合抗体，其具有得自非人抗体的可变区与得自人抗体的恒定区，或包括人源化抗体，其含有得自非人抗体的互补决定区 (CDR)、得自人抗体的框架区 (FR) 和恒定区。使用已知技术即可制备出所述抗体。

[0226] 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)，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. Clin Cancer 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.)。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. (2002). *Oncology*, 63Suppl 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浸润的特异性靶向试剂联用。

[0227] 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。

[0228] 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。

[0229] 因此,可用于本发明上下文的治疗剂包括例如(i)过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(即改变过表达多肽与其结合配对象之间的相互作用的抑制剂、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)。

[0230] 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)。

[0231] 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。

[0232] 抗乳腺癌预防接种:

[0233] 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BRC-相关基因A5657, B9769或C7965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。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BRC-相关基因A5657, B9769或C7965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。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。

[0234] 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BRC-相关基因A5657, B9769或C7965编码的多肽或其片段被认为是受限HLA-A24或HLA-A\*0201的表位肽,它可诱导针对表达BRC-相关基因A5657, B9769或C7965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:

[0235] - 诱导抗肿瘤的细胞毒淋巴细胞,

[0236] - 诱导识别肿瘤的抗体,和

[0237] - 诱导产生抗-肿瘤的细胞因子。

[0238] 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。

[0239] 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。

[0240] 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 <sup>51</sup>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 <sup>3</sup>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。

[0241] 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。

[0242] 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。

[0243] 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。

[0244] 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。

[0245] 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。

[0246] 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。

[0247] 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。

[0248] 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。

[0249] 抑制 BRC 或恶性 BRC 的药物组合物:

[0250] 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。

[0251] 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。

[0252] 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。

[0253] 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。

[0254] 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。

[0255] 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。

[0256] 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。

[0257] 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。

[0258] 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。

[0259] 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。

[0260] 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。

[0261] 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。

[0262] 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。

[0263] 实施例

[0264] 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。

[0265] 患者和组织样品:

[0266] 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 2)。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  $-80^{\circ}\text{C}$ 。

[0267] 组织样品和 LMM :

[0268] 表 2 详细描述了有关肿瘤的临床和病理学信息。将样品包埋于 TissueTekOCT 介质 (Sakura), 然后储藏于  $-80^{\circ}\text{C}$  待用。在低温恒温器中将冷冻样品连续切成  $8\text{-}\mu\text{m}$  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 EZCut 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。

[0269] 细胞系 :

[0270] 人乳腺癌细胞系 HBC4, HBC5, MDA-MB-231, BSY-1 由 Yamori 博士惠赠 (The Japanese Foundation of Cancer Research, Tokyo), MCF7, T47D, SKBR3, HCC1937, MDA-MB-435S, YMB1, HBL100, COS7, NIH3T3 得自 ATCC。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, SKBR3, T47D, YMB1 和 HCC1937 (含  $2\text{mM}$  L-谷氨酰胺); Dulbecco 改进的 Eagle 培养基 (Invitrogen, Carlsbad, CA), 它可用于 HBL100, COS7, NIH3T3; 含  $0.1\text{mM}$  必需氨基酸 (Roche),  $1\text{mM}$  丙酮酸钠 (Roche),  $0.01\text{mg/ml}$  胰岛素 (Sigma) 的 EMEM (Sigma), 它可用于 MCF7; L-15 (Roche), 它可用于 MDA-MB-231 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  $37^{\circ}\text{C}$ , 在不含  $\text{CO}_2$  的湿润空气的气氛中维持 MDA-MB-231 和 MDA-MB-435S 细胞。于  $37^{\circ}\text{C}$ , 在含 5%  $\text{CO}_2$  的湿润空气的气氛中维持其它细胞系。

[0271] RNA 提取和基于 T7 的 RNA 扩增 :

[0272] 从每个经激光捕获的细胞群体中提取总 RNA, 溶于  $350\mu\text{l}$  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  $70^{\circ}\text{C}$  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  $28.8\text{-}329.4\mu\text{g}$  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  $2240.2\mu\text{g}$  和  $2023.8\mu\text{g}$ 。分别在 Cy5-dCTP 和 Cy3-dCTP (Amersham Biosciences) 的存在下逆转录  $2.5\mu\text{g}$  得自每种癌细胞和非癌乳腺导管细胞的 aRNA 等份试样。

[0273] cDNA 微阵列 :

[0274] 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 (AmershamBiosciences) 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。

[0275] 杂交和数据获取:

[0276] 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 ArrayVision 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,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。

[0277] 鉴定和分离新的人基因 A5657, B9769 和 C7965:

[0278] 按上文讨论的方法提取和扩增总 RNA。通过分别用 Cy5-dCTP 和 Cy3-dCTP (Amersham Biosciences, Buckinghamshire, UK) 逆转录来标记得自乳腺癌细胞和正常乳腺导管细胞的经扩增 RNA 的等份试样。按上文所述进行杂交、洗涤和检测。为了检测通常在乳腺癌中被上调的基因,在微阵列上筛选 23,040 个基因的整体表达模式,以选择出分别存在于 >50% 的 i) 所有 81 个乳腺癌病例, ii) 69 例浸润性导管癌, iii) 31 例高度-, iv) 14 例中度-, 或 v) 24 例低度-分化的损害中的表达率 >3.0 的基因。在肿瘤细胞中显示出被上调的总共 102 个基因中,选择 3 个内部鉴定号为 A5657, B9769 和 C7965 的基因进行进一步检查,因为在大于 50% 被确认的乳腺癌病例中,它们的表达率大于 3.0。

[0279] 半-定量 RT-PCR:

[0280] 选择上述 3 个上调基因 A5657, B9769 和 C7965, 通过使用半-定量 RT-PCR 实验来检查其表达水平。具体地说,按上述方法提取、扩增和逆转录总 RNA。通过监测甘油醛-3-磷酸脱氢酶 (GAPDH) 作为定量内部对照,对每个单链 cDNA 进行适当稀释,随后进行 PCR 扩增。PCR 引物序列为:

[0281] 5' -CGACCACTTTGTCAAGCTCA-3' (SEQ ID No. 7) 和

[0282] 5' -GGTTGAGCACAGGGTACTTTATT-3' (SEQ ID No. 8) 用于 GAPDH;

[0283] 5' -CAAATATTAGGTGGAGCCAACAC-3' (SEQ ID No. 9) 和

[0284] 5' -TAGATCACCTTGGCAAAGAACAC-3' (SEQ ID No. 10) 用于 A5657,

[0285] 5' -ACCTCAAGTCCCTCCTGGAA-3' (SEQ ID No. 11) 和

[0286] 5' -TCAGTTTCAACAGGTAAGGCGAT-3' (SEQ ID No. 12) 用于 B9769,

[0287] 5'-AGAGCCATAGAAACTGCTCCTCT-3' (SEQ ID No. 13) 和

[0288] 5'-CATAACTGCATAGACAGCAGTC-3' (SEQ ID No. 14) 用于 C7965。

[0289] Northern-印迹分析：

[0290] 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 A5657, B9769 和 C7965 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。通过使用在半-定量 RT-PCR 部分提及的引物套进行 PCR, 制备出针对 A5657 (541bp), B9769 (499bp) 和 C7965 (238bp) 的特异性探针;

[0291] 5'-CAAATATTAGGTGGAGCCAACAC-3' (SEQ ID No. 9) 和

[0292] 5'-TAGATCACCTTGGCAAAGAACAC-3' (SEQ ID No. 10) 用于 A5657,

[0293] 5'-ACCTCAAGTCCCTCCTGGAA-3' (SEQ ID No. 11) 和

[0294] 5'-TCAGTTTCAACAGGTAAGGCGAT-3' (SEQ ID No. 12) 用于 B9769,

[0295] 5'-GGGAAGAGAAGTCCCGAGTC-3' (SEQ ID No. 15) 和

[0296] 5'-TCCTTATTCTGAATTTCCAGAATC-3' (SEQ ID No. 16) 用于 C7965。

[0297] 快速扩增 cDNA 末端 (RACE)-PCR：

[0298] 为了检测 C7965 的全长转录物,根据厂商说明用 Marathon cDNA 扩增试剂盒 (BD, CLONTECH) 进行 5' RACE。使用制备自乳腺癌细胞系 MCF7 的 cDNA 为模板,用反向引物 (5' -CAAGCAGTCCTACCAGGGTTCGGAAGCTGA-3' ) (SEQ ID No. 17) 进行 1<sup>st</sup>PCR, 检测到多个产物带。在下述条件下进行 1<sup>st</sup>5' RACE-PCR: 起初于 94°C 变性 1 分钟; 94°C, 30 秒, 68°C, 2 分钟, 35 轮循环; 最后于 68°C 延伸 7 分钟。稀释 1<sup>st</sup>-PCR 产物之后, 通过用嵌套反向引物 (5' -CCAGGGTTCGGAAGCTGAGAGGAGCA GTTT-3' ) (SEQ ID No. 18) 进行嵌套式 PCR 来检测强的产物带。在下述条件下进行嵌套式 -PCR: 起初于 94°C 变性 1 分钟; 94°C, 30 秒, 72°C, 2 分钟, 2 轮循环, 94°C, 30 秒, 70°C, 2 分钟, 2 轮循环, 94°C, 30 秒, 68°C, 2 分钟, 15 轮循环; 最后于 68°C 延伸 7 分钟。根据厂商说明使用 Gel 提取试剂盒 (Qiagen) 和 TOPO TA 克隆试剂盒 (Invitrogen) 鉴定序列。

[0299] 构建表达载体：

[0300] 使用 KOD-Plus DNA 聚合酶 (Toyobo, Osaka, Japan) 和下述引物, 通过 PCR 扩增 A5657, B9769 和 C7965cDNA 的完整编码序列：

[0301] A5657 正向; 5'-CCGGAATTCATGCAGAGAGCTTCACGTCTG-3' (SEQ ID No. 19) 和

[0302] A5657 反向; 5'-CCGCTCGAGAACATCAGGATGAAATTTCTTTTC-3' (SEQ ID No. 20),

[0303] B9769- 正向; 5'-CCGGAATTCATGAGCGGTGCGGGGTGGCG-3' (SEQ ID No. 21) 和

[0304] B9769- 反向; 5'-CCGCTCGAGAAGCACTGAGCGATGCAGGCG-3' (SEQ ID No. 22) 和

[0305] C7965- 正向; 5'-CCGGAATTCATGGACGCAGAGCTGGCAGAGGTG

[0306] CG-3' (SEQ ID No. 23) 和

[0307] C7965- 反向 ;5' -CCGCTCGAGGTTGTTCTCCTCTGCACAAAG-3' (SEQID No. 24)。将 PCR 产物分别插入 pCAGGS-3xFLAGn 和 pCAGGSnHA, pcDNA3. 1(+)-Myc/His (Invitrogen) 表达载体的 EcoRI 和 XhoI 位点。

[0308] Kohei Miyazono 博士 (The Cancer Institute of Japanese Foundation for Cancer Research) 惠赠了遍在蛋白结合试验所用的 HA- 遍在蛋白表达载体 (pcdef3-HA-RPS27A)。通过 DNA 测序进一步证实这些构建体 (pCAGGS-A5657-HA, pcDNA3. 1-A5657-Myc/His, pCAGGS-3xFLAG-A5657 和 pcdef3-HA-RPS27A(Ub80a))。为了进行细胞生长促进试验, Motoko Unoki 博士构建了突变 H-ras (G12K) 的完整编码序列。

[0309] 免疫细胞化学染色:

[0310] 为了检查 A5657, B9769 和 C7965 蛋白的亚细胞定位, 以每孔  $5 \times 10^4$  个细胞的密度接种 T47D 细胞 (对 A5657 和 B9769 而言), 以每孔  $1 \times 10^4$  (低密度) 和  $1 \times 10^5$  (高密度) 个细胞的密度接种 COS7 细胞 (对 B9769 和 C7965 而言)。24 小时之后, 根据厂商说明分别用 FuGENE6 转染试剂 (Roche) 将  $1 \mu$  gpCAGGS-A5657-HA 瞬时转染至 T47D 细胞。然后, 用含有 4% 低聚甲醛的 PBS 固定细胞 15 分钟, 于  $4^\circ\text{C}$  用含有 0.1% Triton X-100 的 PBS 透化细胞 2.5 分钟。随后, 于  $4^\circ\text{C}$  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, 分别使经 A5657-HA- 转染的 T47D 细胞与稀释度为 1:1000 的小鼠抗-HA 抗体 (SANTA CRUZ) 或稀释度为 1:1000 的小鼠抗-myc 抗体 (Sigma) 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。另外, 比较 B9769 和其它细胞骨架蛋白的亚细胞定位。用  $1 \mu$  gpcDNA3. 1(+)-B9769-myc-his 转染之后, 使细胞与稀释度为 1:1000 的兔抗-myc 抗体 (SANTACRUZ) 或稀释度为 1:500 的小鼠抗- $\beta$  微管蛋白抗体 (SIGMA) 一起保温。用 PBS 洗涤之后, 用稀释度为 1:5000 的与 Alexa488- 偶联的抗-兔第二抗体 (Molecular Probe), 与 Alexa594- 偶联的抗-小鼠第二抗体 (Molecular Probe), 稀释度为 1:50 的与 Alexa594- 偶联的鬼笔环肽染色经转染的 T47D 细胞。另外, 为了检查 B9769 在低密度或高密度细胞中的亚细胞定位, 用  $1 \mu$  gpCAGGSn3F-B9769-HA 转染 COS7 细胞, 用与上述相同的方法进行免疫细胞化学染色。使经转染的细胞与稀释度为 1:1000 的小鼠抗-HA 抗体 (SANTACRUZ) 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 COS7 细胞。用 4',6'-二脒-2'-苯基吡啶二盐酸化物 (DAPI) 复染细胞核。另外, 为了检查 C7965 在低密度或高密度细胞中的亚细胞定位, 用  $1 \mu$  gpcDNA3. 1(+)-C7965-Myc/His 转染 COS7 细胞, 用与上述相同的方法进行免疫细胞化学染色。使经转染的细胞与稀释度为 1:1000 的小鼠抗-myc 抗体 (SANTACRUZ) 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 COS7 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 SP2A0BS 显微镜 (Leica, Tokyo, Japan) 下获得荧光图像。

[0311] 使用 psiU6X3.0 构建 A5657, B9769 和 C7965 特异性的-siRNA 表达载体:

[0312] 使用文献中所述的 psiU6BX 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 Y. M. Nakamura Y. Involvement of the FGF18 Gene in Colorectal Carcinogenesis, as a Novel Downstream Target of the  $\beta$ -Catenin/T-Cell Factor Complex63, Cancer Res., 63, 6116-20, 2003)。通过

将双链寡核苷酸克隆至 psiU6BX 载体的 BbsI 位点, 制备出针对 A5657 (psiU6BX-A5657), B9769 (psiU6BX-B9769) 和 C7965 (psiU6BX-C7965) 的 siRNA 表达载体。通过将 5' -CACCGA AGCAGCAGACTTCTTCTTCAAGAGAGAAGAAGTCGTGCTGCTTC-3' (SEQ ID No. 25) 和 5' -AAAAGAAGCAG CACGACTTCTTCTCTCTTGAAGAAGAAGTCGTGCTGCTTC-3' (SEQ ID No. 26) 的双链寡核苷酸克隆至 psiU6BX3.0 载体的 BbsI 位点, 制备出对照质粒 psiU6BX-EGFP。

[0313] 表 1

[0314]

	靶序列	SEQ ID No.
siEGFP	5'-GAAGCAGCAGACTTCTT-3'	27
A5657 si2	5'-CATCGCAACTGTGTTGACC-3'	28
A5657 si3	5'-TGCCAGACAGTGGACAGAG-3'	29
B9769 si1	5'-GCCTGCAGTTCCTGCAGCA-3'	30
B9769 si2	5'-GCTTCCAGTCTGTCAAGTC-3'	31
B9769 si4	5'-AGCAGAGGCCTCTAATGCA-3'	32
C7965 si1	5'-ACTGCTCCTCTCAGCTTCC-3'	33
C7965 si3	5'-GTACGCTTACTGGCATCAA-3'	34

[0315] A5657, B9769 和 C7965 的基因 - 沉默作用:

[0316] 将人乳腺癌细胞系 T47D 铺于 10-cm 平皿中 ( $1 \times 10^6$  个细胞 / 皿), 根据厂商推荐的方法, 使用 FuGENE6 试剂 (Roche) 将 psiU6BX-EGFP (作为阴性对照), psiU6BX-A5657 转染至 T47D 细胞。转染 7 天后, 从细胞中提取总 RNA, 然后使用上文所述的针对 A5657, B9769 和 C7965 以及 GAPDH 的特异性引物进行半 - 定量 RT-PCR 以进一步证实 siRNA 的击倒作用。另外, 在含有 0.7mg/ml 新霉素的选择培养基中将使用 T47D 细胞系的能表达 siRNA 的转染子培养 28 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 7 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 (0.01NHCl/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 Reader550 (BioRad) 测定 570nm 下的光吸收值。

[0317] 细胞增殖试验:

[0318] 以每个 150mm 平皿  $3 \times 10^6$  个细胞的密度接种 NIH3T3 细胞。24 小时之后, 使用 FuGENE6 (Roche) 分别将  $16 \mu$  gpcDNA3.1(+)-Myc/His (Invitrogen) (阴性对照), pcDNA3.1(+)-A5657-Myc/His, pcDNA3.1(+)-B9769-Myc/His, pcDNA3.1(+)-C7965 或 pcDNA3.1(+)-H-Ras-mutant-Myc/His (阳性对照) 瞬时转染至 NIH3T3 细胞。接着, 在转染 24 小时之后, 在每孔中重新接种  $2 \times 10^4$  个 NIH3T3 细胞。在重新接种后的第 1, 2, 4 和 6 天, 按上文所述进行 MTT 试验。

[0319] 遍在蛋白结合分析:

[0320] 以每孔  $5 \times 10^5$  个细胞的密度接种 COS7 细胞。24 小时之后, 使用 FuGENE6 (Roche)

将 1  $\mu$ g 经 FLAG 标记的 A5657 和 0.1  $\mu$ g 经 HA 标记的遍在蛋白表达载体 (Ub80a-HA) 共转染至 COS7 细胞。根据 Imamura 等和 Ebisawa 等的方法构建经 HA 标记的遍在蛋白表达载体 (Imamura et al., Nature 389/6651, 622-6, 1997 “Smad6 inhibits signalling by the TGF- $\beta$  super family”; Ebisawa et al., J Biol Chem. 20:276/16, 12477-80, 2001 “Smurf1 interacts with transforming growth factor- $\beta$  type I receptor through Smad7 and induces receptor degradation”)。转染 24 小时之后, 在细胞中加入 5  $\mu$ M 蛋白酶体抑制剂 MG132 或 5  $\mu$ M DMSO, 6 小时后, 用 1ml 1.0% NP40 缓冲液 (1.0% NP40, 50mM Tris-HCl (pH8.0), 150mM NaCl 和全蛋白酶-抑制剂的混合物) 裂解细胞。用与小鼠单克隆抗-FLAG M2-抗体或小鼠抗-HA 抗体偶联的琼脂糖免疫沉淀细胞裂解物。用裂解缓冲液洗涤 6 轮之后, 用 Flag-M2 或 HA 肽 (SIGMA) 洗脱结合蛋白质。在 Western 印迹上使用兔抗-HA 多克隆抗体或兔抗-Flag 多克隆抗体分析免疫沉淀物。将样品缓冲液中的可溶性蛋白质上样于 12% SDS-聚丙烯酰胺凝胶, 并转移至硝酸纤维素膜 (Hybond ECL<sup>®</sup>)。用处于含 0.05% 吐温 20 (TBST) 的 Tris-缓冲盐水中的 BlockAce 粉末 (Yukijirushi, Tokyo, Japan) 封闭膜, 室温下与每种抗体保温 1 小时。然后, 使印迹与 HRP-偶联的第二抗体 (Amersham Biosciences) 杂交, 并使用 ECL 法 (Amersham Biosciences) 进行检测。

#### [0321] 结果

[0322] 将 A5657, B9769 和 C7965 鉴定为乳腺癌细胞中的上调基因:

[0323] 当使用含有 23,040 个人基因的 cDNA 微阵列分析 81 名乳腺癌患者的癌细胞的基因-表达分布图时, 在乳腺癌细胞中鉴定出 102 个通常被上调的基因。从这些上调基因中选择出内部编号如下的基因: A5657, 被称为类似于遍在蛋白缀合酶的 HSPC150 蛋白 (Genbank 登录号为 NM\_014176) (SEQ ID NO:1); B9769, 被称为假拟蛋白 BC016861 (Genbank 登录号为 NM\_138770) (SEQ ID NO:3); 和 C7965, 对应于 EST (SEQ ID NO:5)。在微阵列上, 与正常乳腺导管细胞相比较, 分别评价 49 个乳腺癌细胞中的 38 个, 73 个乳腺癌细胞中的 30 个和 49 个乳腺癌细胞中的 28 个的 A5657, B9769 和 C7965 基因表达。为了进一步证实我们所关注的这些上调基因的表达, 进行半-定量 RT-PCR 以在乳腺癌细胞和包括正常乳腺细胞的正常人组织之间比较这些基因的表达。首先, 我们发现与正常乳腺导管细胞和一些其它正常组织相比, 在 12 例临床乳腺癌样品 (低度分化型) 中的 9 例中, 显示出增加的 A5657 表达, A5657 在所有 6 个乳腺癌细胞系中都过表达 (图 1a)。接着, 我们发现与正常乳腺导管细胞相比, 在 12 例临床乳腺癌样品 (低度分化型) 中的 6 例中, 显示出增加的 B9769 表达, B9769 在 20 个乳腺癌细胞系中的 4 个中过表达 (图 1b)。最后, 我们揭示出与正常乳腺导管细胞相比, 在 12 例临床乳腺癌样品 (高度分化型) 中的 7 例中, 显示出增加的 C7965 表达, C7965 在 20 个乳腺癌细胞系中的 15 个中过表达 (图 1c)。

[0324] 为了进一步检查这些上调基因的表达模式, 以 A5657, B9769 和 C7965 的 cDNA 片段为探针 (见上文), 用多个人组织和乳腺癌细胞系进行 Northern 印迹分析。结果发现除了肺、肝脏、胰腺和外周血白细胞外, A5657 被遍在表达 (图 2a; 上方的一组试验), 然而, 令人惊奇的是, 与其它正常组织, 特别是心脏和骨髓相比, 所有乳腺癌细胞系中的 A5657 皆被过表达 (图 2a; 下方的一组试验), 如图 2a 上方的一组试验所示, 每一种正常组织皆显示出强信号。B9769 只在睾丸和前列腺中表达 (图 2b; 上方的一组试验), 与其它正常组织, 特别是正常人乳腺相比, 一些乳腺癌细胞系中的 B9769 被过表达 (图 2b; 下方的一组试验)。当

以外显子 2 和 3 中的 C7965 片段为探针,使用多个组织和乳腺癌细胞系进行 Northern 印迹分析时,观察到约 1.35kb 的转录物。与包括乳腺组织的正常组织相比,乳腺癌细胞系中的 1.35kb 转录物被特异性表达(图 2c;下方的一组试验),睾丸、骨骼肌和小肠仅微弱表达所述转录物(图 2c;上方的一组试验)。

[0325] A5657, B9769 和 C7965 的基因组结构:

[0326] 为了获得 A5657, B9769 和 C7965 的完整 cDNA 序列,使用制备自乳腺癌细胞系和 T47D 的 cDNA 为模板进行 RT-PCR。被称为类似于遍在蛋白缀合酶的 HSPC150 蛋白的 A5657 由 7 个外显子组成,其位于在基因组中跨越约 10.3kb 的染色体 1q32.1 上。A5657 的全长 cDNA 序列含有 928 个核苷酸。开放读框 (ORF) 始于外显子 2,终止于外显子 7。最后,该转录物编码 197 个氨基酸。

[0327] B9769 由 8 个外显子组成,其位于在基因组中跨越约 5.7kb 的染色体 2q21.2 上。B9769 的全长 cDNA 序列含有 1472 个核苷酸。ORF 始于外显子 1,终止于外显子 8。最后,该转录物编码 378 个氨基酸。

[0328] 为了进一步分离 C7965 转录物的 5' 末端,使用制备自乳腺癌细胞系 MCF7 的 cDNA 为模板进行 cDNA 末端的快速扩增 (RACE) (参见材料和方法)。分离出由 8 个外显子组成的转录物,其分别对应于 Northern 印迹中的 1.35kb (图 2c,下方的一组试验),所述转录物位于在基因组中跨越约 28.8kb 的染色体 9q 上。C7965 的全长 cDNA 序列含有 1315 个核苷酸。C7965cDNA 的 ORF 始于外显子 1,终止于外显子 8。最后,该转录物编码 288 个氨基酸。

[0329] A5657, B9769 和 C7965 的亚细胞定位:

[0330] 为了进一步检查 A5657, B9769 和 C7965 的鉴定,在哺乳动物细胞中检查这些基因产物的亚细胞定位。首先,当我们将表达 A5657 蛋白的质粒 (pCAGGS-A5657-HA) 瞬时转染至 T47D 细胞时,免疫细胞化学染色揭示出外源性 A5657 遍布于约 80% 经转染 T47D 细胞的胞质以及其余约 20% 细胞的细胞核上 (图 3a)。

[0331] 接着,当将表达 B9769 蛋白的质粒 (pCAGGS-Flag-B9769-HA) 瞬时转染至 COS7 细胞时,免疫细胞化学染色揭示出 B9769 蛋白在细胞数目较少 (低密度) 时作为中间丝状网络位于细胞质器上,当细胞数目较多 (高密度) 时位于胞质上 (图 3b),这表明 B9769 对细胞与细胞之间的相互作用起着关键作用。为了更详细地进一步检查该蛋白质的定位情况,通过免疫细胞化学染色比较其它细胞骨架蛋白的定位。结果, B9769 和其它细胞骨架蛋白不象 T47D 细胞中的丝状网络一样共同位于细胞质器上 (图 3c)。

[0332] 最后,将表达 C7965 蛋白的质粒 (pcDNA3.1(+)-C7965-Myc/His) 瞬时转染至 COS7 细胞,免疫细胞化学染色揭示出 C7965 蛋白位于 COS7 细胞的细胞质器上 (图 3d)。

[0333] 被设计用于降低 A5657, B9769 和 C7965 表达的小的 - 干扰 RNA (siRNA) 的生长 - 抑制作用:

[0334] 为了评估 A5657, B9769 和 C7965 的生长 - 促进作用,利用基于哺乳动物载体的 RNA 干扰 (RNAi) 技术 (参见材料和方法) 击倒乳腺癌细胞系 T47D (已在此细胞系中证实了 A5657 的过表达) 中内源性 A5657 和 B9769 的表达 (图 4, 5 和 6)。通过半 - 定量 RT-PCR 实验检查 A5657, B9769 和 C7965 的表达水平。如图 4, 5 和 6 所示,与对照 siRNA 构建体 (psiU6BX-EGFP) 相比, A5657 (si2 和 si3), B9769 (si1 和 si2, si4), 和 C7965 (si1 和 si3) - 特异性 siRNA 显著抑制了每种基因的表达。为了进一步证实 A5657, B9769 和 C7965 - 特异性



siRNA 的细胞生长抑制作用,分别进行集落 - 形成和 MTT 试验。结果,导入 A5657 (si2 和 si3) (图 4b, c), B9769 (si1 和 si2, si4) (图 5b, c), 和 C7965 (si1 和 si3) (图 6b, c) 构建体可抑制 T47D 细胞的生长,这与上述的表达有所降低的结果一致。通过三个独立的实验证实每个结果。因此,这些发现表明 A5657, B9769 和 C7965 在乳腺癌的细胞生长中具有显著作用。

[0335] 在 NIH3T3 细胞中瞬时过表达 A5657, B9769 和 C7965 的生长促进作用:

[0336] 为了进一步检查 A5657, B9769 和 C7965 在细胞生长调节中可能起到的作用,分别将表达 A5657, B9769, C7965, H-ras 突变体 (用作阳性对照) 或 Mock (用作阴性对照) 的质粒瞬时转染至 NIH3T3 细胞,然后进行 MTT 试验 (图 7)。转染 6 天后,与用对照质粒 (空白载体) 转染的 NIH3T3 细胞相比, A5657 (图 7a), B9769 (图 7b) 和 C7965 (图 7c) 的过表达以及 H-ras 突变体的过表达显著增强了细胞生长 (超过 3 倍),这表明 A5657, B9769 和 C7965 可能在细胞生长的增殖中起作用。通过两个独立的实验进一步证实这些结果。

[0337] A5657 蛋白结合遍在蛋白:

[0338] 为了更详细地探查 A5657 的功能,使用被表达 A5657 的质粒转染的 COS7 细胞的全细胞裂解物进行 Western 印迹分析。结果,除了显示出预期分子量的条带外,仅观察到一条额外的迁移较慢的条带 (图 8; 左边的一组试验,全细胞裂解物),暗示着这可能是由翻译后修饰产生的。SMART 程序推测编码 197 个氨基酸的 A5657 可能含有 UBCc 结构域 (遍在蛋白 - 缀合酶 E2, 催化结构域同系物) (5-152 个残基), 这表明 A5657 可能具有潜在的 E2 遍在蛋白酶活性。为了研究这条多出的条带是否由单 - 遍在蛋白化引起,将经设计用于表达经 FLAG 标记的 A5657 (A5657-FLAG) 和经 HA 标记的遍在蛋白 (UB80a-HA) 的质粒 DNA 共转染至 COS7 细胞。使用抗 -FLAG 抗体免疫沉淀 A5657, 用抗 -HA 抗体可检测到沉淀物。观察到一条额外的迁移较慢的条带,这表明 A5657 结合遍在蛋白 (图 8)。另外,当使用抗 -HA 抗体进行免疫沉淀,然后用抗 -FLAG 抗体进行 Western 印迹分析时,也观察到一条额外的迁移较慢的条带 (图 8)。用蛋白酶体抑制剂 MG132 处理对所述结合不起作用。该发现强烈暗示着 A5657 蛋白可能具有经由单 - 遍在蛋白化的 E2 遍在蛋白酶活性。

[0339] 讨论

[0340] 本文通过利用全基因组 cDNA 微阵列测得的乳腺癌精确表达分布图分离出新的基因 A5657, B9769 和 C7965, 与正常人组织相比,乳腺癌细胞中的所述基因被显著过表达。

[0341] 其中, A5657 含有 UBCc 结构域 (遍在蛋白 - 缀合酶 E2, 催化结构域同系物), 因此可结合遍在蛋白。这一发现强烈暗示着 A5657 蛋白具有经由单 - 遍在蛋白化的 E2 遍在蛋白酶活性, 因此可能参与乳腺癌的肿瘤发生。

[0342] 通过免疫化学染色观察到 B9769 蛋白作为中间丝状网络位于胞质中。在细胞群体密度低的条件下,进一步观察到 B9769 蛋白作为中间丝状网络位于细胞质器中,但在细胞群体密度高的条件下,则位于胞质中,这表明 B9769 可能在细胞与细胞的相互作用中起着关键作用。

[0343] 如本文所阐明的那样,用 siRNA 处理乳腺癌细胞可有效抑制所有 3 个靶基因 A5657, B9769 和 C7965 的表达,反过来即可显著抑制乳腺癌细胞 - 肿瘤生长。另外,已在 MTT 试验中证实:当在 NIH3T3 细胞中显著地瞬时过表达这些基因时,所述基因可促进细胞增殖。这些发现暗示着 A5657, B9769 和 C7965 在肿瘤细胞生长增殖中起着关键作用,并且

是开发抗癌药物的理想靶标。

[0344] 表 2 组织临床信息

[0345]

ID	手术 年龄	经绝 状况	T	N	M	病期	组织学 类型	淋巴细 胞浸润	血管侵 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	-	-
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	+	+
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	Ia	0	0	-	+
MMK010613	37	pre	0	0	0	0	Ia	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	Ia	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	Ia	0	0	+	+
MMK010680	58	post	0	0	0	0	Ia	0	0	-	+
MMK010709	33	pre	2	0	0	2	a3	0	2	-	-
MMK010711	51	pre	0	0	0	0	Ia	0	0	-	+
MMK010724	40	pre	2	1	0	2	a3	3	2	+	+
MMK010744	41	pre	0	0	0	0	Ia	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	Ia	0	0	+	+
MMK010846	47	pre	2	0	0	2	a1	0	0	+	+
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	+	+

[0346] 工业实用性

[0347] 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。

[0348] 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。

[0349] 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。

[0350] 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-A0306P1

<160> 34

<170> PatentIn version 3.1

<210> 1

<211> 928

<212> DNA

<213> 人 (Homo sapiens)

<220>

<221> CDS

<222> (127).. (720)

<223>

<400> 1

gcgcgcagcgc ctggtacccc gttgggtccgc gcgttgctgc gttgtgaggg gtgtcagctc 60

agtgcateccc aggcagctct tagtgtggag cagtgaactg tgtgtggttc cttctacttg 120

gggatac atg cag aga gct tca cgt ctg aag aga gag ctg cac atg tta 168

Met Gln Arg Ala Ser Arg Leu Lys Arg Glu Leu His Met Leu

1

5

10

gcc aca gag cca ccc cca ggc atc aca tgt tgg caa gat aaa gac caa 216

Ala Thr Glu Pro Pro Pro Gly Ile Thr Cys Trp Gln Asp Lys Asp Gln

15 20 25 30

atg gat gac ctg cga gct caa ata tta ggt gga gcc aac aca cct tat 264

Met Asp Asp Leu Arg Ala Gln Ile Leu Gly Gly Ala Asn Thr Pro Tyr

35

40

45

gag aaa ggt gtt ttt aag cta gaa gtt atc att cct gag agg tac cca 312

Glu Lys Gly Val Phe Lys Leu Glu Val Ile Ile Pro Glu Arg Tyr Pro

50

55

60

ttt gaa cct cct cag atc cga ttt ctc act cca att tat cat cca aac 360

Phe Glu Pro Pro Gln Ile Arg Phe Leu Thr Pro Ile Tyr His Pro Asn

65

70

75

att gat tct gct gga agg att tgt ctg gat gtt ctc aaa ttg cca cca 408

Ile Asp Ser Ala Gly Arg Ile Cys Leu Asp Val Leu Lys Leu Pro Pro	
80	85
90	
aaa ggt gct tgg aga cca tcc ctc aac atc gca act gtg ttg acc tct	456
Lys Gly Ala Trp Arg Pro Ser Leu Asn Ile Ala Thr Val Leu Thr Ser	
95	100
105	110
att cag ctg ctc atg tca gaa ccc aac cct gat gac ccg ctc atg gct	504
Ile Gln Leu Leu Met Ser Glu Pro Asn Pro Asp Asp Pro Leu Met Ala	
115	120
125	
gac ata tcc tca gaa ttt aaa tat aat aag cca gcc ttc ctc aag aat	552
Asp Ile Ser Ser Glu Phe Lys Tyr Asn Lys Pro Ala Phe Leu Lys Asn	
130	135
140	
gcc aga cag tgg aca gag aag cat gca aga cag aaa caa aag gct gat	600
Ala Arg Gln Trp Thr Glu Lys His Ala Arg Gln Lys Gln Lys Ala Asp	
145	150
155	
gag gaa gag atg ctt gat aat cta cca gag gct ggt gac tcc aga gta	648
Glu Glu Glu Met Leu Asp Asn Leu Pro Glu Ala Gly Asp Ser Arg Val	
160	165
170	
cac aac tca aca cag aaa agg aag gcc agt cag cta gta ggc ata gaa	696
His Asn Ser Thr Gln Lys Arg Lys Ala Ser Gln Leu Val Gly Ile Glu	
175	180
185	190
aag aaa ttt cat cct gat gtt tag gggacttgtc ctggttcac ttagttaatg	750
Lys Lys Phe His Pro Asp Val	
195	
tgttctttgc caaggtgac taagttgect accttgaatt tttttttaa tatattgat	810
gacataattt ttgtgtagtt tatttatctt gtacatatgt attttgaaat cttttaaacc	870
tgaaaaataa atagtcattt aatgttgaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaa	928

&lt;210&gt;2

&lt;211&gt;197

&lt;212&gt;PRT

&lt;213&gt; 人 (Homo sapiens)

&lt;400&gt;2

Met Gln Arg Ala Ser Arg Leu Lys Arg Glu Leu His Met Leu Ala Thr	
1	5
10	15
Glu Pro Pro Pro Gly Ile Thr Cys Trp Gln Asp Lys Asp Gln Met Asp	
20	25
30	
Asp Leu Arg Ala Gln Ile Leu Gly Gly Ala Asn Thr Pro Tyr Glu Lys	

35	40	45	
Gly Val Phe Lys Leu Glu Val Ile Ile Pro Glu Arg Tyr Pro Phe Glu			
50	55	60	
Pro Pro Gln Ile Arg Phe Leu Thr Pro Ile Tyr His Pro Asn Ile Asp			
65	70	75	80
Ser Ala Gly Arg Ile Cys Leu Asp Val Leu Lys Leu Pro Pro Lys Gly			
85	90	95	
Ala Trp Arg Pro Ser Leu Asn Ile Ala Thr Val Leu Thr Ser Ile Gln			
100	105	110	
Leu Leu Met Ser Glu Pro Asn Pro Asp Asp Pro Leu Met Ala Asp Ile			
115	120	125	
Ser Ser Glu Phe Lys Tyr Asn Lys Pro Ala Phe Leu Lys Asn Ala Arg			
130	135	140	
Gln Trp Thr Glu Lys His Ala Arg Gln Lys Gln Lys Ala Asp Glu Glu			
145	150	155	160
Glu Met Leu Asp Asn Leu Pro Glu Ala Gly Asp Ser Arg Val His Asn			
165	170	175	
Ser Thr Gln Lys Arg Lys Ala Ser Gln Leu Val Gly Ile Glu Lys Lys			
180	185	190	
Phe His Pro Asp Val			
195			

&lt;210&gt;3

&lt;211&gt;1472

&lt;212&gt;DNA

&lt;213&gt; 人 (Homo sapiens)

&lt;220&gt;

&lt;221&gt;CDS

&lt;222&gt;(53)..(1189)

&lt;223&gt;

&lt;400&gt;3

ggccactgag ccggggtgca gtggcagcgg gagagtacct ggcgatggcg at atg agc	58	
	Met Ser	
	1	
ggt gcg ggg gtg gcg gct ggg acg cgg ccc ccc agc tcg ccg acc ccg	106	
Gly Ala Gly Val Ala Ala Gly Thr Arg Pro Pro Ser Ser Pro Thr Pro		
5	10	15
ggc tct cgg cgc cgg cgc cag cgc ccc tct gtg ggc gtc cag tcc ttg	154	

Gly Ser Arg Arg Arg Arg Gln Arg Pro Ser Val Gly Val Gln Ser Leu			
20	25		
agg ccg cag agc ccg cag ctc agg cag agc gac ccg cag aaa cgg aac	202		
Arg Pro Gln Ser Pro Gln Leu Arg Gln Ser Asp Pro Gln Lys Arg Asn			
35	40	45	50
ctg gac ctg gag aaa agc ctg cag ttc ctg cag cag cag cac tcg gag	250		
Leu Asp Leu Glu Lys Ser Leu Gln Phe Leu Gln Gln Gln His Ser Glu			
55	60	65	
atg ctg gcc aag ctc cat gag gag atc gag cat ctg aag cgg gaa aac	298		
Met Leu Ala Lys Leu His Glu Glu Ile Glu His Leu Lys Arg Glu Asn			
70	75	80	
aag gat ctc cat tac aag ctc ata atg aat cag aca tca cag aag aaa	346		
Lys Asp Leu His Tyr Lys Leu Ile Met Asn Gln Thr Ser Gln Lys Lys			
85	90	95	
gat ggc ccc tca gga aac cac ctt tcc agg gcc tct gct ccc ttg ggc	394		
Asp Gly Pro Ser Gly Asn His Leu Ser Arg Ala Ser Ala Pro Leu Gly			
100	105	110	
gct cgc tgg gtc tgc atc aac gga gtg tgg gta gag ccg gga gga ccc	442		
Ala Arg Trp Val Cys Ile Asn Gly Val Trp Val Glu Pro Gly Gly Pro			
115	120	125	130
agc cct gcc agg ctg aag gag ggc tcc tca cgg aca cac agg cca gga	490		
Ser Pro Ala Arg Leu Lys Glu Gly Ser Ser Arg Thr His Arg Pro Gly			
135	140	145	
ggc aag cgt ggg cgt ctt gcg ggc ggt agc gcc gac act gtg cgc tct	538		
Gly Lys Arg Gly Arg Leu Ala Gly Gly Ser Ala Asp Thr Val Arg Ser			
150	155	160	
cct gca gac agc ctc tcc atg tca agc ttc cag tct gtc aag tcc atc	586		
Pro Ala Asp Ser Leu Ser Met Ser Ser Phe Gln Ser Val Lys Ser Ile			
165	170	175	
tct aat tca ggc aag gcc agg ccc cag ccc ggc tcc ttc aac aag caa	634		
Ser Asn Ser Gly Lys Ala Arg Pro Gln Pro Gly Ser Phe Asn Lys Gln			
180	185	190	
gat tca aaa gct gac gtc tcc cag aag gcg gac ctg gaa gag gag ccc	682		
Asp Ser Lys Ala Asp Val Ser Gln Lys Ala Asp Leu Glu Glu Glu Pro			
195	200	205	210
cta ctt cac aac agc aag ctg gac aaa gtt cct ggg gta caa ggg cag	730		
Leu Leu His Asn Ser Lys Leu Asp Lys Val Pro Gly Val Gln Gly Gln			
215	220	225	
gcc aga aag gag aaa gca gag gcc tct aat gca gga gct gcc tgt atg	778		

Ala Arg Lys Glu Lys Ala Glu Ala Ser Asn Ala Gly Ala Ala Cys Met			
230	235	240	
ggg aac agc cag cac cag ggc agg cag atg ggg gcg ggg gca cac ccc	826		
Gly Asn Ser Gln His Gln Gly Arg Gln Met Gly Ala Gly Ala His Pro			
245	250	255	
cca atg atc ctg ccc ctt ccc ctg cga aag ccc acc aca ctt agg cag	874		
Pro Met Ile Leu Pro Leu Pro Leu Arg Lys Pro Thr Thr Leu Arg Gln			
260	265	270	
tgc gaa gtg ctc atc cgc gag ctg tgg aat acc aac ctc ctg cag acc	922		
Cys Glu Val Leu Ile Arg Glu Leu Trp Asn Thr Asn Leu Leu Gln Thr			
275	280	285	290
caa gag ctg cgg cac ctc aag tcc ctc ctg gaa ggg agc cag agg ccc	970		
Gln Glu Leu Arg His Leu Lys Ser Leu Leu Glu Gly Ser Gln Arg Pro			
295	300	305	
cag gca gcc ccg gag gaa gct agc ttt ccc agg gac caa gaa gcc acg	1018		
Gln Ala Ala Pro Glu Glu Ala Ser Phe Pro Arg Asp Gln Glu Ala Thr			
310	315	320	
cat ttc ccc aag gtc tcc acc aag agc ctc tcc aag aaa tgc ctg agc	1066		
His Phe Pro Lys Val Ser Thr Lys Ser Leu Ser Lys Lys Cys Leu Ser			
325	330	335	
cca cct gtg gcg gag cgt gcc atc ctg ccc gca ctg aag cag acc ccg	1114		
Pro Pro Val Ala Glu Arg Ala Ile Leu Pro Ala Leu Lys Gln Thr Pro			
340	345	350	
aag aac aac ttt gcc gag agg cag aag agg ctg cag gca atg cag aaa	1162		
Lys Asn Asn Phe Ala Glu Arg Gln Lys Arg Leu Gln Ala Met Gln Lys			
355	360	365	370
cgg cgc ctg cat cgc tca gtg ctt tga gccaccccaa tctggtcagt	1209		
Arg Arg Leu His Arg Ser Val Leu			
375			
gccaggccca ccaacctgca gctggagact ggctctctat agcatttcct gatacttccg	1269		
ctacttttag gcctggctaa attccaagac agataacact caagatagat aaagtacttg	1329		
atctccaaac tgacaaactg tttattttct agctgttatt ttgctatttg gcatttacat	1389		
aaaagcacac gatgaagcag gtatcgcctt acctgttgaa actgaaaata aagcttggtt	1449		
atttccaaaa aaaaaaaaaa aaa	1472		

&lt;210&gt;4

&lt;211&gt;378

&lt;212&gt;PRT

&lt;213&gt;人 (Homo sapiens)



&lt;400&gt;4

Met Ser Gly Ala Gly Val Ala Ala Gly Thr Arg Pro Pro Ser Ser Pro  
 1                   5                   10                   15  
 Thr Pro Gly Ser Arg Arg Arg Arg Gln Arg Pro Ser Val Gly Val Gln  
                   20                   25                   30  
 Ser Leu Arg Pro Gln Ser Pro Gln Leu Arg Gln Ser Asp Pro Gln Lys  
                   35                   40                   45  
 Arg Asn Leu Asp Leu Glu Lys Ser Leu Gln Phe Leu Gln Gln Gln His  
                   50                   55                   60  
 Ser Glu Met Leu Ala Lys Leu His Glu Glu Ile Glu His Leu Lys Arg  
 65                   70                   75                   80  
 Glu Asn Lys Asp Leu His Tyr Lys Leu Ile Met Asn Gln Thr Ser Gln  
                   85                   90                   95  
 Lys Lys Asp Gly Pro Ser Gly Asn His Leu Ser Arg Ala Ser Ala Pro  
                   100                   105                   110  
 Leu Gly Ala Arg Trp Val Cys Ile Asn Gly Val Trp Val Glu Pro Gly  
                   115                   120                   125  
 Gly Pro Ser Pro Ala Arg Leu Lys Glu Gly Ser Ser Arg Thr His Arg  
                   130                   135                   140  
 Pro Gly Gly Lys Arg Gly Arg Leu Ala Gly Gly Ser Ala Asp Thr Val  
 145                   150                   155                   160  
 Arg Ser Pro Ala Asp Ser Leu Ser Met Ser Ser Phe Gln Ser Val Lys  
                   165                   170                   175  
 Ser Ile Ser Asn Ser Gly Lys Ala Arg Pro Gln Pro Gly Ser Phe Asn  
                   180                   185                   190  
 Lys Gln Asp Ser Lys Ala Asp Val Ser Gln Lys Ala Asp Leu Glu Glu  
                   195                   200                   205  
 Glu Pro Leu Leu His Asn Ser Lys Leu Asp Lys Val Pro Gly Val Gln  
                   210                   215                   220  
 Gly Gln Ala Arg Lys Glu Lys Ala Glu Ala Ser Asn Ala Gly Ala Ala  
 225                   230                   235                   240  
 Cys Met Gly Asn Ser Gln His Gln Gly Arg Gln Met Gly Ala Gly Ala  
                   245                   250                   255  
 His Pro Pro Met Ile Leu Pro Leu Pro Leu Arg Lys Pro Thr Thr Leu  
                   260                   265                   270  
 Arg Gln Cys Glu Val Leu Ile Arg Glu Leu Trp Asn Thr Asn Leu Leu  
                   275                   280                   285

Gln Thr Gln Glu Leu Arg His Leu Lys Ser Leu Leu Glu Gly Ser Gln  
 290 295 300  
 Arg Pro Gln Ala Ala Pro Glu Glu Ala Ser Phe Pro Arg Asp Gln Glu  
 305 310 315 320  
 Ala Thr His Phe Pro Lys Val Ser Thr Lys Ser Leu Ser Lys Lys Cys  
 325 330 335  
 Leu Ser Pro Pro Val Ala Glu Arg Ala Ile Leu Pro Ala Leu Lys Gln  
 340 345 350  
 Thr Pro Lys Asn Asn Phe Ala Glu Arg Gln Lys Arg Leu Gln Ala Met  
 355 360 365  
 Gln Lys Arg Arg Leu His Arg Ser Val Leu  
 370 375

<210>5

<211>1315

<212>DNA

<213> 人 (Homo sapiens)

<220>

<221>CDS

<222>(251).. (1114)

<223>

<400>5

agagaaagta tagccactgc ttagacagcc agggaaacgt gtgcggggaa gtggaggact 60  
 caggctctcg tgcgagagcg gagttggacg tgcagggccg ctggggtcac gcgagactct 120  
 cccgcctccc ctccgcgtga gctctgggat ggtccgcgcc gggagcgcgc gcgaggcttg 180  
 aagcgcgggt gaagcgcgca ggteggagtg acagctgcgc tgccggcccc gctgcggtca 240  
 gcaacgcgcc atg gac gca gag ctg gca gag gtg cgc gcc ttg caa get 289  
 Met Asp Ala Glu Leu Ala Glu Val Arg Ala Leu Gln Ala  
 1 5 10  
 gag atc gcg gcc ctg cgg cga gcg tgt gag gac cca ccg gcg ccc tgg 337  
 Glu Ile Ala Ala Leu Arg Arg Ala Cys Glu Asp Pro Pro Ala Pro Trp  
 15 20 25  
 gaa gag aag tcc cga gtc caa aaa tct ttt caa gcc ata cac caa ttc 385  
 Glu Glu Lys Ser Arg Val Gln Lys Ser Phe Gln Ala Ile His Gln Phe  
 30 35 40 45  
 aat ttg gaa gga tgg aag tct tca aaa gat ctg aaa aat cag ctt gga 433  
 Asn Leu Glu Gly Trp Lys Ser Ser Lys Asp Leu Lys Asn Gln Leu Gly

cat tta gaa tca gaa ctt tca ttt cta agt acg ctt act ggc atc aat	50	55	60	481
His Leu Glu Ser Glu Leu Ser Phe Leu Ser Thr Leu Thr Gly Ile Asn				
	65	70	75	
ata aga aat cac tcc aag cag aca gaa gac cta aca agc act gag atg				529
Ile Arg Asn His Ser Lys Gln Thr Glu Asp Leu Thr Ser Thr Glu Met				
	80	85	90	
aca gaa aag agt att aga aaa gtt cta cag aga cac aga tta tca gga				577
Thr Glu Lys Ser Ile Arg Lys Val Leu Gln Arg His Arg Leu Ser Gly				
	95	100	105	
aat tgc cac atg gtt aca ttt caa ctt gaa ttt cag att ctg gaa att				625
Asn Cys His Met Val Thr Phe Gln Leu Glu Phe Gln Ile Leu Glu Ile				
	110	115	120	
cag aat aag gag aga tta tct tct gct gtt act gac ctc aac ata ata				673
Gln Asn Lys Glu Arg Leu Ser Ser Ala Val Thr Asp Leu Asn Ile Ile				
	130	135	140	
atg gag ccc aca gaa tgc tca gaa tta agt gaa ttt gtg tct aga gca				721
Met Glu Pro Thr Glu Cys Ser Glu Leu Ser Glu Phe Val Ser Arg Ala				
	145	150	155	
gaa gag aga aaa gat ctg ttc atg ttt ttc cga agc ctg cat ttt ttt				769
Glu Glu Arg Lys Asp Leu Phe Met Phe Phe Arg Ser Leu His Phe Phe				
	160	165	170	
gtg gag tgg ttt gaa tat cgt aag cgc acg ttt aaa cat ctc aag gaa				817
Val Glu Trp Phe Glu Tyr Arg Lys Arg Thr Phe Lys His Leu Lys Glu				
	175	180	185	
aag tac cca gat gcc gtg tac ctc tcg gag ggg ccc tcc tcc tgc tcc				865
Lys Tyr Pro Asp Ala Val Tyr Leu Ser Glu Gly Pro Ser Ser Cys Ser				
	190	195	200	
atg ggg atc cgc agc gcc agc cgg cca ggg ttt gaa tta gtc att gtt				913
Met Gly Ile Arg Ser Ala Ser Arg Pro Gly Phe Glu Leu Val Ile Val				
	210	215	220	
tgg agg ata caa ata gat gaa gat ggg aag gtt ttt cca aag ctg gat				961
Trp Arg Ile Gln Ile Asp Glu Asp Gly Lys Val Phe Pro Lys Leu Asp				
	225	230	235	
ctt ctc acc aaa gtc cca cag cga gcc ctg gag ctg gac aag aac aga				1009
Leu Leu Thr Lys Val Pro Gln Arg Ala Leu Glu Leu Asp Lys Asn Arg				
	240	245	250	
gcc ata gaa act gct cct ctc agc ttc cga acc ctg gta gga ctg ctt				1057
Ala Ile Glu Thr Ala Pro Leu Ser Phe Arg Thr Leu Val Gly Leu Leu				

255	260	265	
gga atc gaa gct gct ctg gaa agc ctg ata aaa tcg ctt tgt gca gag			1105
Gly Ile Glu Ala Ala Leu Glu Ser Leu Ile Lys Ser Leu Cys Ala Glu			
270	275	280	285
gag aac aac tagttccaaa acagtgaacg tggaggatga agatgctgcg			1154
Glu Asn Asn			
tggaggaaca tgcaatttta ttcaatataa acatttgcta ttttctgctt agaaaccaca			1214
ccctgaagac gtgctgtcta tgcagttatg gcacattata tggaaactct catgacatga			1274
aaaataaata caactagtta agtataaaat gccaaaaaaa a			1315

&lt;210&gt;6

&lt;211&gt;288

&lt;212&gt;PRT

&lt;213&gt;人 (Homo sapiens)

&lt;400&gt;6

Met Asp Ala Glu Leu Ala Glu Val Arg Ala Leu Gln Ala Glu Ile Ala			
1	5	10	15
Ala Leu Arg Arg Ala Cys Glu Asp Pro Pro Ala Pro Trp Glu Glu Lys			
	20	25	30
Ser Arg Val Gln Lys Ser Phe Gln Ala Ile His Gln Phe Asn Leu Glu			
	35	40	45
Gly Trp Lys Ser Ser Lys Asp Leu Lys Asn Gln Leu Gly His Leu Glu			
	50	55	60
Ser Glu Leu Ser Phe Leu Ser Thr Leu Thr Gly Ile Asn Ile Arg Asn			
65	70	75	80
His Ser Lys Gln Thr Glu Asp Leu Thr Ser Thr Glu Met Thr Glu Lys			
	85	90	95
Ser Ile Arg Lys Val Leu Gln Arg His Arg Leu Ser Gly Asn Cys His			
	100	105	110
Met Val Thr Phe Gln Leu Glu Phe Gln Ile Leu Glu Ile Gln Asn Lys			
	115	120	125
Glu Arg Leu Ser Ser Ala Val Thr Asp Leu Asn Ile Ile Met Glu Pro			
	130	135	140
Thr Glu Cys Ser Glu Leu Ser Glu Phe Val Ser Arg Ala Glu Glu Arg			
145	150	155	160
Lys Asp Leu Phe Met Phe Phe Arg Ser Leu His Phe Phe Val Glu Trp			
	165	170	175

Phe Glu Tyr Arg Lys Arg Thr Phe Lys His Leu Lys Glu Lys Tyr Pro  
 180 185 190  
 Asp Ala Val Tyr Leu Ser Glu Gly Pro Ser Ser Cys Ser Met Gly Ile  
 195 200 205  
 Arg Ser Ala Ser Arg Pro Gly Phe Glu Leu Val Ile Val Trp Arg Ile  
 210 215 220  
 Gln Ile Asp Glu Asp Gly Lys Val Phe Pro Lys Leu Asp Leu Leu Thr  
 225 230 235 240  
 Lys Val Pro Gln Arg Ala Leu Glu Leu Asp Lys Asn Arg Ala Ile Glu  
 245 250 255  
 Thr Ala Pro Leu Ser Phe Arg Thr Leu Val Gly Leu Leu Gly Ile Glu  
 260 265 270  
 Ala Ala Leu Glu Ser Leu Ile Lys Ser Leu Cys Ala Glu Glu Asn Asn  
 275 280 285

<210>7

<211>20

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物

<400>7

cgaccacttt gtcaagctca

20

<210>8

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物

<400>8

ggttgagcac aggttacttt att

23

<210>9

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物

<400>9

caaatattag gtggagccaa cac

23

<210>10

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物

<400>10

tagatcacct tggcaaagaa cac

23

<210>11

<211>20

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物

<400>11

acctcaagtc cctcctggaa

20

<210>12

<211>23

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 RT-PCR 引物

<400>12

tcagtttcaa caggtaagge gat

23

<210>13	
<211>23	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 RT-PCR 引物	
<400>13	
agagccatag aaactgctcc tct	23
<210>14	
<211>23	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 RT-PCR 引物	
<400>14	
cataactgca tagacagcac gtc	23
<210>15	
<211>20	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 RT-PCR 引物	
<400>15	
gggaagagaa gtcccgagtc	20
<210>16	
<211>24	
<212>DNA	
<213> 人工的	

<220>

<223> 人工合成的 RT-PCR 引物

<400>16

tccttattct gaatttccag aatc

24

<210>17

<211>30

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 5' RACE 引物

<400>17

caagcagtcc taccagggtt cggaagctga

30

<210>18

<211>30

<212>DNA

<213> 人工的

<220>

<223> 人工合成的嵌套 PCR 引物

<400>18

ccagggttcg gaagctgaga ggagcagttt

30

<210>19

<211>30

<212>DNA

<213> 人工的

<220>

<223> 人工合成的 PCR 引物

<400>19

ccggaattca tgcagagagc ttcacgtctg

30



<210>20	
<211>33	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 PCR 引物	
<400>20	
ccgctcgaga acatcaggat gaaatttett ttc	33
<210>21	
<211>30	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 PCR 引物	
<400>21	
ccggaattca tgagcgggtgc gggggtggcg	30
<210>22	
<211>30	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 PCR 引物	
<400>22	
ccgctcgaga agcactgagc gatgcaggcg	30
<210>23	
<211>35	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 PCR 引物	

<400>23	
ccggaattca tggacgcaga gctggcagag gtgcg	35
<210>24	
<211>30	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 PCR 引物	
<400>24	
ccgctcgagg ttgttctcct ctgcacaaag	30
<210>25	
<211>51	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 siRNA 寡核苷酸	
<400>25	
caccgaagca gcacgacttc ttcttcaaga gagaagaagt cgtgctgctt c	51
<210>26	
<211>51	
<212>DNA	
<213> 人工的	
<220>	
<223> 人工合成的 siRNA 寡核苷酸	
<400>26	
aaaagaagca gcacgacttc ttctctcttg aagaagaagt cgtgctgctt c	51
<210>27	
<211>18	
<212>DNA	

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>27

gaagcagcac gacttctt

18

<210>28

<211>19

<212>DNA

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>28

catcgcaact gtgttgacc

19

<210>29

<211>19

<212>DNA

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>29

tgccagacag tggacagag

19

<210>30

<211>19

<212>DNA

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>30

gcctgcagtt cctgcagca 19

<210>31

<211>19

<212>DNA

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>31

gcttcagtc tgtcaagtc 19

<210>32

<211>19

<212>DNA

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>32

agcagaggcc tctaatgca 19

<210>33

<211>19

<212>DNA

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>33

actgctctc tcagcttcc 19

<210>34

<211>19

<212>DNA

<213> 人工的

<220>

<223>siRNA 的靶序列

<400>34

gtacgcttac tggcatcaa

19

**a**

12 份临床样品

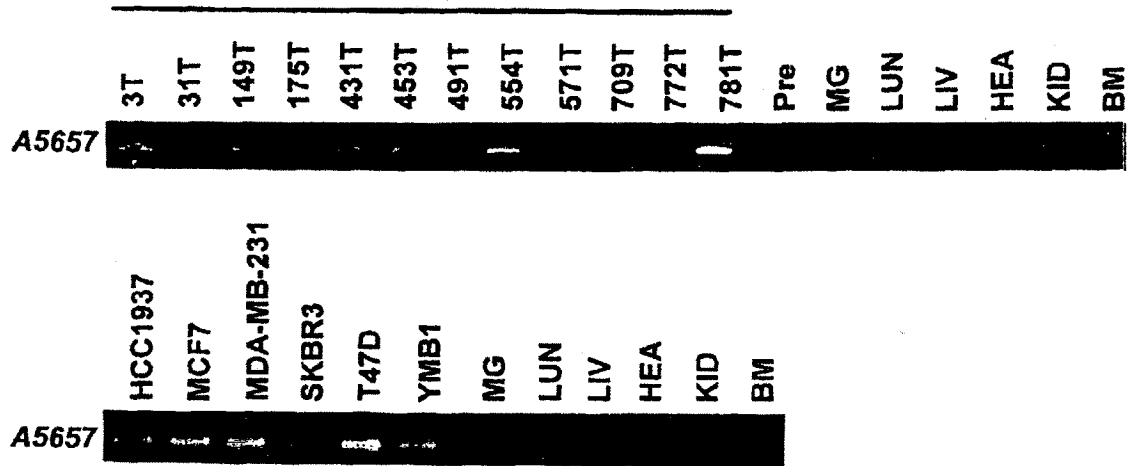


图 1a

**b**

12 份临床样品

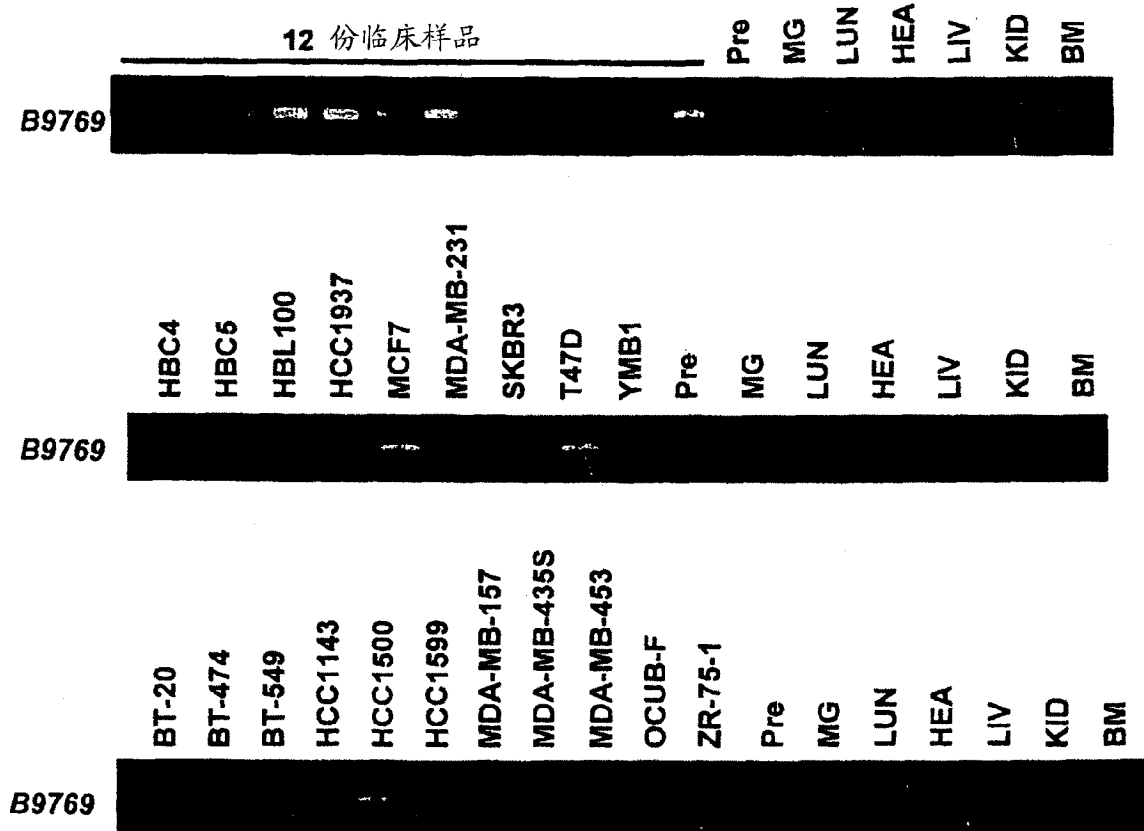


图 1b

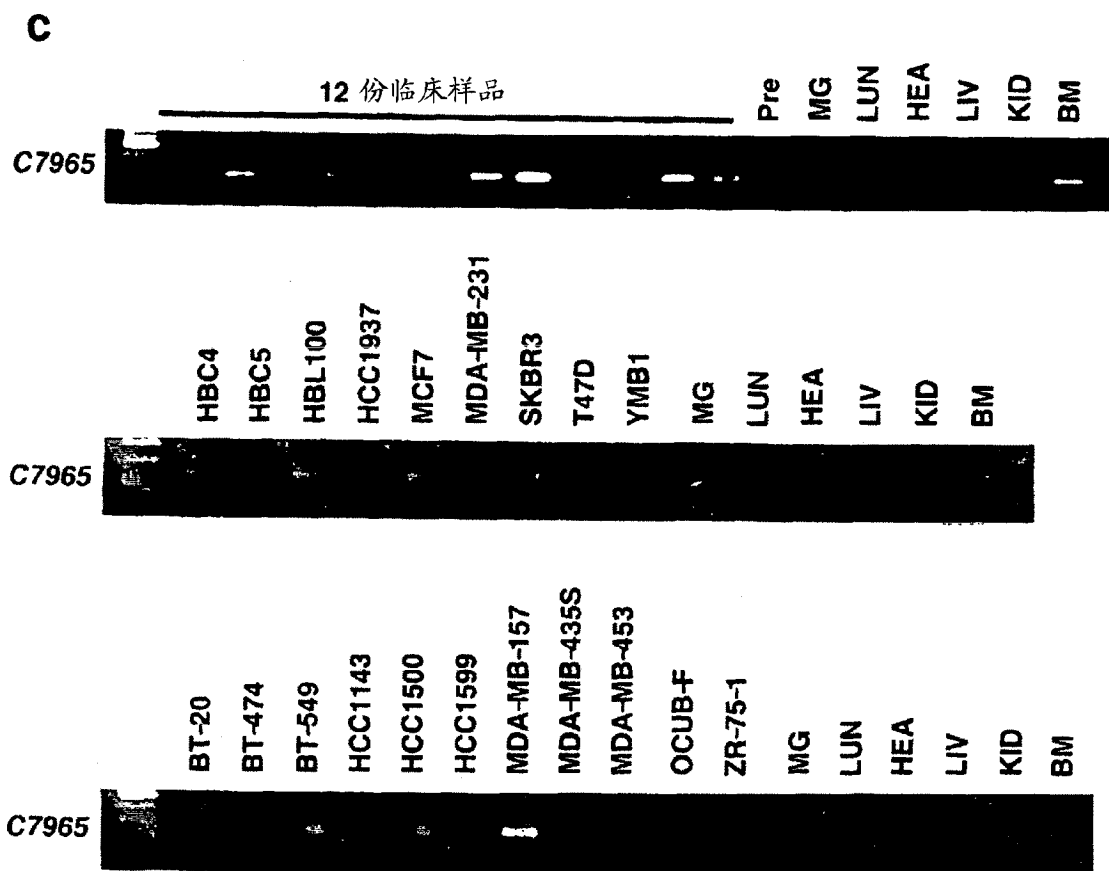


图 1c

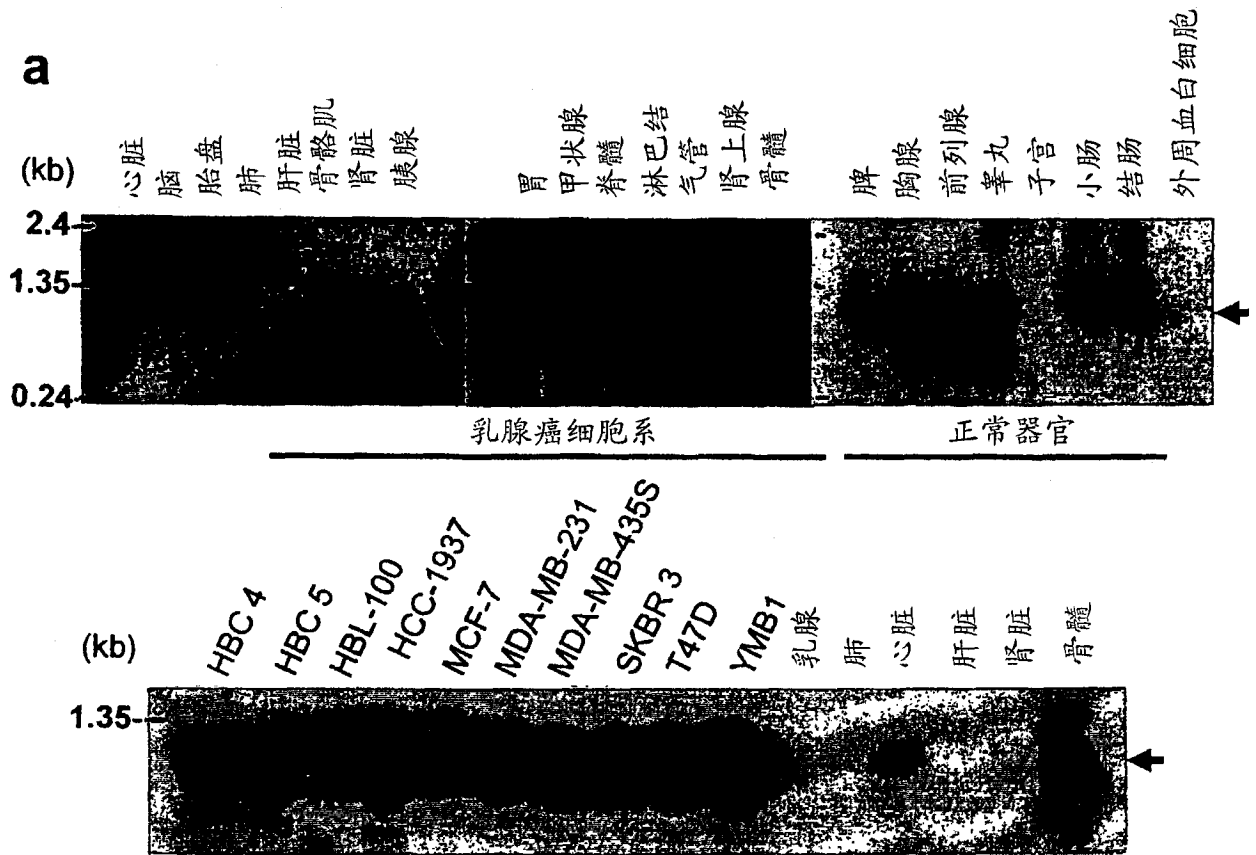


图 2a



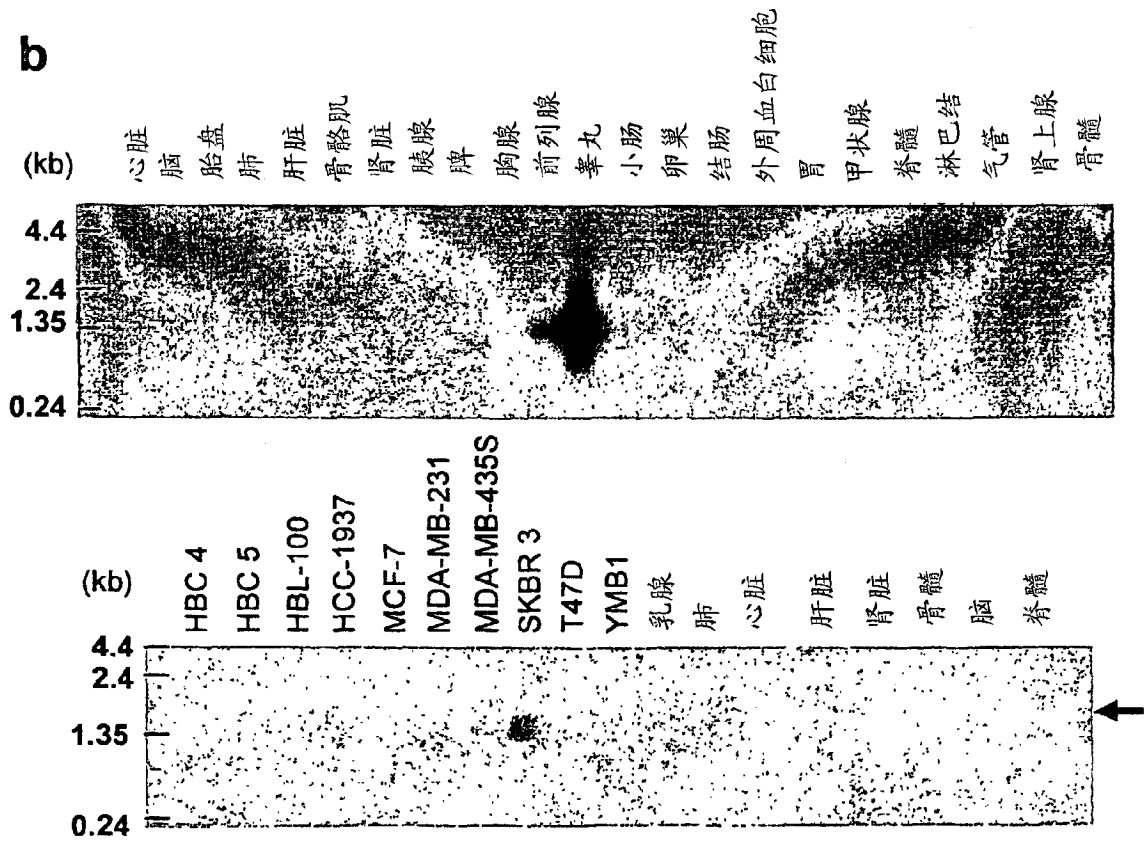


图 2b

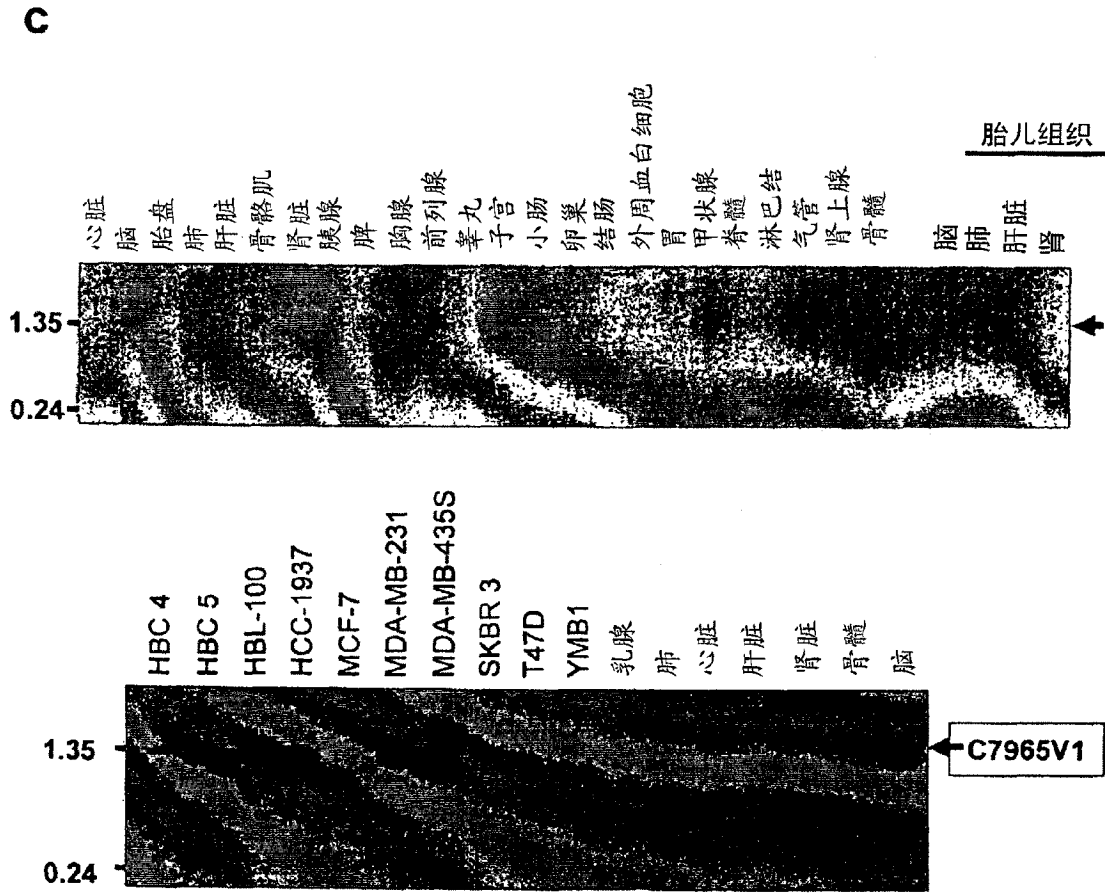


图 2c

**a** T47D 乳腺癌细胞

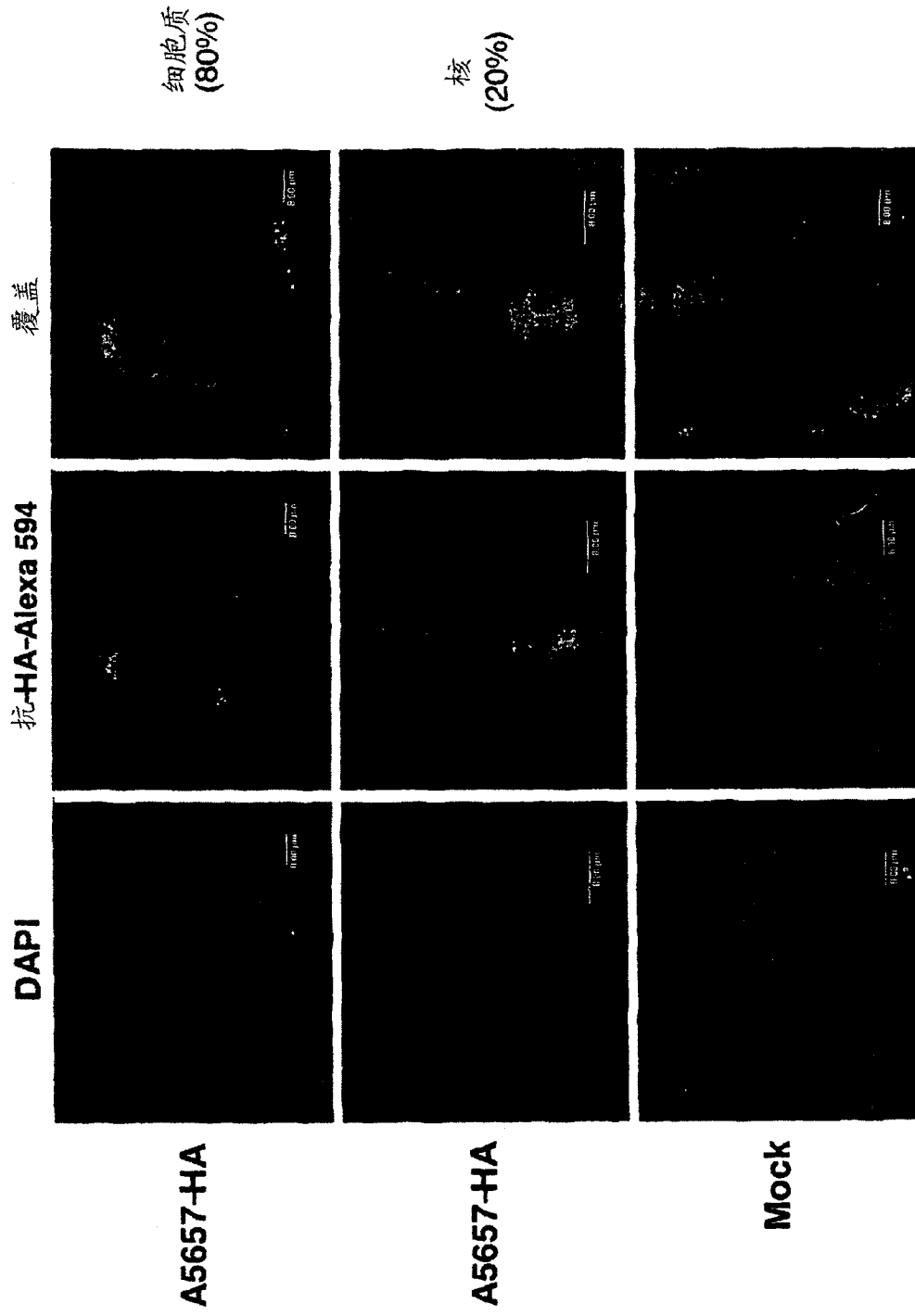
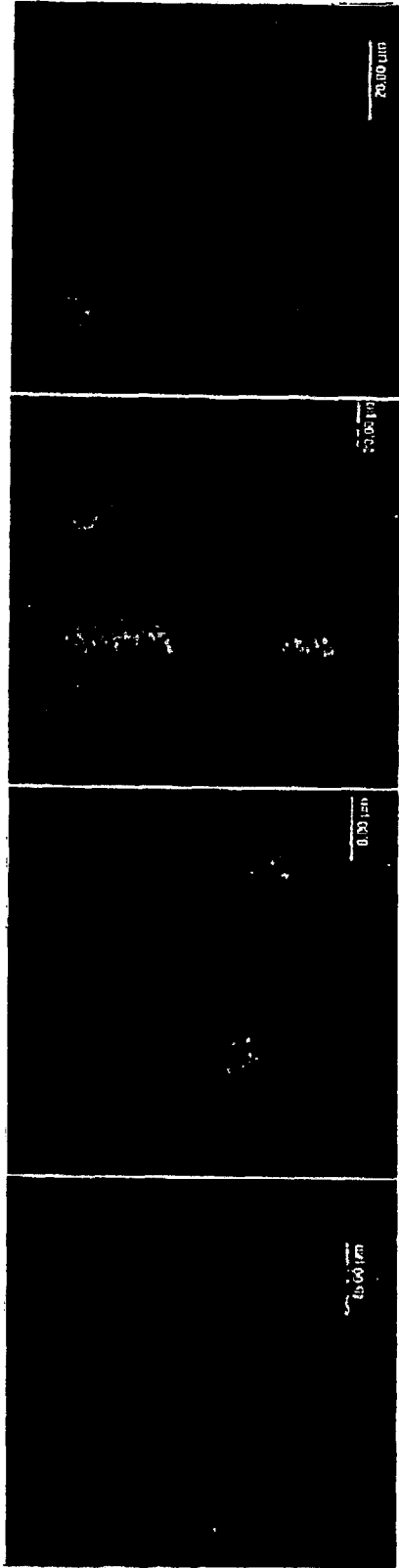


图 3a

b

6 孔平板  $1 \times 10^4$  细胞/孔



6 孔平板  $1 \times 10^5$  细胞/孔

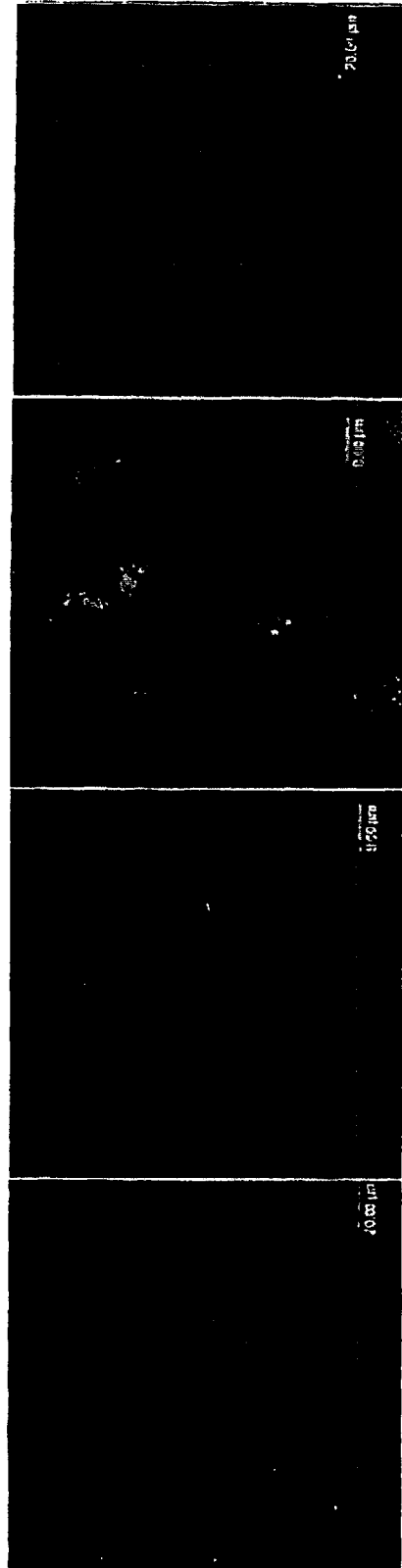
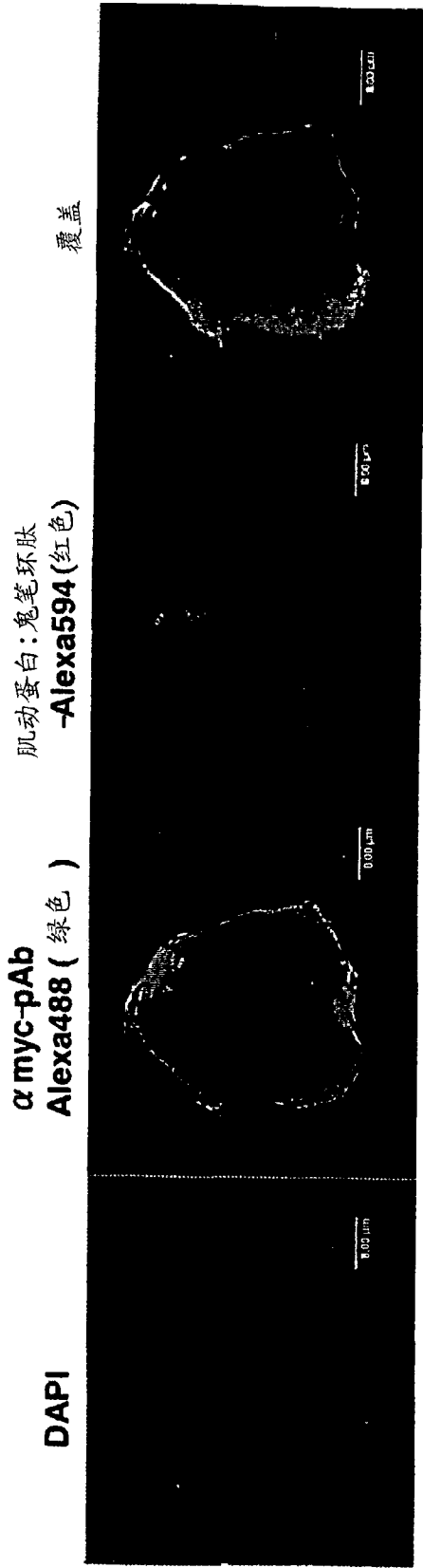


图 3b

C



69

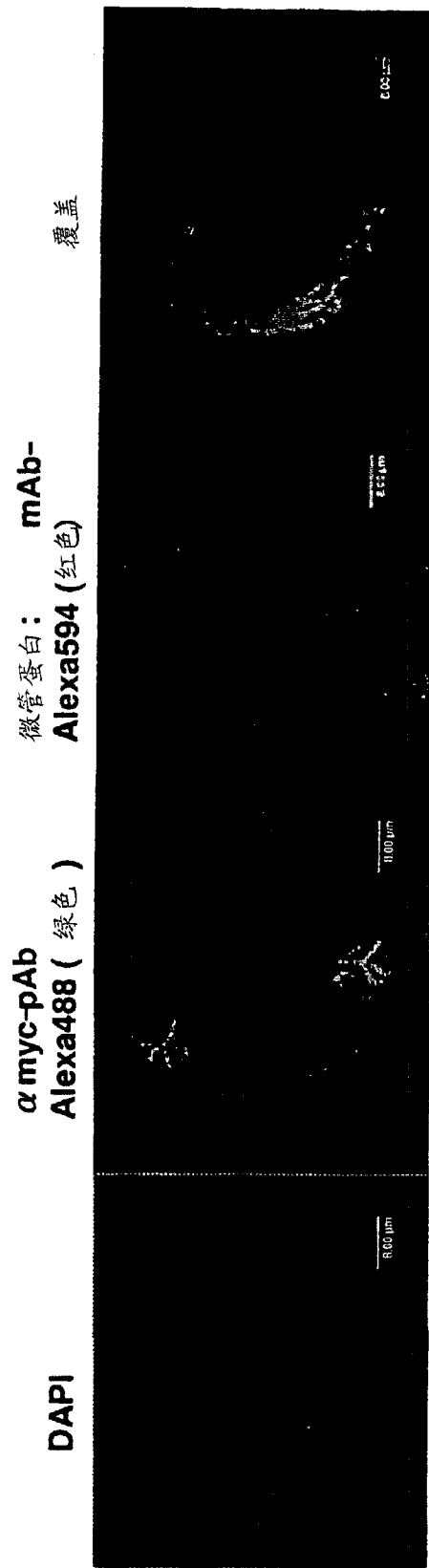


图 3c

d

抗-myc:FITC

合并

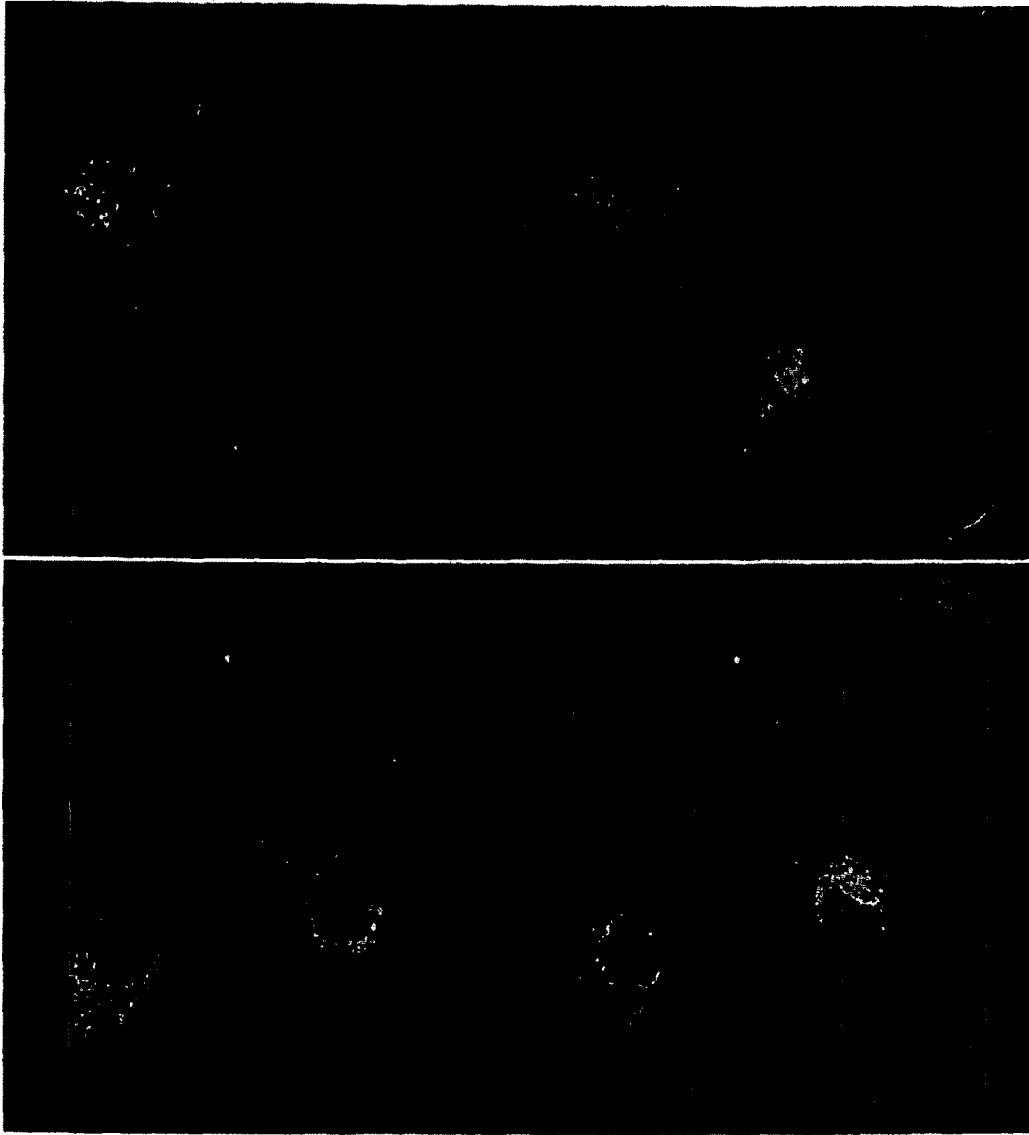


图 3d

(a)

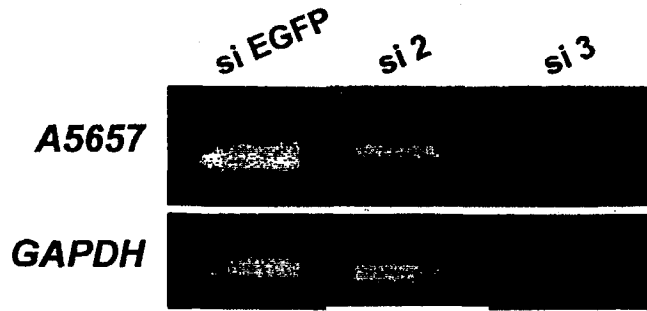


图 4a

(b)

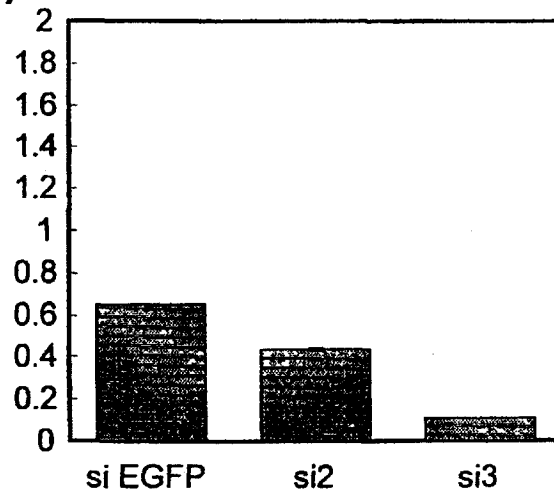


图 4b

(c)

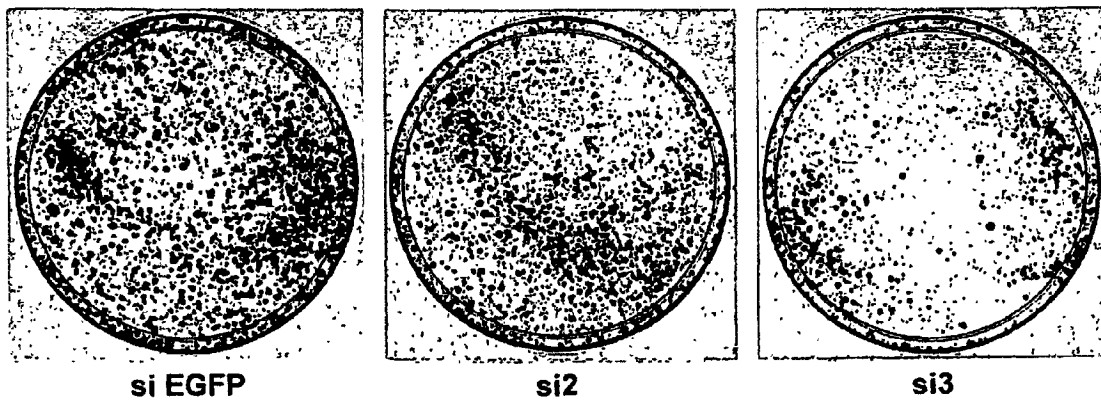


图 4c

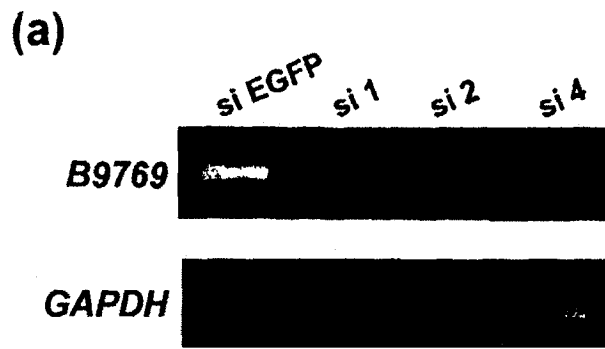


图 5a

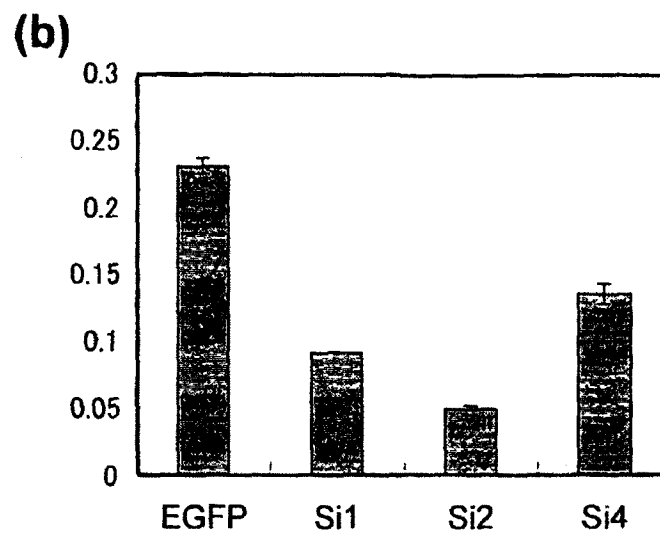


图 5b



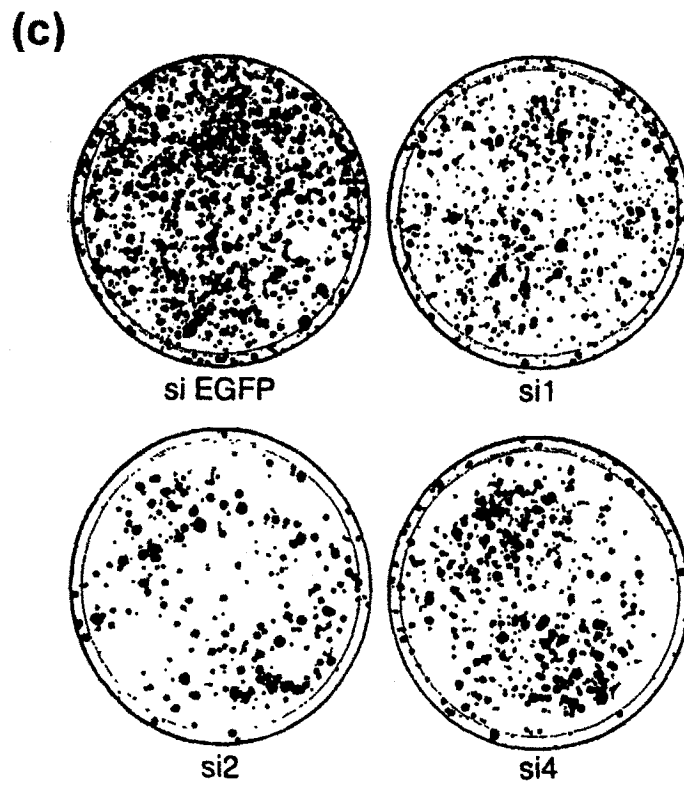


图 5c



图 6a

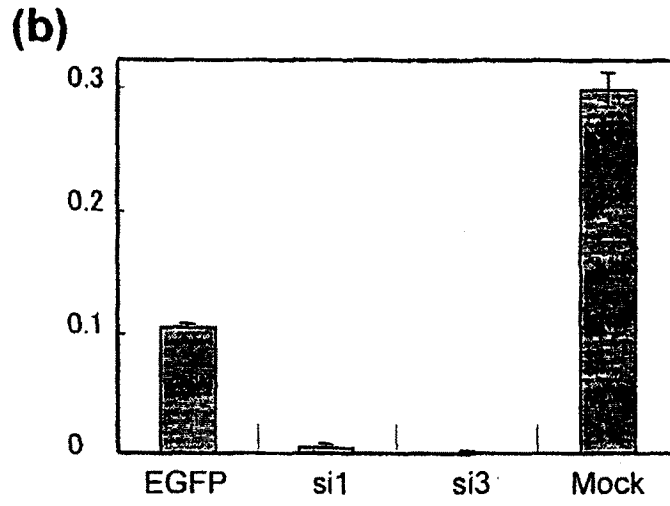


图 6b

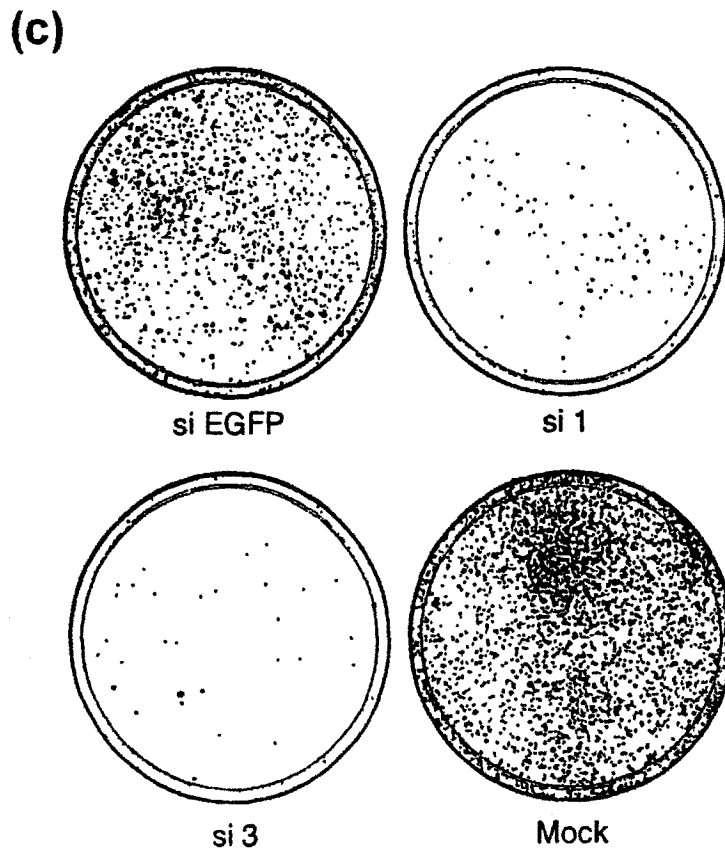


图 6c

a

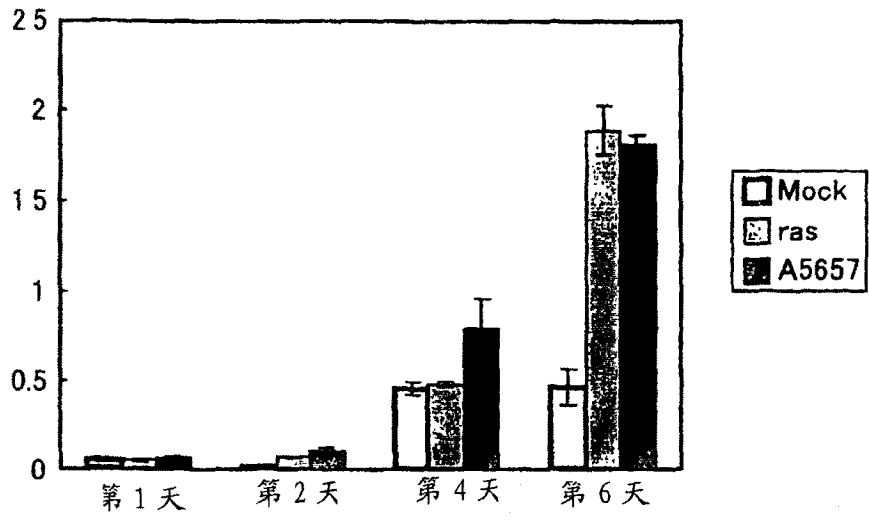


图 7a

b

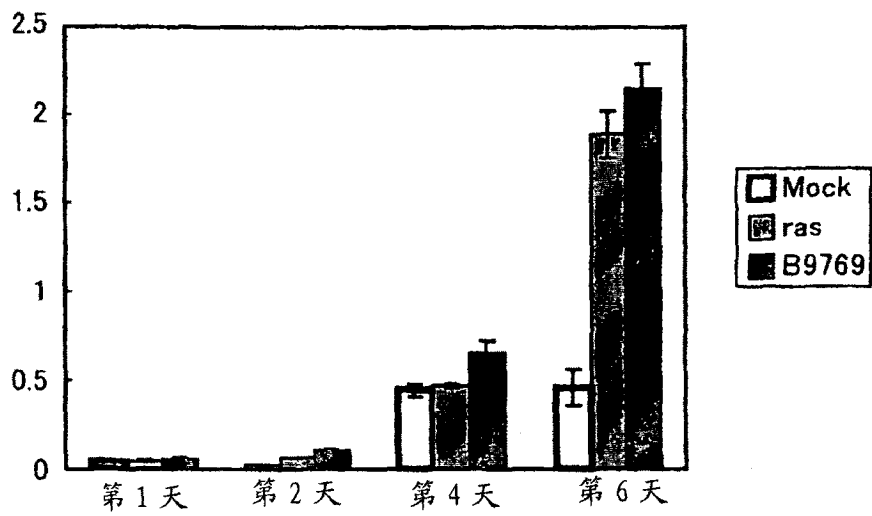


图 7b

**C**

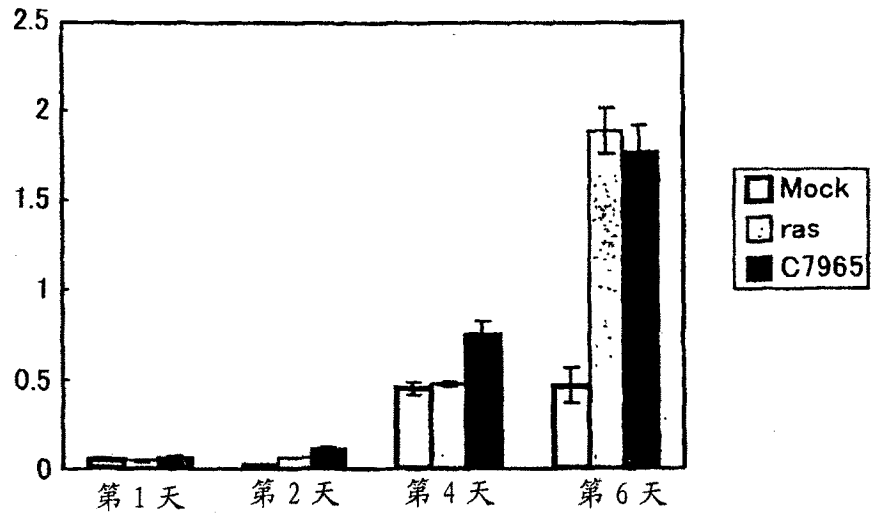


图 7c

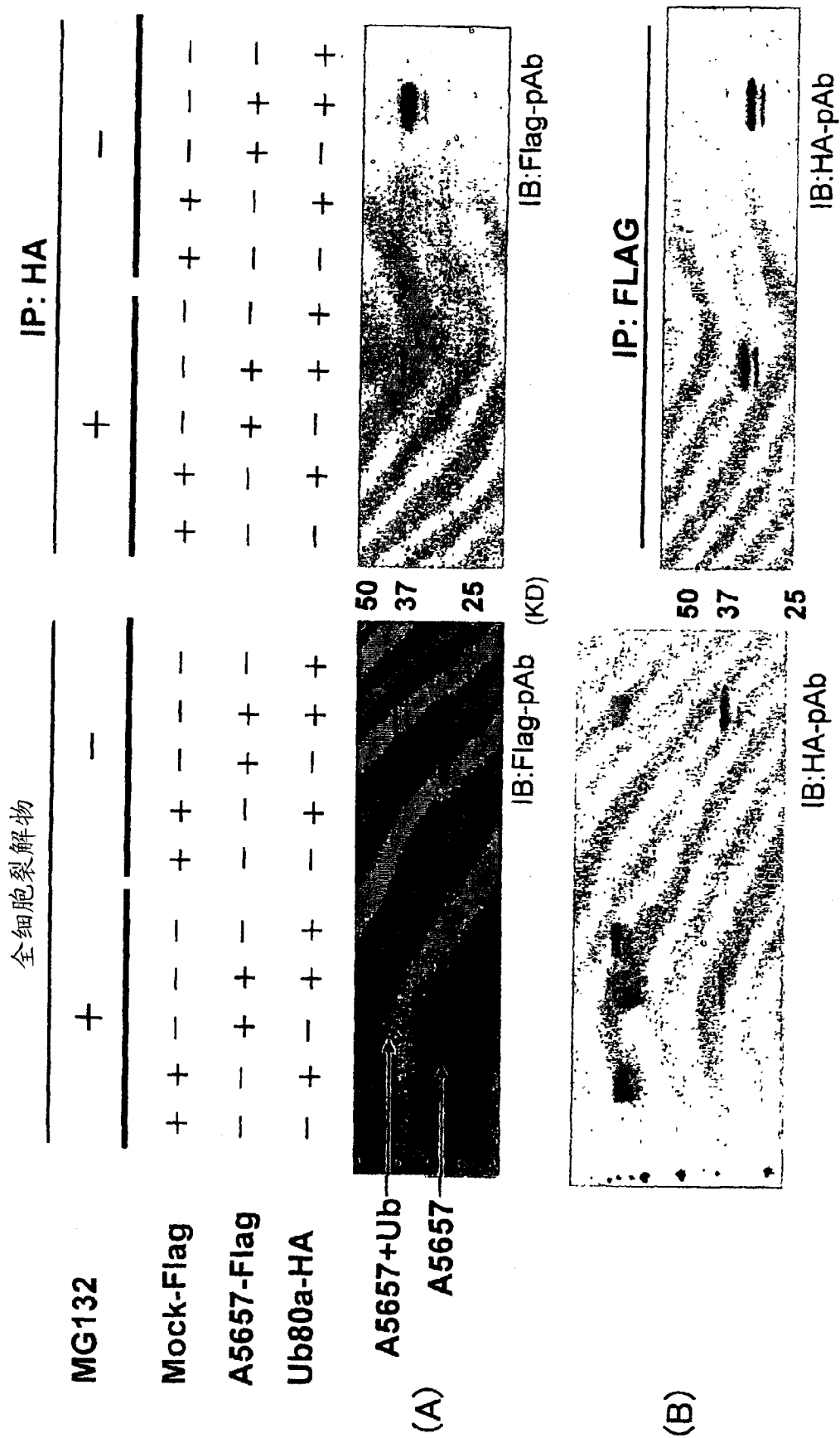


图 8