

[19] 中华人民共和国国家知识产权局



# [12] 发明专利说明书

专利号 ZL 01808488.5

[51] Int. Cl.

*C07K 16/18 (2006.01)*  
*A61K 38/17 (2006.01)*  
*A61K 39/00 (2006.01)*  
*A61K 39/395 (2006.01)*  
*G01N 33/50 (2006.01)*  
*G01N 33/53 (2006.01)*

[45] 授权公告日 2008 年 4 月 30 日

[11] 授权公告号 CN 100384875C

[51] Int. Cl. (续)

*A61K 48/00 (2006.01)*

*C12Q 1/68 (2006.01)*

*A61P 35/00 (2006.01)*

[22] 申请日 2001.2.21 [21] 申请号 01808488.5

[30] 优先权

[32] 2000. 2. 25 [33] GB [31] 0004576.5

[32] 2000. 12. 21 [33] GB [31] 0031341.1

[86] 国际申请 PCT/GB2001/000748 2001.2.21

[87] 国际公布 WO2001/062784 英 2001.8.30

[85] 进入国家阶段日期 2002.10.23

[73] 专利权人 牛津格莱克科学(英国)有限公司

地址 英国阿宾登

[72] 发明人 R·S·博伊德

A·C·斯塔姆普斯

J·A·特雷特

[56] 参考文献

WO9947669A2 1999.9.23

WO9807749A1 1998.2.26

审查员 宋智刚

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 徐雁漪

权利要求书 1 页 说明书 127 页 附图 10 页

[54] 发明名称

蛋白质

[57] 摘要

本发明提供了用于乳腺癌筛选、诊断和预后、监测乳腺癌治疗功效以及药物开发的方法和组合物。

1.对 BCMP-17 具有免疫特异性的单克隆抗体或其包含所述抗体的结合结构域的片段在生产乳腺癌治疗药物中的用途，其中抗体与治疗部分缀合。

2.权利要求 1 的用途，其中抗体是人源化抗体。

3.权利要求 1 的用途，其中抗体是完全人抗体。

4.权利要求 1 的用途，其中治疗部分是毒素。

5.权利要求 1 至 4 中任一项的用途，其中抗体通过聚乙二醇修饰。

## 蛋白质

### 引言

本发明涉及先前未曾报道的人类乳腺癌细胞中膜蛋白的鉴定，所述膜蛋白可以构成生物靶，可以制备针对所述靶的治疗抗体或其它药物，或者所述膜蛋白具有作为乳腺癌和乳腺癌转移癌的诊断和预后标记的效用。

### 发明背景

乳腺癌是一种在美国妇女中最常诊断出的非皮肤癌。在癌症相关的死亡中，它仅次于肺癌。1997年将诊断出大约180,000个乳腺癌新病例，预期约44,000名妇女死于该疾病(National Cancer Institute, [www.nci.nih.gov](http://www.nci.nih.gov), USA, 1999)。在英国，乳腺癌是迄今妇女最常见的癌症，1998年有34,600个新病例(Cancer Research Campaign, [www.crc.org.uk](http://www.crc.org.uk), UK, 2000)。99%的乳腺癌发生在妇女中。发生乳腺癌的危险随年龄稳定增加；终身发生乳腺癌的危险在美国估计为每8个妇女中有1个。在美国，乳腺癌治疗的年费用大约是\$100亿(Fuqua等2000, American Association for Cancer Research, [www.aacr.org](http://www.aacr.org), USA)。在过去50年间，乳腺癌发生率一直在上升，但最近变为平稳。这可能反映出通过乳房X射线照相术较早检测出乳腺癌的时期。多种已建立的因素可以增加妇女患此病的危险。这些因素包括年纪较大、先前有乳腺癌病史、显著暴露于辐射、强的乳腺癌家族史、社会经济等级较高、未经产、初潮早、绝经晚或初次妊娠时年龄超过30岁。在生命较早期延长应用口服避孕药似乎略微增加危险。延长的绝经后雌激素替代疗法增加危险20-40%。已经推测，初潮时年龄的降低、改变分娩形式、或外源雌激素应用增加，导致乳腺癌发病率增加(Fuqua等2000,

American Association for Cancer Research, [www.aacr.org](http://www.aacr.org), USA)。

### 乳腺癌的病因

乳腺癌是一种异质性疾病。虽然雌性激素在驱动许多乳房肿瘤的起源和进化方面起重要作用，但也涉及许多其它已被识别的因子和未知因子。已鉴定癌基因的微扰包括 HER-2 和表皮生长因子受体基因的扩增和细胞周期蛋白 D1 的过量表达。可以将这些癌基因的过量表达与明显预后差相联系。同样，在乳腺癌方面大量记载了遗传的改变或者肿瘤抑制基因例如 p53 基因的丧失，它们也与预后较差有关。研究人员已经鉴定出两种基因，称为 BRCA1 和 BRCA2，它们是绝经前家族性乳腺癌的预兆。遗传危险的评价现在是可能的，它们可能加强供化学保护试验用的候选者的鉴定(Fuqua 等 2000, American Association for Cancer Research, [www.aacr.org](http://www.aacr.org), USA)。

### 诊断

乳腺癌的早期诊断对于保证最佳治疗结果是至关重要的。具有先进保健系统的许多国家制定了乳腺癌的筛选计划。这通常在 50-60 年龄段采用乳房的定期 X 射线检查(乳房 X 射线照相术)，其中显示了这种干涉的最大益处。某些专家主张将这种计划扩大至 60 岁以上和 40-49 年龄组。许多国家的健康专家也宣传了妇女定期乳房自查的重要性。在这些筛选过程中检出的异常和表现出症状的病例通常可以通过针吸活组织检查细胞学、用立体定位(stereotatic)技术或超声技术对不可触知的病灶进行核心针吸活组织检查、或者切开性或切除性活组织检查加以证实。同时，可以测定与治疗选择和预后相关的其它信息例如雌激素(ER)和孕酮受体(PR)状况(National Cancer Institute, USA, 2000, Breast Cancer PDQ, [www.nci.nih.gov](http://www.nci.nih.gov))。

## 疾病分期和预后

疾病分期是查明癌症扩散到多远的一种方法。the American Joint Committee on Cancer (AJCC)的疾病分期系统也称为 TNM 系统，是乳腺癌最常用的系统。用于疾病分期的 TNM 系统给出三类关键的信息：

后接 0-4 的数字的字母 T 描述了肿瘤的大小以及扩散至乳房下皮肤或胸部。较大的数字是指较大的肿瘤和/或更多地扩散至乳房附近的组织。

后接 0-3 的数字的字母 N 指出癌症是否扩散至乳房附近的淋巴结，如果是，那么受影响的淋巴结是否附着于臂下其它结构。

后接 0 或 1 的字母 M 表明癌症是否转移到机体其它器官或转移到不在乳房附近的淋巴结。

为了使该信息略微清楚，TNM 描述可以结合为一组更简单的病期，标为 0 期至 IV 期(0-4)。一般而言，数字越小，癌症扩散就越少。数字越大，例如 IV 期(4)，意味着癌症越严重(American Cancer Society, 2000, USA, [www.cancer.org](http://www.cancer.org))。

### 按病期的乳腺癌存活率

病期	5 年相对存活率
0	100%
I	98%
IIA	88%
IIB	76%
IIIA	56%
IIIB	49%
IV	16%

(American Cancer Society, 2000, USA, [www.cancer.org](http://www.cancer.org))

尽管解剖学病期(原发性肿瘤的大小、涉及腋淋巴结)是一个重要的预后因素，但其它特征可能具有预测价值。例如，得自 the National Surgical Adjuvant Breast and Bowel Project (NSABP)和 the International

Breast Cancer Study Group (IBCSG)的研究已经表明, 肿瘤细胞核分级和组织学分级分别是乳腺癌辅佐治疗后结果的重要指标。有重要证据表明, 雌激素受体状况和原发性肿瘤增殖能力的测量(胸苷标记指数或S期和倍性的流式细胞术测量)可能有重要的独立预测价值。在II期疾病中, PR状况可能比ER状况的预后价值更大。肿瘤血管形成、c-erbB-2、c-myc、p53表达和淋巴管侵入也可能是乳腺癌患者的预后指标(National Cancer Institute, USA, 2000, Breast Cancer PDQ, [www.nci.org](http://www.nci.org)和其中的参考文献)。

## 治疗

外科手术、放射疗法、激素疗法和化学疗法是最常见的乳腺癌疗法。

外科手术 大多数患有乳腺癌的妇女将接受某种类型的外科手术。外科手术的目的是尽可能多地除去癌症。这可能以肿块切除术形式, 或者更多的是根治性乳房切除术加上乳房整复。外科手术也可以与其它疗法如化学疗法、激素疗法或放射疗法联合。也可以实施外科手术, 以查明乳腺癌是否扩散至臂下淋巴结(腋窝解剖)、恢复更为正常的外观(整复外科)或缓解晚期癌症的症状。

化学疗法 化学疗法是应用抗癌药杀伤癌细胞。在外科手术后给予化学疗法(辅佐疗法)时, 可能减少癌症复发的机会。化学疗法也可以用作癌症发现时已广为扩散、或者在初次治疗后广泛扩散的妇女的主要疗法。在外科手术之前给出新佐剂化学疗法, 常常使肿瘤缩小并使其更易于切除。化学疗法以周期给予, 在每个治疗期后是一个恢复期。总疗程持续3-6个月。应用几种药物常常比仅用一种药物更为有效。最常用的组合是: 环磷酰胺、氨甲蝶呤和氟尿嘧啶(CMF), 环磷酰胺、多柔比星(阿霉素)和氟尿嘧啶(CAF), 多柔比星(阿霉素)和环磷酰胺(AC), 加上或不加紫杉醇(紫杉酚)多柔比星(阿霉素), 然后是CMF。

**放射疗法** 放射疗法最常用于乳腺癌的治疗。它可以用来在外科手术之前减小肿瘤的大小，或在外科手术之后破坏乳房、胸壁或腋下区域中残余的癌细胞。

**激素疗法和化学防护** 激素 - 雌激素可以增加某些妇女中乳腺癌细胞的生长。给予阻断雌激素效应的药物例如他莫昔芬，以对抗这种生长。另一种新药 - 雷洛昔芬也阻断雌激素对乳腺组织和乳腺癌的效应。有越来越多的证据表明，这些抗雌激素疗法也可能在高危个体中在乳腺癌的化学防护方面有作用。

**免疫疗法** Trastuzumab (Herceptin)是一种新的免疫治疗药，它与称为 c-erbB2/HER2/neu 的生长因子受体结合，所述受体以少量存在于正常乳腺细胞表面，而在某些乳腺癌中以高得多的水平存在。这种蛋白可以使癌生长并且扩散更快。Herceptin 可以阻止 c-erbB2/HER2/neu 蛋白对乳腺癌细胞生长的促进作用。它也可能帮助免疫系统更好地攻击癌症。目前在标准激素疗法或化学疗法不再有效之后开始 Herceptin (American Cancer Society, 2000, USA, [www.cancer.org](http://www.cancer.org))。

### 治疗挑战

乳腺癌疗法的主要挑战是提高早期检出率、发现新的可以用来跟踪疾病进程并鉴定复发的非侵入性标记、以及发现改善且毒性较小的疗法，尤其对于 5 年存活率仍非常差的更晚期疾病而言。非常需要鉴定出对于癌细胞更具特异性的靶，理想的是在肿瘤细胞表面表达的靶，使得它们可以用有前景的新方法如免疫治疗药和靶向毒素来攻击。

### **发明概述**

本发明提供用于乳腺癌筛选、诊断、预后和治疗、用于监测乳腺癌治疗疗效以及用于乳腺癌治疗药物开发的方法和组合物。

我们已经应用质谱法鉴定出通过实验室培养的人乳腺细胞系膜蛋白提取物的凝胶电泳和胰蛋白酶消化产生的肽。将肽序列与现有的

cDNA 数据库进行了比较, 鉴定出相应的基因序列。这些序列中的许多序列先前在乳腺细胞膜中未曾报道过, 代表一组新的有潜在诊断和/或治疗价值的蛋白质。

因此, 本发明的第一个方面提供乳腺癌的诊断方法, 所述方法包括通过单向电泳测定至少一种乳腺癌相关膜蛋白(BCMP) (例如本文公开的一种或多种 BCMP 或其任何组合), 分析乳腺组织样品。这些方法也适用于筛选、预后、治疗结果的监测、药物开发和用于药物治疗的新靶的发现。

本发明的第二个方面提供治疗乳腺癌的方法, 所述方法包括给予患者治疗有效量的一种化合物, 所述化合物调节(例如正调节或负调节)或者互补乳腺癌患者体内 BCMP 的表达或生物活性(或这两者), 以便(a)防止乳腺癌发作或发生; (b)防止乳腺癌发展; 或(c)缓解乳腺癌的症状。

本发明的第三个方面提供调节(例如正调节或负调节) BCMP 表达或生物活性的化合物的筛选方法。

本发明的第四个方面提供能够与 BCMP (例如本文公开的 BCMP) 免疫特异性结合的单克隆抗体和多克隆抗体。

因此, 在第五个方面, 本发明提供一种人类受治疗者乳腺癌的筛选和/或诊断方法, 所述方法包括下述步骤: 鉴定在得自所述人类受治疗者的生物样品中本文表 1 和表 2 中限定的一种或多种 BCMP 是否存在。

第六个方面, 本发明提供一种人类受治疗者中乳腺癌治疗的监测和/或评价方法, 所述方法包括下述步骤: 鉴定在得自所述人类受治疗者的生物样品中本文表 1 或表 2 中限定的一种或多种 BCMP 是否存在。

第七个方面, 本发明提供一种鉴定在得自所述人类受治疗者的生物样品中是否存在转移性乳腺癌细胞的方法, 所述方法包括下述步骤: 鉴定本文表 1 或表 2 中限定的一种或多种 BCMP 是否存在。

第八个方面，本发明提供一种人类受治疗者中乳腺癌治疗的监测和/或评价方法，所述方法包括下述步骤：鉴定在得自患者的生物样品中本文表1或表2中限定的一种或多种 BCMP 是否增加/减少。

所用的生物样品可以来自任何来源，例如血清样品或组织样品，例如乳腺组织。例如，当寻找转移性乳腺癌的证据时，人们可以考虑乳腺转移癌的主要部位，例如淋巴结、肝、肺和/或骨。

最好是，本发明的方法不基于检查表1和表2中限定的所有 BCMP 的存在与否，而是着眼于“类(cluster)”或其组。

下文和本文的权利要求书中叙述了本发明的其它方面。

### 附图简述

图1显示了 BCMP 81 的核苷酸序列和预测的氨基酸序列。指定至所预测蛋白的质谱以粗体和下划线或虚线下划线表示。串联质谱以粗体和斜体显示。

图2显示了 BCMP 81 的基因组结构，显示出所鉴定的胰蛋白酶肽的位置。所述肽(GDAEKPEEELEEDDDEELDETLSEK)跨越外显子1和外显子2。

图3显示了 BCMP 81 mRNA 的组织分布。正常组织、乳腺癌细胞系和前列腺癌细胞系中 mRNA 的水平通过实时 RT-PCR 定量。mRNA 水平以 ng-1 cDNA 的拷贝数表示。表达水平非常低，平均小于 150 个拷贝/ng cDNA。

图4(a)显示了 BCMP 11 的核苷酸序列和预测的氨基酸序列。预测的 N 末端信号序列以下划线显示。指定至所预测蛋白的质谱以粗体和下划线显示。串联质谱以粗体和斜体表示。图4(b)显示了与 hAG-2 的氨基酸序列(下部) (Thompson, D.A.和 Weigel, R.J.)对齐的 BCMP 11 的预测的氨基酸序列(上部)。hAG-2 是有爪蟾蜍(*Xenopus laevis*)胶粘腺基因 XAG-2 的人类同系物，在乳腺癌细胞系中与雌激素受体共表达。Biochem. Biophys. Res. Commun. 251, 111-116 (1998)。通过垂直的实线

指出相同的氨基酸。

图 5 显示了 BCMP 11 mRNA 的组织分布。正常组织和乳腺癌细胞系中 mRNA 的水平通过实时 RT-PCR 定量。mRNA 水平以 ng-1 cDNA 的拷贝数表示；

图 6 显示了 BCMP 84 的核苷酸序列和预测的氨基酸序列。串联质谱以粗体和斜体显示。MALDI 质谱以粗体和下划线表示；

图 7 显示了 BCMP 84 mRNA 的组织分布。正常组织和乳腺癌细胞系中 mRNA 的水平通过实时 RT-PCR 定量。mRNA 水平以 ng-1 cDNA 的拷贝数表示；

图 8 显示了各种组织中 BCMP 7 的 mRNA 表达水平；

图 9 显示了各种组织中 BCMP 17 的 mRNA 表达水平；且

图 10 显示了各种组织中 BCMP 23 的 mRNA 表达水平。

## 发明详述

以下详细描述的本发明提供了用于以下目的的方法和组合物：用于哺乳动物受治疗者乳腺癌的临床筛选、诊断和预后；用于鉴定最有可能对特定治疗性治疗有反应的患者；用于监测乳腺癌治疗的结果；用于药物筛选和药物开发。本发明也包括将治疗组合物给予哺乳动物受治疗者，以治疗或预防乳腺癌。所述哺乳动物受治疗者可以是非人类哺乳动物，但优选为人，更优选为成人，即至少 21 岁(更优选至少 35 岁、至少 50 岁、至少 60 岁、至少 70 岁或至少 80 岁)的人类受治疗者。为了阐明公开内容并且并不作为限制，本发明将在乳腺组织样品的分析方面进行描述。然而，本领域技术人员会认识到，下述的测定和技术可以适用于其它类型的患者样品，包括体液(例如血液、血清、血浆或唾液)、得自有患乳腺癌危险的患者的组织样品(例如活检样品，例如乳腺组织活检样品)或其匀浆。本发明的方法和组合物尤其适用于活受治疗者的筛选、诊断和预后，但也可以用于受治疗者的尸体剖检诊断，例如以鉴定有发生相同疾病危险的家族成员。

本文所用的乳腺组织是指乳腺本身以及乳腺下层的邻近组织和/或所述层内的组织。

### 乳腺癌相关膜蛋白(BCMP)

在本发明的一个方面,应用单向电泳分析得自受治疗者(最好是活受治疗者)的乳腺组织,以便测量一种或多种乳腺癌相关膜蛋白(BCMP)的表达,以筛选或诊断乳腺癌,以决定乳腺癌患者的预后、以监测乳腺癌治疗的疗效、或用于药物开发。本文所用的“单向电泳”(1D-电泳)是指包括变性电泳的技术;这产生含有多种经分离的蛋白的单向凝胶(1D-凝胶)。最好是,变性电泳步骤应用十二烷基硫酸钠存在下的聚丙烯酰胺电泳(SDS-PAGE)。尤其优选在 WO 98/23950 中描述的高度准确的自动化方法和装置(“the Preferred Technology”),该文献通过引用全部结合到本文中,尤其参考第 19-29 页的优选方案。简而言之,the Preferred Technology 提供了用于鉴定、选择和特征鉴定生物样品中生物分子的有效计算机辅助方法和装置。通过将生物分子在单向凝胶上根据其电泳迁移率分离,产生单向排列。产生一个所述排列的计算机产生的数字分布,代表在所述单向排列中检出的多种生物分子的身份、表观分子量,从而允许对得自多种生物样品的分布进行计算机介导的比较以及所分离的目的蛋白的计算机辅助的切取。

用于检测荧光标记蛋白的优选扫描器在 WO 96/36882 和 David A. Basiji 的博士论文(题为“利用内部反射光学和相位灵敏检测的高通量荧光扫描器的开发(全内部反射,电泳)”, University of Washington (1997), Dissertation Abstracts International 的第 58/12-B 卷,第 6686 页中有描述,所述文献的内容通过引用结合到本文中。这些文件描述了为高速自动综合操作特别设计的图象扫描器。所述扫描器可以将已经用荧光染料或银染料染色的凝胶成像以及存储荧光屏。Basiji 的论文提供了一个相位灵敏检测系统,以供分辨经调制的荧光与由于激光散射所致的基线噪声或均匀荧光,但所述扫描器也可以以非相位灵敏方式进

行操作。与传统的荧光成像系统相比，这种相位灵敏检测能力能够将仪器的灵敏度提高一个数量级或更多。灵敏度提高，将减少上游仪器的样品制备上样量，而图象质量的提高简化了该方法中下游的图象分析。

一种更为高度优选的扫描器是 Apollo 2 扫描器 (Oxford Glycosciences, Oxford, UK)，这是一种上述扫描器的改进形式。在 Apollo 2 扫描器中，凝胶在精密引导螺杆驱动系统上传送通过扫描器。这最好将玻璃板置于 Basiji 论文中所述的皮带传动系统上，因为它提供一种准确传送凝胶通过成像光学系统的可再现方法。

在 Apollo 2 扫描器中，凝胶用三个对准止动件固定，将玻璃板大致保持在已知位置。借此并且连同上述精密传送系统，可以预测并记录凝胶的绝对位置。这确保可以更准确地测量凝胶上每个特征(feature)的坐标，如果需要，将其传送至切割机械人，以切取所述特征。在 Apollo 2 扫描器中，承载凝胶的载体有 4 个一体化荧光标记，用来校正图象的几何形状。这些标记是证实扫描正确进行的质量控制特征。

与 Basiji 论文中所述的扫描器相比，Apollo 2 扫描器的光学部件倒置。在 Apollo 2 扫描器中，激光、反射镜、波导和其它光学部件在被扫描的玻璃板之上。在 Basiji 论文中所述的扫描器中，这些部件在所述玻璃板之下。在 Apollo 2 扫描器中，玻璃板固定在扫描器上，凝胶面向下，使得光路仍通过玻璃板。借此，离开玻璃板的任何凝胶颗粒将落在仪器的底部，而不是落在光学器件上。这不影响该系统的功能性，而提高了其可靠性。

再更优选的是 Apollo 3 扫描器，其中信号输出被数字化为全 16 比特数据，而没有任何峰饱和，或者没有信号的平方根编码。补偿算法也已经用来校正沿扫描束路径的检测灵敏度的任何变化。这种变化是光学系统的异常和波导两端收集效率的差异所致。用具有均匀荧光通量的有机玻璃板进行校准。从该板扫描接受的数据用来确定将所述信号从每个像素电平提高至靶电平所需的放大系数。这些放大系数然

后用于随后的凝胶扫描，以除去任何内部光学变化。

本文所用术语“乳腺癌相关膜蛋白”(BCMP)是指通过乳腺组织样品的 1D 电泳可检测的、存在于得自乳腺癌受治疗者的乳腺组织中的一个特征(例如 1D 凝胶中的一个条带)。

已经在实验室培养的人乳腺细胞系的膜蛋白提取物中，通过 the Preferred Technology 的方法和装置(一般为 1D 凝胶电泳和实验室培养的人乳腺细胞系膜蛋白提取物的胰蛋白酶消化)，鉴定出本文公开的 BCMP。将肽序列与 SWISS-PROT 和 trEMBL 数据库(the Swiss Institute of Bioinformatics (SIB)和 the European Bioinformatics Institute (EBI)维护，它们可在<http://www.expasy.com/>得到)和 GenBank 数据库(由 the National Institute of Health (NIH) 维护，它可在<http://www.ncbi.nlm.nih.gov/GenBank/>得到)进行了比较，鉴定出相应的基因。搜索了已发表报道和数据库，包括在正常人乳腺腔上皮细胞中表达的蛋白质数据库(Page 等, *Proc Natl Acad Sci USA*. 1999 96(22):12589-94; 英国专利申请号 9919258.5)，以确立所述鉴定的基因中的任一种的产物是否先前已经证明在人乳腺细胞膜或人乳腺癌细胞膜中表达。鉴定了两组 BCMP: (1) 匹配未曾描述过蛋白质或生物功能、本发明确定所述蛋白产物的存在及其在人乳腺癌细胞膜中定位的 cDNA 的概念翻译的蛋白质序列; (2) 先前未在乳腺细胞膜方面描述过的已知蛋白，本发明表明所述蛋白可能还涉及人乳腺癌。所有这些 BCMP 都可用作乳腺细胞、尤其是乳腺癌细胞的标记。属于第一类的那些蛋白的氨基酸序列示于以下表 1，属于第二类的那些蛋白的氨基酸序列示于以下表 2。表 1 和表 2 中的每种蛋白用 Swiss Prot 或 GenBank 登记号来识别，每种序列通过引用结合到本文中。在这些表中也列出了通过串联质谱和数据库搜索(例如实施例描述的，参见下文)鉴定的这些 BCMP 的胰蛋白酶消化肽的表观分子量和氨基酸序列。

表 1: 匹配未曾描述过蛋白质或生物功能的 cDNA 的概念翻译的蛋白质序列

BCMP #	MW (kDa)	胰蛋白酶消化肽的氨基酸序列	所识别序列的登记号
1	35.8	ELFPIQMEGVK, FVNWQVDGEYR	AF050154
2	55.6	QGHLSLQR	U93562
3	31.7	YQETFNVIER	Q13011
4	19.3	INPDGSQSVVEVPYAR	NP_055070
5	48.9	QQQLNEENLR	NP_060658
6	52.7	NVDLSTFYQNR	NP_055635
8	26.5	YDDAIQLYDR	Q15006
9	25.1	VGDVVEFR, EDLEELQAR	AAF17227
10	49.0	LENGEIETIAR	CAB75499
12	106.3	SGAGVPAVILR	Q92542
13	32.8	SNPEDQILYQTER	Q14165
14	46.7	VGAENVAIVEPSER	AF042284
15	14.5	NGNQAFNEDNLK, FILMDCMEGR	AF070659
16	62.5	FDGILTEGEGPR	AF081886

表 2: 先前未在乳腺细胞膜方面描述过的已知蛋白

BCMP #	MW (kDa)	胰蛋白酶消化肽的氨基酸序列	所识别序列的登记号
17	15.0	LQDASAEVER	Q10589
18	28.9	LSCAYSGFSSPR,	NP_058642

BCMP #	MW (kDa)	胰蛋白酶消化肽 的氨基酸序列	所识别序列的登记号
		FDQGDTR, ITASYEDR, IPENNPVK	
19	20.7	QLVEQVEQIQK, FSLFAGGMLR	CAA62380
20	16.4	IPDQLVILDMK, LKPLEVELR	P49755
21	25.6	HDWINR, AYEYVECPIR, DYQFTILDVR, SWMGYELPFDR	P53701
22	12.9	TEDLEATSEHFK	AF053233
23	105.0	EVVSPQEFK, NINSDGGPYVCR, AVDGFTFTEGDK	O15394
24	34.3	IEEACEIYAR	P54920
25	48.7	SPQQVLPSPDGR	P22570
26	14.7	ELAPLFEELR	AAF17239
27	26.0	STIGVEFATR	P24410
28	26.0	DEFLIQASPR	P51149
29	28.3	DNEGSWFR	NP_003625
30	25.6	LQIWDTAGQER, TITSSYYR, QWLQEIDR	P11476
31	25.6	TQIDHYVGIAR	NP_006045
32	46.1	QIAVEAQEILR	AAA65729
33	26.0	FLTQPQVVAR	O43760
34	41.9	VAEDEAEAAAAAK	P08195
35	127.0	IQTQPGYANTLR	Q00325
36	42.5	VVVTVEQTEELER	O15269

BCMP #	MW (kDa)	胰蛋白酶消化肽 的氨基酸序列	所识别序列的登记号
37	33.5	QGYVLSSIEGR	O43684
38	23.7	TQLQLDER	P53007
39	26.1	EVSFQSTGESEWK	Q07021
40	26.8	LSELQLR, QLVEQVEIQK	X90872
41	31.9	DNVDDPTGNFR	AAA19572
42	81.8	VEEVGPYTYR	Q14108
43	42.8	ATVLESEGTR, DVQGT DASLDEELDR	NP_038470
44	34.4	FNASQLITQR	AAF44345
45	34.4	HSEIQQLER, NILSSADYVER	Q12846
46	15.3	DLQQYQSQAK	AF047442
47	27.8	AEEWGVQYVETSAK, QVPVEEAR, VFFDLMR	P11234
48	29.4	DLECVTNLQEVAR	P10619
49	25.8	WPYAGTAEAIK, EVVEAHVDQK	P30042
50	22.9	YDPTIEDSYR	P10113
51	30.6	HLIPAANTGESK, YLAEFATGNDR	P42655
52	24.9	EQNSPIYISR, VLQSEFCNAVR	AK002077
53	22.1	YLECSALTQR	P15154

以上表1和表2中描述的BCMP可以单独应用或者与以下更详细提及的乳腺癌诊断和治疗联合应用。另外,这些BCMP中的一种或多种可以与以下表3中所述蛋白质中的一种或多种联合应用。

表3: 可以与表1和表2的蛋白质联合应用的蛋白质

BCMP #	MW (kDa)	胰蛋白酶消化肽 的氨基酸序列
7	17.9	IMFVDPSLTVR, HLSPDGQYVPR
11	16.6	NLSPDGQYVPR
81	20.6	LQMEQQQLQQR
84	13.6	SANAEDAQEFSDVER

在 WO98/07749 中公开了 BCMP 7 和 BCMP 11 的序列。如本文实施例 3 中所述克隆了 BCMP 11。BCMP81 的序列公开于 WO99/18202，而 BCMP84 的序列公开于 WO99/47669。

对于任何给定的 BCMP 而言，相对于分析得自无乳腺癌受治疗者的乳腺组织时获得的检测水平而言，分析得自乳腺癌受治疗者的乳腺组织时获得的检测水平将取决于所用的具体分析方案和检测技术，条件是这类 BCMP 在正常组织和疾病组织之间差异表达。因此，本发明考虑到，每个实验室将根据应用中的分析方案和检测技术，为无乳腺癌受治疗者中的每种 BCMP 建立一个参比范围，在诊断领域通常也是如此。最好是，在每批分析的试验样品中，包括至少一个得自己知有乳腺癌的受治疗者的对照阳性乳腺组织样品、或者至少一个得自己知无乳腺癌的受治疗者的对照阴性乳腺组织样品(更优选既包括阳性对照样品，也包括阴性对照样品)。

在一个实施方案中，测定相对于本底值的蛋白质表达水平，本底值定义为得自图象邻近区的信号水平，所述图象(a)以面积计等于所研究的特定特征，和(b)不含可检出的蛋白质特征。

BCMP 可以用于乳腺癌的检测、预后、诊断或监测或者用于药物开发。在本发明的一个实施方案中，通过 1D 电泳分析得自受治疗者(例如怀疑患有乳腺癌的受治疗者)的乳腺组织，以检测表 1 和表 2 中的一

种或多种 BCMP。相对于一个或多个无乳腺癌受治疗者的乳腺组织(例如,对照样品)或者先前测定的参比范围而言得自所述受治疗者的乳腺组织中所述一种或多种 BCMP 丰度的减少,指示乳腺癌的存在与否。

在一个优选实施方案中,分析得自受治疗者的乳腺组织,以定量测定表 1、2 和 3 中的多类 BCMP。

正如本领域技术人员将显而易见的,可以按照为表 1 和表 2 中的 BCMP 提供的数据,描述给定的 BCMP。所述 BCMP 是一种包含有关该 BCMP 所述的肽序列(最好包含多个、更优选所有有关该 BCMP 所述的肽序列)的蛋白质。

在一个实施方案中,分析得自受治疗者的乳腺组织,以定量测定表 1 和表 2 中 BCMP 中的一种或多种或其任何组合,其中相对于得自一个或多个无乳腺癌受治疗者的乳腺组织(例如对照样品或先前测定的参比范围)而言,得自所述受治疗者的乳腺组织中所述一种或多种 BCMP (或其任何组合)丰度的改变指示存在乳腺癌。

在一个优选实施方案中,分析得自受治疗者的乳腺组织,以定量测定表 1、表 2 和表 3 中的多类 BCMP。

对于每种 BCMP 而言,本发明还提供:(a) 包含所述分离的 BCMP 的制剂;(b) 包含所述 BCMP 的一个或多个片段的制剂;和(c) 与所述 BCMP 结合、与所述片段结合或与所述 BCMP 并且与所述片段结合的抗体。本文所用的 BCMP 是“分离的”,当它存在于制剂中时,基本上不含污染蛋白,即其中低于 10% (优选低于 5%,更优选低于 1%) 的所存在的总蛋白为污染蛋白的制剂。污染蛋白是通过质谱分析测定,其氨基酸序列显著不同于所述分离的 BCMP 的氨基酸的蛋白质。本文所用的“显著不同的”序列是通过按照 the Reference Protocol 进行的质谱分析,使得能够分辨所述污染蛋白质与所述 BCMP 的序列。

本发明的 BCMP 可以通过本领域技术人员已知的任何方法分析,所述方法包括但不限于本文所述的 the Preferred Technology、激酶测定、酶测定、结合测定和其它功能分析、免疫测定和蛋白质印迹法。

在一个实施方案中,根据其 MW 在 1-D 凝胶上分离所述 BCMP,并且通过将凝胶染色来显现。在一个实施方案中,所述 BCMP 用荧光染料染色,并用荧光扫描器成像。Sypro Red (Molecular Probes, Inc., Eugene, Oregon)是一种用于此目的合适染料。一种优选的荧光染料公开于 1999 年 10 月 5 日申请的美国申请第 09/412,168 号,该申请通过引用全部结合到本文中。

或者,可以用免疫测定检测 BCMP。在一个实施方案中,如下进行免疫测定:通过使得自受试受治疗者的样品与抗 BCMP 抗体接触,其接触条件使得当所述 BCMP 存在时可以发生免疫特异性结合;并且检测或测量所述抗体的任何免疫特异性结合的量。抗 BCMP 抗体可以通过本文所述的方法和技术来产生。

在一个实施方案中,组织切片中的抗体结合可以用来检测异常的 BCMP 定位或一种或多种 BCMP 的异常水平。在一个具体实施方案中,抗 BCMP 抗体可以用来分析患者组织(例如乳腺组织活检样品)中所述 BCMP 的水平,其中 BCMP 的异常水平指示乳腺癌。本文所用的“异常水平”是指与无乳腺癌受治疗者的水平或参比水平相比增加或减少的水平。如果需要,可以与得自同一受治疗者、取自不受乳腺癌影响的机体部分的匹配样品进行比较。

可以使用任何合适的免疫测定,包括但不限于利用诸如以下技术的竞争性测定系统和非竞争性测定系统:蛋白质印迹法、放射免疫测定、ELISA (酶联免疫吸附测定)、“夹心”免疫测定、免疫沉淀测定、沉淀素反应、凝胶扩散沉淀素反应、免疫扩散测定、凝集测定、补体结合测定、免疫放射分析、荧光免疫测定和 A 蛋白免疫测定。

例如,可以通过两步夹心测定,测定体液样品(例如血液、尿或乳腺组织匀浆)中的 BCMP。在第一步中,用捕获试剂(例如抗 BCMP 抗体)来捕获所述 BCMP。所述捕获试剂可以任选地固定化于固相上。在第二步中,用直接或间接标记的检测试剂来检测所捕获的 BCMP。在一个实施方案中,所述检测试剂是凝集素。与具有与所述 BCMP 相同

核心蛋白的其它同种型相比或与共享所述抗体所识别的抗原决定簇的其它蛋白相比, 优先与所述 BCMP 结合的任何凝集素, 都可以用于此目的。在一个优选实施方案中, 所选凝集素与所述 BCMP 结合的亲合性比具有与所述 BCMP 相同核心蛋白的所述其它同种型或与共享所述抗体所识别的抗原决定簇的所述其它蛋白高至少 2 倍、更优选高至少 5 倍, 再更优选高至少 10 倍。根据本文的描述, 适用于检测给定 BCMP 的凝集素可以用本领域众所周知的方法容易地鉴定, 例如通过测试 Sumar 等, 作为疾病相关糖形指示的凝集素, Gabius H-J 和 Gabius S (编著), 1993, *Lectins and Glycobiology*, 第 158-174 页(该文献通过引用全部结合到本文中)的第 158-159 页表 I 中列举的一种或多种凝集素。在一个可供选择的实施方案中, 所述检测试剂是一种抗体, 例如免疫特异性地检测其它翻译后修饰的抗体, 诸如免疫特异性地与磷酸化氨基酸结合的抗体。这类抗体的实例包括与磷酸酪氨酸结合的抗体(BD Transduction Laboratories, 产品目录号: P11230-050/P11230-150; P11120; P38820; P39020)、与磷酸丝氨酸结合的抗体(Zymed Laboratories, Inc., South San Francisco, CA, 产品目录号: 61-8100)和与磷酸苏氨酸结合的抗体(Zymed Laboratories, Inc., South San Francisco, CA, 产品目录号: 71-8200, 13-9200)。

如果需要, 在杂交分析中也可以使用编码 BCMP 的基因、相关基因或相关核酸序列或亚序列, 包括互补序列。编码 BCMP 的核苷酸或其包含至少 8 个核苷酸、优选至少 12 个核苷酸、最优选至少 15 个核苷酸的亚序列, 可以用作杂交探针。杂交分析可以用于与 BCMP 编码基因异常表达相关的病症、疾病或病状的检测、预后、诊断或监测, 或用于具有暗示乳腺癌的体征或症状的受治疗者的鉴别诊断。具体而言, 这样一种杂交分析可以通过这样一种方法进行, 所述方法包括使受治疗者的含有核酸的样品与能够与编码 BCMP 的 DNA 或 RNA 杂交的核酸探针在使得可以发生杂交的条件下接触, 并且检测或测量任何所产生的杂交。核苷酸可以用于乳腺癌受治疗者的治疗, 如下所述。

本发明也提供诊断试剂盒,所述诊断试剂盒包含一种抗 BCMP 抗体。另外,这样一种试剂盒可以任选地包含以下的一种或多种:(1)关于使用所述抗 BCMP 抗体进行诊断、预后、治疗监测或这些应用任何组合的说明;(2)所述抗体的标记结合配偶体;(3)抗 BCMP 在其上固定化的固相(例如试剂纸条);和(4)标签或插页,表明管理机构批准用于诊断、预后或治疗应用或其任何组合。如果没有提供所述抗体的标记结合配偶体,则所述抗 BCMP 抗体本身可以用可检测标记(例如化学发光部分、酶部分、荧光部分或放射性部分)标记。

本发明也提供一种包含能够与编码 BCMP 的 RNA 杂交的核酸探针的试剂盒。在一个具体实施方案中,试剂盒在一个或多个容器中包含一对引物(例如每种引物的大小范围为 6-30 个核苷酸,更优选 10-30 个核苷酸,再更优选 10-20 个核苷酸),所述一对引物在合适的反应条件下可以引发编码 BCMP 的核酸的至少一部分扩增,例如通过聚合酶链式反应(参见例如 Innis 等,1990,PCR Protocols, Academic Press, Inc., San Diego, CA)、连接酶链式反应(参见 EP 320,308)、应用 Q $\beta$ 复制酶、环状探针反应或本领域已知的其它方法。

也提供允许检测多种 BCMP 或分别编码一种 BCMP 的多种核酸的试剂盒。试剂盒可以任选地还包含预定量的一种分离 BCMP 蛋白或编码一种 BCMP 的一种核酸,例如用作标准或对照。

### 临床研究方面的应用

本发明的诊断方法和组合物可以有助于监测临床研究,例如以评价用于治疗乳腺癌的药物。在一个实施方案中,测试候选分子将乳腺癌受治疗者的 BCMP 水平恢复至无乳腺癌受治疗者中的水平的能力,或者测试候选分子在经治疗的受治疗者体内(例如在用紫杉酚或多柔比星治疗后)将 BCMP 水平保持在无乳腺癌数值水平或其附近的能力。可以分析一种或多种 BCMP 的水平。

在另一实施方案中,本发明的方法和组合物可以用来筛选可供临

床研究以鉴定患乳腺癌个体的候选者；然后将所述个体从该项研究中排除，或者可以将其置于单独的群组，以进行治疗或分析。如果需要，可以同时筛选所述候选者，以鉴定患乳腺癌个体；用于这些筛选的方法是本领域众所周知的。

### BCMP 的纯化

在许多特定方面，本发明提供分离的 BCMP (最好是人 BCMP) 以及包含抗原决定簇(即可为抗体识别)或者具有其它功能活性的其片段和衍生物、以及编码前述 BCMP、片段和衍生物的核酸序列。本文所用的“功能活性”是指物质显示出与全长(野生型) BCMP 相关的一种或多种功能活性，例如与 BCMP 底物或 BCMP 结合配偶体结合、抗原性(与抗靶抗体结合)、免疫原性、酶活性等。

在具体的实施方案中，本发明提供 BCMP 的肽片段，所述肽片段包含至少 5 个氨基酸、至少 10 个氨基酸、至少 50 个氨基酸或至少 75 个氨基酸。也提供缺乏 BCMP 的某些区或所有区的片段，也提供包含这类片段的蛋白质(例如融合蛋白)。提供编码上述肽片段或上述蛋白质的核酸。

一旦鉴定出编码所述 BCMP、所述 BCMP 的一部分或所述 BCMP 前体的重组核酸，则可以分析所述基因产物。这可以通过基于所述产物的物理特性或功能特性的测定来达到，所述测定包括所述产物的放射性标记，然后进行凝胶电泳、免疫测定等分析。

本文鉴定的 BCMP 可以用标准方法分离和纯化，所述方法包括色谱(例如离子交换色谱、亲和色谱和大小柱色谱)、离心、差别溶解性或通过用于蛋白质纯化的的任何其它标准技术。

或者，一旦鉴定出编码所述 BCMP 的重组核酸，则可以从所述重组核酸中包含的所述基因编码区的核苷酸序列，推导出所述 BCMP 的完整氨基酸序列。结果，可以通过本领域已知的标准化学方法，合成所述蛋白质(参见例如 Hunkapiller 等, 1984, Nature 310:105-111)。

在另一个可供选择的实施方案中，可以通过标准方法例如上述方法(例如免疫亲和纯化)，从天然来源中纯化天然 BCMP。

因此，本发明提供分离的 BCMP、分离的 BCMP 相关多肽和 BCMP 或 BCMP 相关多肽的分离的衍生物或片段；任何上述物质都可以通过重组 DNA 技术或化学合成方法来生产。

### 编码 BCMP 的 DNA 的分离

以下介绍用于克隆 BCMP 编码基因的具体实施方案，所述实施方案仅作为实例，而非限制性的。

本发明的核苷酸序列(包括 DNA 和 RNA 并且包含编码 BCMP 或其片段或 BCMP 相关多肽的序列)可以采用本领域已知的方法合成，例如采用常规的化学方法或聚合酶链式反应(PCR)扩增。本发明的核苷酸序列也使得有可能鉴定和克隆编码 BCMP 同系物或 BCMP 直向同系物的基因，包括例如通过筛选 cDNA 文库、基因组文库或表达文库。

例如，为了通过 PCR 技术克隆编码 BCMP 的基因，可以为鉴定为 BCMP 蛋白部分的所有 BCMP 肽片段设计锚定简并寡核苷酸(或一组最有可能寡核苷酸)。可以用得自一种或多种物种的相关 cDNA 和基因组 DNA (例如得自乳腺组织或得自免疫系统细胞)，进行多种条件下的 PCR 反应。也可以用上述寡核苷酸(最好是嵌套式的)，在任何可得到的 cDNA 和基因组 DNA 上进行 vectorette 反应。vectorette PCR 是在仅已知一种引物序列的情况下使特定 DNA 片段能够扩增的一种方法。因此，它将 PCR 的应用延伸至仅可获得一端的序列信息的 DNA 序列段。(Arnold C, 1991, PCR Methods Appl. 1(1):39-42; Dyer KD, Biotechniques, 1995, 19(4):550-2)。可以用例如为编码 BCMP 肽片段的锚定简并寡核苷酸(或最可能的寡核苷酸)的探针，用基因组文库或 cDNA 文库的库(pool)作为模板，进行 vectorette PCR。

可以为所有 BCMP 肽片段设计锚定简并寡核苷酸(和最可能的寡核苷酸)。这些寡核苷酸可以被标记，并且与含有 cDNA 文库和基因组

DNA 文库的滤膜杂交。针对得自同一蛋白的不同肽的寡核苷酸将通常鉴定所述文库的相同成员。cDNA 文库和基因组 DNA 文库可以得自任何合适或所需的哺乳动物物种，例如得自人类。

包含编码本发明 BCMP 或 BCMP 片段的核苷酸序列的核苷酸序列，尤其它们能够与编码其它蛋白的基因的互补序列段选择性杂交，因而是有用的。根据应用，可以使用各种各样的杂交条件，以获得与编码 BCMP 的核苷酸序列至少 30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、99% 或 100% 相同的核苷酸序列。

对于高度的选择性，用相对严格的条件形成双链体，例如低盐或高温条件。本文所用的“高度严格条件”是指在 0.5 M NaHPO<sub>4</sub>、7% 十二烷基硫酸钠(SDS)、1 mM EDTA 中于 65°C 下与滤膜结合的 DNA 杂交，并且在 0.1 × SSC/0.1% SDS 中于 68°C 洗涤(Ausubel, F.M.等编著, 1989, Current Protocols in Molecular Biology, 第 I 卷, Green Publishing Associates, Inc., 和 John Wiley & Sons, Inc., New York, 第 2.10.3 页; 所述文献通过引用全部结合到本文中)。对于某些应用，需要较低严格条件来形成双链体。本文所用的“中等严格条件”是指在 0.2 × SSC/0.1% SDS 中于 42°C 洗涤(Ausubel 等, 1989, 参见上文)。通过提高甲酰胺的量以使杂种双链体不稳定，也可以使得杂交条件更为严格。因此，可以容易地操作具体的杂交条件，一般根据所需结果来选择。一般而言，在 50% 甲酰胺存在下方便的杂交温度是：对于与编码 BCMP 的基因片段有 95-100% 同一性的探针而言为 42°C，对于有 90-95% 同一性的探针而言为 37°C，而对于有 70-90% 同一性的探针而言为 32°C。

在基因组文库的制备中，产生 DNA 片段，其中某些 DNA 片段将编码 BCMP 的部分或完全的 BCMP。用于制备 DNA 片段的任何合适的方法都可以用于本发明。例如，可以用各种限制性酶在特定位点切割所述 DNA。或者，人们可以在锰存在下使用 DNA 酶，使 DNA 片段化，或者可以例如通过超声处理，物理剪切所述 DNA。然后，用标

准技术, 包括但不限于琼脂糖凝胶电泳和聚丙烯酰胺凝胶电泳、柱色谱和蔗糖梯度离心, 可以根据大小分离所述 DNA 片段。然后将所述 DNA 片段插入到合适的载体中, 所述载体包括但不限于质粒、粘粒、噬菌体 $\lambda$ 或 T<sub>4</sub> 和酵母人工染色体(YAC)。(参见例如 Sambrook 等, 1989, Molecular Cloning, A Laboratory Manual, 第 2 版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M.(编著), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. 第 I、II 卷; Ausubel F.M. 编著, 1989, Current Protocols in Molecular Biology, 第 I 卷, Green Publishing Associates, Inc., 和 John Wiley & Sons, Inc., New York)。可以通过与标记探针进行核酸杂交, 来筛选基因组文库(Benton 和 Davis, 1977, Science 196:180; Grunstein 和 Hogness, 1975, Proc. Natl. Acad. Sci. U.S.A. 72:3961)。

根据本发明的描述, 可以用本领域众所周知的最佳方法, 用对应于所述 BCMP 任何肽的氨基酸序列的标记简并寡核苷酸探针, 筛选基因组文库。所用的任何探针至少长 10 个核苷酸、至少 15 个核苷酸、至少 20 个核苷酸、至少 25 个核苷酸、至少 30 个核苷酸、至少 40 个核苷酸、至少 50 个核苷酸、至少 60 个核苷酸、至少 70 个核苷酸、至少 80 个核苷酸或至少 100 个核苷酸。探针优选长 10 个核苷酸或更长, 更优选长 15 个核苷酸或更长。

在以上表 1 和表 2 中, 发现本文公开的某些 BCMP 对应于先前鉴定的由公众已知其序列的基因编码的蛋白质的同种型。(用实施例中所描述的方法进行 BCMP 的序列分析和蛋白质鉴定)。为了筛选这样一种基因, 可以使用与所述基因或其互补物互补的任何探针; 所述探针优选长 10 个核苷酸或更长, 更优选长 15 个核苷酸或更长。SWISS-PROT 和 trEMBL 数据库(由 the Swiss Institute of Bioinformatics (SIB)和 the European Bioinformatics Institute (EBI) 维护, 它们可在 <http://www.expasy.ch/>得到)和 GenBank 数据库(由 the National Institute of Health (NIH)维护, 它可在 <http://www.ncbi.nlm.nih.gov/>得到)提供了表 1

和表 2 中所列 BCMP 的蛋白序列，每个序列通过引用结合到本文中。

当筛选文库时，具有编码所述 BCMP 或其片段的插入 DNA 的克隆，将与一组对应的简并寡核苷酸探针(或其互补物)的一个或多个成员杂交。用本领域已知的方法进行这类寡核苷酸探针与基因组文库的杂交。例如，可以在如上定义的高度严格或中等严格条件下，进行与上述简并寡核苷酸探针组中的一组探针或其互补物(或该组探针的任何成员或其互补物)的杂交，或者所述杂交可以在  $2 \times \text{SSC}$ 、1.0% SDS 中于  $50^\circ\text{C}$  下进行，并且用上文所述的用于高度严格或中等严格杂交的洗涤条件进行洗涤。

在本发明的再一方面，含有编码完整 BCMP、BCMP 的片段、BCMP 相关多肽或 BCMP 相关多肽的片段的核苷酸序列的克隆，也可以通过筛选表达文库来获得。例如，分离出得自相关来源的 DNA，制备随机片段，并连接到表达载体(例如噬菌体、质粒、噬菌粒或粘粒)中，使得所述载体中所插入的序列能够由所述载体随后所导入的宿主细胞表达。然后，可以用各种筛选测定，根据对所表达的 BCMP 或 BCMP 相关多肽来进行选择。在一个实施方案中，本发明的各种抗 BCMP 抗体可以用来采用本领域已知的方法鉴定所需克隆。参见例如 Harlow 和 Lane, 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 附录 IV。使得自所述文库的菌落或噬斑与所述抗体接触，以鉴定结合抗体的那些克隆。

在一个实施方案中，含有编码 BCMP、BCMP 的片段、BCMP 相关多肽或 BCMP 相关多肽的片段的 DNA 的菌落或噬斑，可以用 DYNA 微珠，按照 Olsvick 等, 29<sup>th</sup> ICAAC, Houston, Tex. 1989 (该文献通过引用结合到本文中)所述来检测。使抗 BCMP 抗体与甲苯磺酰化 DYNA 微珠 M280 交联，然后使这些含抗体的微珠与表达重组多肽的菌落或噬斑接触。表达 BCMP 或 BCMP 衍生多肽的菌落或噬斑鉴定为结合所述微珠的那些菌落或噬斑中的任一种。

或者，可以将所述抗 BCMP 抗体非特异性地固定化至合适的支持

体上，例如二氧化硅或 Celite®树脂。然后这种材料用来吸附至表达本文所述的 BCMP 蛋白或 BCMP 相关多肽的菌落。

另一方面，可以用 PCR 扩增，从基因组 DNA 中分离编码所述完整 BCMP 或其部分的基本纯的 DNA (即基本上不含污染核酸的 DNA)。这种 DNA 的纯度优选至少为 95%，更优选纯度至少为 99%。对应于本文公开的 BCMP 的肽序列的简并的或者其它的寡核苷酸序列，可以用作引物。

可以例如利用 Perkin-Elmer Cetus 热循环仪和 Taq 聚合酶(Gene Amp®或 AmpliTaq DNA 聚合酶)进行 PCR。人们可以选择合成几种不同的简并引物，以供用于 PCR 反应。也有可能改变引发 PCR 反应所用的杂交条件的严格性，以提供所述简并引物和所述 DNA 中对应序列之间更高或更低的核苷酸序列相似性程度。在成功扩增 BCMP 编码序列一个区段后，可以将该区段进行分子克隆并测序，并且用作探针以分离完整的基因组克隆。这又将允许测定所述基因的完整核苷酸序列、其表达的分析以及生产其蛋白产物以供功能分析，如下文所述。

编码 BCMP 的基因也可以通过 mRNA 选择，通过核酸杂交，然后进行体外翻译来鉴定。在这种方法中，利用片段，通过杂交来分离互补的 mRNA。这种 DNA 片段可以代表可得到的、经纯化的、编码另一物种(例如小鼠、人类)的 BCMP 的 DNA。所分离 mRNA 的分离产物的体外翻译产物的免疫沉淀分析或功能分析(例如体外聚集能力；与受体结合)，鉴定出所述 mRNA，并因此鉴定出含有所需序列的互补 DNA 片段。另外，通过使从细胞中分离的多核糖体吸附至特异性识别一种 BCMP 的固定化抗体上，可以选择特定的 mRNA。用所述经选择的 mRNA (得自吸附的多核糖体)作为模板，可以合成编码 BCMP 的放射性标记的 cDNA。然后可以用所述放射性标记的 mRNA 或 cDNA 作为探针，从其它基因组 DNA 片段中鉴定编码 BCMP 的 DNA 片段。

分离编码 BCMP 的基因组 DNA 的替代方法包括但不限于：根据已知序列化学合成所述基因序列本身、或者从编码所述 BCMP 的

mRNA 制备 cDNA。例如，可以从表达 BCMP 的细胞中，分离出 RNA 以供编码所述 BCMP 的基因的 cDNA 克隆。根据本文的描述，本领域技术人员会理解，可以使用其它方法，并且这些方法在本发明范围内。

任何合适的真核细胞都可以作为可供 BCMP 编码基因的分子克隆的核酸来源。可以从脊椎动物、哺乳动物、灵长类、人类、猪、牛、猫、鸟类、马、犬或鼠类来源，分离出编码所述 BCMP 的核酸序列。可以通过本领域已知的标准方法，通过化学合成、cDNA 克隆或基因组 DNA 或其片段的克隆，从由所需细胞中纯化的克隆的 DNA (例如 DNA“文库”)中获得所述 DNA。(参见例如 Sambrook 等, 1989, *Molecular Cloning, A Laboratory Manual*, 第 2 版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M.(编著), 1985, *DNA Cloning: A Practical Approach*, MRL Press, Ltd., Oxford, U.K. 第 I、II 卷)。来源于基因组 DNA 的克隆除编码区外，可能还含有调节 DNA 区和内含子 DNA 区；来源于 cDNA 的克隆将仅含有外显子序列。

然后，可以将所述经鉴定并分离的基因或 cDNA 插入任何合适的克隆载体中。可以使用本领域已知的许多载体-宿主系统。正如本领域技术人员会认识到的，唯一的限制是所选的载体系统与所用的宿主细胞匹配。这类载体包括但不限于：噬菌体例如 $\lambda$ 噬菌体、质粒例如 pBR322 或 pUC 质粒衍生物或 Bluescript 载体(Stratagene)或经修饰的病毒例如腺病毒、腺伴随病毒或反转录病毒。例如通过将所述 DNA 片段连接到具有互补粘性末端的克隆载体中，可以完成在克隆载体的插入。然而，如果用来使所述 DNA 片段化的互补限制位点在所述克隆载体中不存在，那么可以用酶修饰所述 DNA 分子的末端。或者，可以通过将核苷酸序列(接头)连接到所述 DNA 末端，产生所需的任何位点；这些连接的接头可以包含特定的化学合成的编码限制性内切核酸酶识别序列的寡核苷酸。在一个可供选择的方法中，可以通过同聚物加尾反应，修饰所经切割的载体和编码 BCMP 的基因。可以将重组分子通过转化、转染、感染、电穿孔等导入宿主细胞中，使得产生许多拷贝

的所述基因序列。

在具体实施方案中，用加入编码所述 BCMP 的分离基因的重组 DNA 分子、cDNA 或合成 DNA 序列转化宿主细胞，使得能够产生多拷贝的所述基因。因此，通过培养转化子，从转化子中分离所述重组 DNA 分子，必要时从所分离的重组 DNA 中提取所插入的基因，可以大量获得所述基因。

本发明的核苷酸序列包括编码与天然 BCMP 氨基酸序列基本相同的氨基酸序列的核苷酸序列、编码具有功能等同氨基酸的氨基酸序列的核苷酸序列、编码 BCMP、BCMP 的片段、BCMP 相关多肽或 BCMP 相关多肽的片段的核苷酸序列。

在一个具体实施方案中，通过在 BCMP 的核苷酸序列中引入一个或多个核苷酸取代、添加或缺失，使得在所编码的蛋白质中引入一个或多个氨基酸取代、添加或缺失，可以构建编码 BCMP 相关多肽的分离的核酸分子。可以用本领域技术人员已知的标准技术引入突变，所述技术包括例如定点诱变和 PCR 介导的诱变。最好在一种或多种预测非必需的氨基酸残基上进行保守氨基酸取代。“保守氨基酸取代”是其中所述氨基酸残基被具有带相似电荷的侧链的氨基酸残基取代的氨基酸取代。在本领域中，已经确定了具有带相似电荷的侧链的氨基酸的多个类别。这些类别包括具有碱性侧链的氨基酸(例如赖氨酸、精氨酸、组氨酸)、具有酸性侧链的氨基酸(例如天冬氨酸、谷氨酸)、具有不带电荷的极性侧链的氨基酸(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)、具有非极性侧链的氨基酸(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)、具有 $\beta$ 分支侧链的氨基酸(例如苏氨酸、缬氨酸、异亮氨酸)和具有芳族侧链的氨基酸(例如酪氨酸、苯丙氨酸、色氨酸、组氨酸)。或者，可以例如通过饱和诱变，沿整个编码序列或部分编码序列随机引入突变，并且可以根据生物活性筛选所产生的突变体，以鉴定保留活性的突变体。在诱变之后，可以表达所编码的蛋白质，并且可以测

定所述蛋白质的活性。

### 编码 BCMP 的 DNA 的表达

可以将编码 BCMP、BCMP 类似物、BCMP 相关肽或者任何上述物质的片段或其它衍生物的核苷酸序列，插入到合适的表达载体中，所述载体例如含有所插入的蛋白编码序列的转录和翻译的必需元件的载体。所述必需转录信号和翻译信号也可以由编码所述 BCMP 的天然基因或其侧翼区来供应，或者由编码所述 BCMP 相关多肽的天然基因或其侧翼区来供应。在本发明中可以利用多种宿主-载体系统来表达所述蛋白编码序列。这些包括但不限于用病毒(例如痘苗病毒、腺病毒等)感染的哺乳动物细胞系统；用病毒(例如杆状病毒)感染的昆虫细胞系统；微生物，例如含有酵母载体的酵母；或用噬菌体、DNA 或质粒 DNA 或粘粒 DNA 转化的细菌。载体的表达元件在其强度和特异性方面不同。根据所用的宿主-载体系统，可以使用多种合适转录元件和翻译元件中的任何一种。在具体实施方案中，表达编码人类基因的核苷酸序列(或编码人类 BCMP 的功能活性部分的核苷酸序列)。在另一实施方案中，表达包含所述 BCMP 结构域的 BCMP 片段。

先前描述用于将 DNA 片段插入载体中的任何方法，都可以用来构建含有由合适转录和翻译控制信号和所述蛋白编码序列组成的嵌合基因的表达载体。这些方法可以包括体外重组 DNA 技术和合成技术以及体内重组体(基因重组)。编码 BCMP 或其片段的核酸序列的表达，可以由第二种核酸序列来调节，使得所述 BCMP 或片段在用所述重组 DNA 分子转化的宿主中表达。例如，BCMP 的表达，可以由本领域已知的任何启动子或增强子元件来控制。可以用来控制编码 BCMP 或 BCMP 相关多肽的基因表达的启动子包括但不限于：SV40 早期启动子区(Bernoist 和 Chambon, 1981, Nature 290:304-310)、在劳斯肉瘤病毒的 3' 长末端重复中含有的启动子(Yamamoto 等, 1980, Cell 22:787-797)、疱疹胸苷激酶启动子(Wagner 等, 1981, Proc. Nat. Acad. Sci. U.S.A.

78:1441-1445)、金属硫蛋白基因的调节序列(Brinster 等, 1982, Nature 296:39-42)、四环素(Tet)启动子(Gossen 等, 1995, Proc. Natl. Acad. Sci. USA 89:5547-5551); 原核表达载体, 例如 $\beta$ -内酰胺酶启动子(Villa-Kamaroff 等, 1978, Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731)或 *tac* 启动子(DeBoer 等, 1983, Proc. Natl. Acad. Sci. U.S.A. 80:21-25; 也参见“得自重组细菌的有用蛋白质”, Scientific American, 1980, 242:74-94); 含有胭脂碱合成酶启动子区的植物表达载体(Herrera-Estrella 等, Nature 303:209-213)或花椰菜花叶病毒 35S RNA 启动子(Gardner 等, 1981, Nucl. Acids Res. 9:2871)、以及光合酶核酮糖二磷酸羧化酶的启动子(Herrera-Estrella 等, 1984, Nature 310: 115-120); 得自酵母和其它真菌的启动子元件, 例如 Gal 4 启动子、ADC (醇脱氢酶)启动子、PGK (磷酸甘油激酶)启动子、碱性磷酸酶启动子和以下动物转录控制区, 所述控制区表现出组织特异性, 并且已经用于转基因动物中: 弹性蛋白酶 I 基因控制区, 它在胰腺腺泡细胞中有活性(Swift 等, 1984, Cell 38:639-646; Ornitz 等, 1986, Cold Spring Harbor Symp. Quant. Biol. 50:399-409; MacDonald, 1987, Hepatology 7:425-515); 胰岛素基因控制区, 它在胰腺 $\beta$ 细胞中有活性(Hanahan, 1985, Nature 315:115-122)、免疫球蛋白基因控制区, 它在淋巴样细胞中有活性(Grosschedl 等, 1984, Cell 38:647-658; Adames 等, 1985, Nature 318:533-538; Alexander 等, 1987, Mol. Cell. Biol. 7:1436-1444)、小鼠乳腺癌病毒控制区, 它在睾丸细胞、乳腺细胞、淋巴样细胞和肥大细胞中有活性(Leder 等, 1986, Cell 45:485-495)、清蛋白基因控制区, 它在肝中有活性(Pinkert 等, 1987, Genes and Devel. 1:268-276)、甲胎蛋白基因控制区, 它在肝中有活性(Krumlauf 等, 1985, Mol. Cell. Biol. 5:1639-1648; Hammer 等, 1987, Science 235:53-58)、 $\alpha$ 1-抗胰蛋白酶基因控制区, 它在肝中有活性(Kelsey 等, 1987, Genes and Devel. 1:161-171)、 $\beta$ -珠蛋白基因控制区, 它在骨髓性细胞中有活性(Mogram 等, 1985, Nature 315:338-340; Kollias 等, 1986, Cell 46:89-94); 髓鞘碱性蛋白基因控制区, 它在脑中少突胶质细胞中有活性(Readhead

等, 1987, *Cell* 48:703-712); 肌球蛋白轻链-2 基因控制区, 它在骨骼肌中有活性(Sani, 1985, *Nature* 314:283-286); 神经元特异性烯醇化酶(NSE), 它在神经元细胞中有活性(Morelli 等, 1999, *Gen. Virol.* 80:571-83); 脑衍生神经营养因子(BDNF)基因控制区, 它在神经元细胞中有活性(Tabuchi 等, 1998, *Biochem. Biophysic. Res. Com.* 253:818-823); 胶质细胞原纤维酸性蛋白(GFAP)启动子, 它在星形胶质细胞中有活性(Gomes 等, 1999, *Braz J Med Biol Res* 32(5):619-631; Morelli 等, 1999, *Gen. Virol.* 80:571-83)和促性腺激素释放激素基因控制区, 它在下丘脑中有活性(Mason 等, 1986, *Science* 234:1372-1378)。

在一个具体实施方案中, 使用一种载体, 所述载体包含一个与 BCMP 编码核酸有效连接的启动子、一个或多个复制起点和任选的一个或多个选择标记(例如抗生素抗性基因)。

在一个具体实施方案中, 通过将 BCMP 或 BCMP 相关多肽编码序列亚克隆到三种 pGEX 载体(谷胱甘肽 S-转移酶表达载体; Smith 和 Johnson, 1988, *Gene* 7:31-40)中任一种的 *EcoRI* 限制位点中, 制备一种表达构建体。这保证地来自以正确读框亚克隆的 BCMP 产物或 BCMP 相关多肽的表达。

在哺乳动物宿主细胞中, 可以利用许多基于病毒的表达系统。在使用腺病毒作为表达载体的情况下, 可以将所述 BCMP 编码序列或 BCMP 相关多肽编码序列与腺病毒转录/翻译控制复合物(例如晚期启动子和三联前导序列)连接。然后将这种嵌合基因通过体外或体内重组, 插入到腺病毒基因组中。在病毒基因组的非必需区(例如 E1 区和 E3 区)中插入, 将产生有活力的并且能够在被感染细胞中表达所述抗体分子的重组病毒。(参见例如 Logan 和 Shenk, 1984, *Proc. Natl. Acad. Sci. USA* 81:355-359)。所插入抗体编码序列的有效翻译, 也可能需要特定的起始信号。这些信号包括 ATG 起始密码子和相邻序列。此外, 所述起始密码子必须符合所需编码序列的读框, 以确保翻译完整的插入片段。这些外源的翻译控制信号和起始密码子可以有多种来源, 可

以是天然的，也可以是合成的。通过包括合适的转录增强子元件、转录终止子等，可以提高表达效率。(参见 Bittner 等, 1987, *Methods in Enzymol.* 153:51-544)。

含有 BCMP 或 BCMP 相关多肽编码基因的插入片段的表达载体可以通过三个通用方法来鉴定：(a)核酸杂交，(b)“标记”基因功能是否存在，和(c)所插入序列的表达。在第一个方法中，用包含与所插入的 BCMP 编码基因同源的序列的探针，通过核酸杂交，可以检测表达载体中插入的 BCMP 编码基因的存在。在第二个方法中，根据由于在所述载体中插入 BCMP 编码基因引起的某些“标记”基因功能(例如胸苷激酶活性、抗生素抗性、转化表型、在杆状病毒中形成包含体等)存在与否，可以鉴定和选择所述重组载体/宿主系统。例如，如果所述 BCMP 编码基因插入到该载体的标记基因序列中，则根据所述标记基因功能的缺乏，可以鉴定出含有所述 BCMP 编码基因插入片段的重组体。在第三个方法中，通过分析由所述重组体表达的所述基因产物(即 BCMP)，可以鉴定重组表达载体。这类方法可以基于例如体外测定系统中所述 BCMP 的物理特性或功能特性，例如与抗 BCMP 抗体的结合。

另外，可以选择调节所插入序列表达、或者以所需的特定方式修饰和加工所述基因产物的宿主细胞株。来自某些启动子的表达在存在某些诱导物的情况下可以增加；因而可以控制遗传工程 BCMP 或 BCMP 相关多肽的表达。此外，不同的宿主细胞具有翻译和翻译后加工和修饰(例如蛋白质的糖基化、磷酸化)的特征性和特定机制。可以选择合适的细胞系或宿主系统，以确保所表达的外源蛋白的所需修饰和加工。例如，在细菌系统中表达将产生非糖基化产物，在酵母中表达将产生糖基化产物。可以使用具有正确加工初级转录物、基因产物的糖基化和磷酸化的细胞机器的真核宿主细胞。这类哺乳动物宿主细胞包括但不限于 CHO、VERY、BHK、Hela、COS、MDCK、293、3T3 和 WI38。此外，不同的载体/宿主表达系统可能在不同程度上影响加

工反应。

对于重组蛋白长期高产量的生产,优选稳定的表达。例如,可以对稳定表达差异表达蛋白或途径基因蛋白的细胞系进行工程改造。不使用含有病毒复制起点的表达载体,而是用受合适表达控制元件(例如启动子、增强子序列、转录终止子、聚腺苷酸化位点等)控制的DNA和选择标记转化宿主细胞。在导入外源DNA后,让工程细胞在富集培养基中生长1-2天,然后转换到选择培养基上。重组质粒中的选择标记赋予对所述选择的抗性,使得细胞能够将所述质粒稳定地整合到其染色体中,生长形成转化灶,可以进而将转化灶克隆,扩充为细胞系。该方法可以有利地用来对表达差异表达的或途径基因蛋白的细胞系进行工程改造。这类工程细胞系可能在影响所述差异表达的或途径基因蛋白的内源活性的化合物的筛选和评价方面特别有用。

可以使用许多选择系统,包括但不限于在tk-细胞、hgp<sup>r</sup>t-细胞或ap<sup>r</sup>t-细胞中可以分别使用单纯疱疹病毒胸苷激酶(Wigler等,1977,Cell 11:223)、次黄嘌呤-鸟嘌呤磷酸核糖转移酶(Szybalska和Szybalski,1962,Proc. Natl. Acad. Sci. USA 48:2026)和腺嘌呤磷酸核糖转移酶(Lowy等,1980,Cell 22:817)基因。此外,抗代谢物抗性可以用作对以下基因的选择基础:dhfr它赋予对氨基蝶呤的抗性(Wigler等,1980,Natl. Acad. Sci. USA 77:3567; O'Hare等,1981,Proc. Natl. Acad. Sci. USA 78:1527); gpt,它赋予对霉酚酸的抗性(Mulligan和Berg,1981,Proc. Natl. Acad. Sci. USA 78:2072); neo,它赋予对氨基糖苷G-418的抗性(Colberre-Garapin等,1981,J. Mol. Biol. 150:1); hypro,它赋予对潮霉素的抗性(Santerre等,1984,Gene 30:147)。

在其它具体实施方案中,所述BCMP、片段、类似物或衍生物可以作为融合蛋白产物或嵌合蛋白产物(包含通过肽键与一种异源蛋白序列连接的所述蛋白、片段、类似物或衍生物)表达。例如,本发明的多肽可以与免疫球蛋白(IgA、IgE、IgG、IgM)的恒定区或其部分(CH1、CH2、CH3,或其任何组合或其部分)融合,产生嵌合多肽。这类融合

蛋白可以有助于纯化、增加体内半寿期和增强抗原穿越表皮屏障传递至免疫系统。已经表明由人 CD4-多肽的前两个结构域和哺乳动物免疫球蛋白重链或轻链恒定区的不同结构域组成的嵌合蛋白的体内半寿期增加并且有利于纯化。参见例如 EP 394,827; Traunecker 等, *Nature*, 331:84-86 (1988)。已经证明, 对与 FcRn 结合配偶体(例如 IgG 或 Fc 片段)缀合的抗原(例如胰岛素)而言, 其抗原穿越表皮屏障至免疫系统的传递增强(参见例如 PCT 公开说明书 WO 96/22042 和 WO 99/04813)。

编码 BCMP、BCMP 的片段、BCMP 相关多肽、BCMP 相关多肽片段的核酸, 可以与附加表位(例如血凝素(“HA”)标志或 flag 标志)融合, 以有助于所表达多肽的检测和纯化。例如, Janknecht 等描述的系统, 提供在人细胞系中表达的非变性融合蛋白的快速纯化(Janknecht 等, 1991, *Proc. Natl. Acad. Sci. USA* 88:8972-897)。

融合蛋白可以如下制备: 用本领域已知的方法, 将编码所需氨基酸序列的合适核酸序列在正确的编码框内相互连接, 并且通过本领域周知的方法表达所述嵌合产物。或者, 可以通过蛋白合成技术, 例如利用肽合成仪, 制备融合蛋白。

cDNA 序列和基因组序列都可以克隆和表达。

### BCMP 的结构域结构

某些 BCMP 的结构域是本领域已知的, 并且描述于科学文献中。此外, 可以采用本领域技术人员已知的技术鉴定 BCMP 的结构域。例如, 通过采用一种或多种以下程序可以鉴定 BCMP 的一个或多个结构域: ProDom、TMpred 和 SAPS。ProDom 将多肽的氨基酸序列与编译的结构域的数据库进行比较(参见例如 <http://www.toulouse.inra.fr/prodom.html>; Corpet F., Gouzy J.和 Kahn D., 1999, *Nucleic Acids Res.*, 27:263-267)。TMprep 预测多肽跨膜区及其定向。该程序使用基于 TMbase 的统计学分析的算法, TMbase 是天然存在的跨膜蛋白的数据库(参见例如 [http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html);

Hofmann 和 Stoffel. (1993) “TMbase - 跨膜蛋白区段的数据库”。Biol. Chem. Hoppe-Seyler 347, 166)。SAPS 程序分析多肽的统计学显著性特征, 如电荷-簇(charge-cluster)、重复序列、疏水区、复合结构域(参见例如 Brendel 等, 1992, Proc. Natl. Acad. Sci. USA 89:2002-2006)。因此, 基于本发明的描述, 技术人员可以鉴定具有酶活性或结合活性的 BCMP 结构域, 还可以鉴定编码这类结构域的核苷酸序列。这些核苷酸序列然后可以用于保留所述 BCMP 酶活性或结合活性的 BCMP 片段的重组表达。

基于本发明的描述, 技术人员可以鉴定具有酶活性或结合活性的 BCMP 结构域, 还可以鉴定编码这类结构域的核苷酸序列。这些核苷酸序列然后可以用于保留所述 BCMP 酶活性或结合活性的 BCMP 片段的重组表达。

在一个实施方案中, BCMP 具有与已知多肽的已鉴定结构域足够相似的氨基酸序列。本文所用的术语“足够相似”是指第一种氨基酸序列或核苷酸序列含有足够数目的与第二种氨基酸序列或核苷酸序列相同或等同(例如具有相似侧链)的氨基酸残基或核苷酸, 使得所述第一种和第二种氨基酸序列或核苷酸序列具有或者编码共同的结构域或共同的功能活性或共同的结构域和共同的功能活性。

可以采用本领域技术人员众所周知的技术, 评价 BCMP 结构域的功能。例如, 可以采用技术人员已知的技术, 评价结构域的激酶活性或者其与 DNA 结合的能力。例如通过测量多肽将底物磷酸化的能力, 可以评价激酶活性。例如通过在电泳迁移率变动测定中测量多肽与 DNA 结合元件结合的能力, 可以评价 DNA 结合活性。

### 抗 BCMP 抗体的产生

按照本发明, BCMP、BCMP 类似物、BCMP 相关蛋白或任何上述物质的片段或衍生物, 可以用作免疫原来产生免疫特异性结合这种免疫原的抗体。这类免疫原可以通过任何常规方法包括上述方法来分

离。本发明的抗体包括但不限于多克隆抗体、单克隆抗体、双特异性抗体、人源化抗体或嵌合抗体、单链抗体、Fab 片段和 F(ab')片段、通过 Fab 表达文库产生的片段、抗独特型(抗 Id)抗体和任何上述抗体或片段的表位结合片段。本文所用的术语“抗体”是指免疫球蛋白分子和免疫球蛋白分子的免疫活性部分,即含有特异性结合抗原的抗原结合部位的分子。本发明的免疫球蛋白分子可以是免疫球蛋白分子的任何类别(例如 IgG、IgE、IgM、IgD 和 IgA)或亚类。

在一个实施方案中,识别 BCMP 编码基因的基因产物的抗体是公众可得到的。在另一实施方案中,用本领域技术人员已知的方法产生识别 BCMP、BCMP 类似物、BCMP 相关多肽或任何上述物质的片段或衍生物的抗体。

在本发明的一个实施方案中,产生针对 BCMP 的一个特定结构域的抗体。在一个具体实施方案中,BCMP 的亲水性片段用作供抗体生产用的免疫原。

在抗体的生产中,所需抗体的筛选可以通过本领域已知的技术例如 ELISA (酶联免疫吸附测定)来完成。例如,为了选择识别 BCMP 特定结构域的抗体,人们可以分析所产生的杂交瘤中与含这种结构域的 BCMP 片段结合的产物。关于选择特异性结合第一种 BCMP 同系物、但不特异性地结合(或结合急剧减少)第二种 BCMP 同系物的抗体,人们可以根据与所述第一种 BCMP 同系物的结合阳性且缺乏与所述第二种 BCMP 同系物结合(或结合减少)来进行选择。同样,对于选择特异性结合一种 BCMP、但不特异性地结合(或结合急剧减少)同一蛋白的不同同种型(例如具有与 BCMP 相同核心肽的不同糖形)的抗体,人们可以根据与所述 BCMP 的结合阳性且缺乏与所述不同同种型(例如不同糖形)结合(或结合减少)来进行选择。因而,本发明提供与 BCMP 结合的亲和性高于(优选亲和性高至少 2 倍、更优选至少 5 倍、再更优选至少 10 倍)与所述 BCMP 的一种或多种不同同种型(例如糖形)结合的亲和性的抗体(最好是单克隆抗体)。

可用于本发明方法的多克隆抗体是来源于免疫动物血清的不均一的抗体分子群体。也可以使用未分级分离的免疫血清。本领域已知的各种方法可以用来产生针对 BCMP、BCMP 的片段、BCMP 相关多肽或 BCMP 相关多肽片段的多克隆抗体。在一个具体实施方案中,可以获得针对 BCMP 或 BCMP 相关多肽的表位的兔多克隆抗体。例如,为了生产多克隆抗体或单克隆抗体,可以通过注射天然形式或合成(例如重组)形式的 BCMP、BCMP 的片段、BCMP 相关多肽或 BCMP 相关多肽片段,免疫各种宿主动物,包括但不限于兔子、小鼠、大鼠等。本文所述的 the Preferred Technology 提供了适用于这类免疫的分离的 BCMP。如果所述 BCMP 通过凝胶电泳纯化,则可以在将所述 BCMP 用于免疫之前,从聚丙烯酰胺凝胶中进行或不进行提取。根据宿主物种,可以用各种佐剂来增强免疫应答,所述佐剂包括但不限于弗氏完全或不完全佐剂、矿物凝胶例如氢氧化铝、表面活性剂例如溶血卵磷脂、多聚醇、聚阴离子、肽、油乳液、匙孔鳃血蓝蛋白、二硝基酚和佐剂例如 BCG (卡介苗)或小棒杆菌(*Corynebacterium parvum*)。其它佐剂也是本领域众所周知的。

对于针对 BCMP、BCMP 的片段、BCMP 相关多肽或 BCMP 相关多肽片段的单克隆抗体(mAb)的制备,可以使用提供用培养的连续细胞系生产抗体分子的任何技术。例如,最早由 Kohler 和 Milstein (1975, Nature 256:495-497)开发的杂交瘤技术以及 trioma 技术-人 B 细胞杂交瘤技术(Kozbor 等, 1983, Immunology Today 4:72)和产生人单克隆抗体的 EBV-杂交瘤技术(Cole 等, 1985, Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., 第 77-96 页)。这类抗体可以是任何免疫球蛋白类别(包括 IgG、IgM、IgE、IgA、IgD)和其任何亚类。产生本发明 mAb 的杂交瘤可以体外或体内培养。在本发明的另一实施方案中,可以利用已知技术,在无菌动物中生产单克隆抗体(PCT/US90/02545, 该文献通过引用结合到本文中)。

所述单克隆抗体包括但不限于人单克隆抗体和嵌合单克隆抗体

(例如人-小鼠嵌合体)。嵌合抗体是其中不同部分来源于不同动物种的分子，例如具有人免疫球蛋白恒定区和来源于鼠 mAb 的可变区的分子。(参见例如 Cabilly 等，美国专利第 4,816,567 号；和 Boss 等，美国专利第 4,816,397 号，所述文献通过引用全部结合到本文中)。人源化抗体是来自非人类物种的抗体分子，这种分子具有一个或多个得自所述非人类物种的互补性决定区(CDR)和一个得自人免疫球蛋白分子的构架区。(参见例如 Queen，美国专利第 5,585,089 号，该专利通过引用全部结合到本文中)。

嵌合单克隆抗体和人源化单克隆抗体可以通过本领域已知的重组 DNA 技术生产，例如采用在以下文献中描述的方法：PCT 公开说明书 WO 87/02671；欧洲专利申请 184,187；欧洲专利申请 171,496；欧洲专利申请 173,494；PCT 公开说明书 WO 86/01533；美国专利第 4,816,567 号；欧洲专利申请 125,023；Better 等，1988, *Science* 240:1041-1043；Liu 等，1987, *Proc. Natl. Acad. Sci. USA* 84:3439-3443；Liu 等，1987, *J. Immunol.* 139:3521-3526；Sun 等，1987, *Proc. Natl. Acad. Sci. USA* 84:214-218；Nishimura 等，1987, *Canc. Res.* 47:999-1005；Wood 等，1985, *Nature* 314:446-449；和 Shaw 等，1988, *J. Natl. Cancer Inst.* 80:1553-1559；Morrison，1985, *Science* 229:1202-1207；Oi 等，1986, *Bio/Techniques* 4:214；美国专利第 5,225,539 号；Jones 等，1986, *Nature* 321:552-525；Verhoeyan 等(1988) *Science* 239:1534；和 Beidler 等，1988, *J. Immunol.* 141:4053-4060。

完全人抗体，对于人类受治疗者的治疗性治疗而言是尤其理想的。这类抗体可以采用转基因小鼠来生产，所述转基因小鼠不能表达内源免疫球蛋白重链和轻链基因，但它可以表达人类重链和轻链基因。所述转基因小鼠用所选抗原(例如完整的本发明 BCMP 或本发明 BCMP 的一部分)以正常方式免疫。可以采用常规杂交瘤技术获得针对所述抗原的单克隆抗体。所述转基因小鼠带有的免疫球蛋白基因在 B 细胞分化期间重排，随后经历类别转换和体细胞突变。因此，

采用这种技术,有可能生产治疗上有用的 IgG、IgA、IgM 和 IgE 抗体。有关生产人抗体的这种技术的综述,参见 Lonberg 和 Huszar (1995, *Int. Rev. Immunol.* 13:65-93)。有关生产人抗体和人单克隆抗体的这种技术和生产这类抗体的方案的详细讨论,参见例如美国专利 5,625,126; 美国专利 5,633,425; 美国专利 5,569,825; 美国专利 5,661,016; 和美国专利 5,545,806。另外,可以雇用例如 Abgenix, Inc. (Freemont, CA)和 Genpharm (San Jose, CA)的公司,采用与上述技术相似的技术提供针对所选抗原的人抗体。

可以采用称为“指导选择(guided selection)”的技术,产生识别所选表位的完全人抗体。在这种方法中,用所选非人类单克隆抗体(例如小鼠抗体)来指导识别同一表位的完全人抗体的选择。(Jespers 等(1994) *Bio/technology* 12:899-903)。

本发明的抗体也可以采用本领域已知的各种噬菌体展示方法来产生。在噬菌体展示方法中,功能抗体结构域展示在携带编码所述抗体的多核苷酸序列的噬菌体颗粒表面。特别是,这类噬菌体可以用来展示由所有组成部分或联合抗体文库(例如人或鼠)表达的抗原结合结构域。可以用抗原,例如采用标记抗原或结合或捕获在固体表面或微珠上的抗原,选择或鉴定表达结合目的抗原的抗原结合结构域的噬菌体。用于这些方法中的噬菌体通常是丝状噬菌体,包括 fd 和 M13,将由噬菌体表达的结合结构域与 Fab、Fv 或二硫键稳定的 Fv 抗体结构域与或者噬菌体基因 III 或者基因 VIII 蛋白重组融合。可以用来制备本发明抗体的噬菌体展示方法包括公开于下述文献中的方法: Brinkman 等, *J. Immunol. Methods* 182:41-50 (1995); Ames 等, *J. Immunol. Methods* 184: 177-186 (1995); Kettleborough 等, *Eur. J. Immunol.* 24:952-958 (1994); Persic 等, *Gene* 187:9-18 (1997); Burton 等, *Advances in Immunology* 57:191-280 (1994); PCT 申请 PCT/GB91/01134; PCT 公开说明书 WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; 和美国专利第 5,698,426、

5,223,409、5,403,484、5,580,717、5,427,908; 5,750,753、5,821,047、5,571,698、5,427,908、5,516,637、5,780,225、5,658,727、5,733,743 和 5,969,108 号, 所述文献中的每个都通过引用全部结合到本文中。

正如上述参考文献中所述的, 在噬菌体选择后, 如下详述, 可以分离得自噬菌体的所述抗体编码区, 将其用来产生完整抗体包括人抗体或任何其它所需的抗原结合片段, 然后在任何所需的宿主(包括哺乳动物细胞、昆虫细胞、植物细胞、酵母和细菌等)中表达。例如, 也可以采用重组产生 Fab、Fab'和 F(ab')<sub>2</sub> 片段的技术, 采用例如在以下文献中公开的本领域已知的方法: PCT 公开说明书 WO 92/22324; Mullinax 等, *BioTechniques* 12(6):864-869 (1992); 和 Sawai 等, *AJRI* 34:26-34 (1995) 和 Better 等, *Science* 240:1041-1043 (1988) (所述参考文献通过引用全部结合到本文中)。

可以用来产生单链 Fv 和抗体的技术的实例包括以下文献中公开的那些技术: 美国专利 4,946,778 和 5,258,498; Huston 等, *Methods in Enzymology* 203:46-88 (1991); Shu 等, *PNAS* 90:7995-7999 (1993); 和 Skerra 等, *Science* 240:1038-1040 (1988)。

本发明还提供双特异性抗体的应用, 所述双特异性抗体可以用本领域已知的方法制备。全长双特异性抗体传统上的生产, 基于两个免疫球蛋白重链-轻链对的共表达, 其中所述两条链具有不同的特异性 (Milstein 等, 1983, *Nature* 305:537-539)。因为免疫球蛋白重链和轻链随机分配, 所以这些杂交瘤(quadromas)产生 10 种不同抗体分子的潜在混合物, 其中仅一种具有正确的双特异性结构。正确分子的纯化(通常通过亲和色谱步骤来进行)相当麻烦, 并且产物收率低。在于 1993 年 5 月 13 日公开的 WO 93/08829 和 Traunecker 等, 1991, *EMBO J.* 10:3655-3659 中公开了相似的方法。

按照一种不同的更优选的方法, 将具有所需结合特异性(抗体-抗原结合部位)抗体可变区与免疫球蛋白恒定区序列融合。所述融合体最好具有免疫球蛋白重链恒定区, 包含至少部分铰链区、CH2 区和 CH3

区。最好在至少一种所述融合体中存在含有轻链结合所必需部位的第一重链恒定区(CH1)。将编码所述免疫球蛋白重链融合体和(必要时)免疫球蛋白轻链的DNA插入到单独的表达载体中,共转染到合适的宿主生物中。当在构建中所用的三种多肽链不等的比率提供最佳收率时,这在调节实施方案中所述三种多肽片段的相互比率方面提供很大的灵活性。然而,当至少两种多肽链等比率的表达导致高收率时,或者当所述比率没有特别的重要性时,有可能将编码两种或所有三种多肽链的编码序列插入到一种表达载体中。

在这种方法的一个优选实施方案中,所述双特异性抗体由在一个臂中一条具有第一结合特异性的杂种免疫球蛋白重链和在另一个臂中一条杂种免疫球蛋白重链-轻链对(提供第二结合特异性)组成。发现这种不对称结构有利于从不想要的免疫球蛋白链组合中分离出所需双特异性化合物,因为仅在所述双特异性分子中的一半中存在免疫球蛋白轻链,提供了一种灵活的分离途径。这种方法公开于1994年3月3日公开的WO 94/04690中。有关产生双特异性抗体的其它细节,参见例如Suresh等, *Methods in Enzymology*, 1986, 121:210。

本发明提供所述抗BCMP免疫球蛋白分子的功能活性片段、衍生物或类似物。功能活性是指所述片段、衍生物或类似物能够引发识别所述片段、衍生物或类似物所来源的抗体所识别的相同抗原的抗独特型抗体(即第三抗体)。具体地说,在一个优选实施方案中,通过缺失构架序列和位于特异性识别所述抗原的CDR序列C端的CDR序列,可以增强所述免疫球蛋白分子的独特型的抗原性。为了确定哪些CDR序列结合所述抗原,可以用本领域已知的任何结合测定方法,将含所述CDR序列的合成肽用于与所述抗原的结合测定中。

本发明提供抗体片段,例如但不限于 $F(ab')_2$ 片段和Fab片段。识别特定表位的抗体片段可以用已知技术来产生。 $F(ab')_2$ 片段由可变区、轻链恒定区和重链的CH1区组成,通过胃蛋白酶消化所述抗体分子产生 $F(ab')_2$ 片段。通过还原 $F(ab')_2$ 片段的二硫键产生Fab片段。本

发明也提供本发明抗体的重链和轻链的二聚体、或其任何最小片段，例如 Fv 或单链抗体(SCA) (参见，如以下文献中所述：美国专利 4,946,778; Bird, 1988, Science 242:423-42; Huston 等, 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; 和 Ward 等, 1989, Nature 334:544-54), 或具有与本发明抗体相同特异性的任何其它分子。通过将 Fv 区的重链和轻链片段通过一个氨基酸桥连接产生单链多肽，形成单链抗体。可以使用在大肠杆菌(*E. coli*)中装配功能 Fv 片段的技术(Skerra 等, 1988, Science 242:1038-1041)。

在其它实施方案中，本发明提供本发明免疫球蛋白的融合蛋白(或其功能活性片段)，例如其中所述免疫球蛋白通过一个共价键(例如肽键)融合于另一种非免疫球蛋白的蛋白(或其部分，最好是所述蛋白的至少 10 个、20 个或 50 个氨基酸的部分)的氨基酸序列的或者 N 末端或者 C 末端。所述免疫球蛋白或其片段最好在恒定区的 N 末端共价连接于所述其它蛋白。如上所述，这类融合蛋白可以有利于纯化，增加体内半寿期并且增强抗原穿越表皮屏障传递至免疫系统。

本发明的免疫球蛋白包括或者经修饰(即通过共价连接任何类型的分子，只要这种共价连接不损害免疫特异性结合)的类似物和衍生物。例如但不作为限制，所述免疫球蛋白的衍生物和类似物包括例如通过糖基化、乙酰化、PEG 化、磷酸化、酰胺化、用已知保护基/封闭基团衍生化、蛋白酶剪切、与细胞配体或其它蛋白连接等进行进一步修饰的衍生物和类似物。许多化学修饰中的任一种都可以用已知技术来进行，所述已知技术包括但不限于特异性化学切割、乙酰化、甲酰化等。另外，所述类似物或衍生物可以含有一个或多个非典型氨基酸。

上述抗体可以用于本领域已知的涉及本发明 BCMP 定位和活性的方法中，例如使这些蛋白成像、测量其在合适生理样品中的水平，用于诊断方法中等。

## 抗体的表达

本发明的抗体可以通过本领域用于合成抗体已知的任何方法尤其是通过化学合成或重组表达来生产，最好通过重组表达技术来生产。

抗体、或其片段、衍生物或类似物的重组表达需要构建编码所述抗体的核酸。如果所述抗体的核苷酸序列是已知的，则可以由化学合成的寡核苷酸(例如如 Kutmeier 等, 1994, *BioTechniques* 17:242 中所述)来装配编码所述抗体的核酸，简而言之，这包括合成含有抗体编码序列的部分的重叠寡核苷酸，使那些寡核苷酸退火和连接，然后通过 PCR 扩增所连接的寡核苷酸。

或者，可以通过克隆所述抗体，获得编码所述抗体的核酸。如果含有编码所述特定抗体的核酸的克隆不能获得，但已知所述抗体分子的序列，则采用可与所述序列的 3'端和 5'端杂交的合成引物，从合适的来源(例如抗体 cDNA 文库，或由表达所述抗体的任何组织或细胞中产生 cDNA 文库)通过 PCR 扩增，或者用对所述特定基因序列特异性的寡核苷酸探针通过克隆，可以获得编码所述抗体的核酸。

如果特异性识别特定抗原的抗体分子(或者用于克隆编码这种抗体的核酸的 cDNA 文库的来源)不能获得，则可以通过本领域已知的任何方法产生对特定抗原特异性的抗体，例如通过免疫动物(例如兔子)产生多克隆抗体，或者更优选通过产生单克隆抗体。或者，通过在 Fab 表达文库(例如如 Huse 等, 1989, *Science* 246:1275-1281 中所述)中筛选结合所述特定抗原的 Fab 片段的克隆，或者通过筛选抗体文库(参见例如 Clackson 等, 1991, *Nature* 352:624; Hane 等, 1997, *Proc. Natl. Acad. Sci. USA* 94:4937)，可以获得编码所述抗体的至少 Fab 部分的克隆。

一旦获得编码所述抗体分子的至少可变区的核酸后，则可以将其引入含有编码所述抗体分子恒定区的核苷酸序列的载体中(参见例如 PCT 公开说明书 WO 86/05807; PCT 公开说明书 WO 89/01036; 和美国专利第 5,122,464 号)。也可得到含有所述完整轻链或重链以供与所述

核酸共表达、以允许表达完整抗体分子的载体。然后，编码所述抗体的核酸可以用来引入用不含巯基的任何氨基酸残基取代(或缺失)参与链内二硫键的一个或多个可变区半胱氨酸残基所必需的核苷酸取代或缺失。这类修饰可以通过用于在核苷酸序列中引入特定突变或缺失的本领域已知的任何方法来进行，所述方法例如但不限于化学诱变、体外定点诱变(Hutchinson 等, 1978, *J. Biol. Chem.* 253:6551)、基于 PCR 的方法等。

另外，可以使用开发用于产生“嵌合抗体”的技术(Morrison 等, 1984, *Proc. Natl. Acad. Sci.* 81:851-855; Neuberger 等, 1984, *Nature* 312:604-608; Takeda 等, 1985, *Nature* 314:452-454)，所述技术包括将得自小鼠的具有合适抗原特异性的抗体分子的基因与得自人类的具有合适生物活性的抗体分子的基因剪接在一起。正如上文所述，嵌合抗体是一种其中不同部分来源于不同动物种的分子，例如具有来源于鼠 mAb 的可变区和人抗体恒定区的抗体，例如人源化抗体。

一旦获得编码本发明抗体分子的核酸后，则可以采用本领域众所周知的技术，通过重组 DNA 技术产生可供用于生产所述抗体分子的载体。因此，本文描述了用于通过表达含有所述抗体分子序列的核酸制备本发明蛋白的方法。本领域技术人员众所周知的方法可以用来构建含有抗体分子编码序列和合适转录和翻译控制信号的表达载体。这些方法包括例如体外重组 DNA 技术、合成技术和体内遗传重组。参见例如在 Sambrook 等(1990, *Molecular Cloning, A Laboratory Manual*, 第 2 版, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY)和 Ausubel 等(编著, 1998, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY)中描述的技术。

通过常规技术将所述表达载体转染到宿主细胞中，然后通过常规技术培养所述转染细胞，以生产本发明的抗体。

用来表达本发明重组抗体的宿主细胞可以或者是细菌细胞例如大肠杆菌，或者最好是真核细胞，尤其是对于表达完整的重组抗体分

子而言。特别是，哺乳动物细胞例如中国仓鼠卵巢细胞(CHO)与载体例如来自人巨细胞病毒的主要立即早期基因启动子元件是一个抗体的有效表达系统(Foecking 等, 1986, *Gene* 45:101; Cockett 等, 1990, *Bio/Technology* 8:2)。

可以使用多种宿主-表达载体系统来表达本发明的抗体分子。这类宿主-表达系统代表可以用来生产并随后纯化目的编码序列的载体,但也代表当用所述合适核苷酸编码序列转化或转染后可以原位表达本发明抗体分子的细胞。这些包括但不限于微生物,例如用含有抗体编码序列的重组噬菌体 DNA、质粒 DNA 或粘粒 DNA 表达载体转化的细菌(例如大肠杆菌、枯草杆菌(*B. subtilis*));用含有抗体编码序列的重组酵母表达载体转化的酵母(例如酵母属(*Saccharomyces*)、毕赤酵母属(*Pichia*));用含有所述抗体编码序列的重组病毒表达载体(例如杆状病毒)感染的昆虫细胞系统;用含有抗体编码序列的重组病毒表达载体(例如花椰菜花叶病毒 CaMV;烟草花叶病毒 TMV)感染或用含有抗体编码序列的重组质粒表达载体(例如 Ti 质粒)转化的植物细胞系统;或带有含来源于哺乳动物细胞基因组的启动子(例如金属硫蛋白启动子)或得自哺乳动物病毒基因组的启动子(例如腺病毒晚期启动子;痘苗病毒 7.5K 启动子)的重组表达构建体的哺乳动物细胞系统(例如 COS、CHO、BHK、293、3T3 细胞)。

在细菌系统中,根据所表达抗体分子的计划用途,可以有利地对多种表达载体进行选择。例如,当需生产大量的这种蛋白以供生产包含抗体分子的药用组合物时,可能理想的是指导高水平的易纯化的融合蛋白产物表达的载体。这类载体包括但不限于:大肠杆菌表达载体 pUR278 (Ruther 等, 1983, *EMBO J.* 2:1791),其中可以将所述抗体编码序列单独连接到载体中,并且与 *lac Z* 编码区符合读框,使得产生融合蛋白;pIN 载体(Inouye 和 Inouye, 1985, *Nucleic Acids Res.* 13:3101-3109; Van Heeke 和 Schuster, 1989, *J. Biol. Chem.* 24:5503-5509)等。pGEX 载体也可以用来表达作为具有谷胱甘肽 S-转移酶(GST)的融合蛋白的外

源多肽。一般而言，这类融合蛋白是可溶性的，可以通过吸附和与基质谷胱甘肽-琼脂糖微珠结合，然后通过游离谷胱甘肽存在下洗脱，容易地从裂解细胞中纯化。设计 pGEX 载体，使其包括凝血酶或因子 Xa 蛋白酶切割位点，使得可以从 GST 部分释放出所克隆的靶基因产物。

在昆虫系统中，苜蓿银纹夜蛾(*Autographa californica*)核型多角体病毒(AcNPV)用作表达外源基因的载体。所述病毒在草地贪夜蛾(*Spodoptera frugiperda*)细胞中生长。所述抗体编码序列可以单独克隆到该病毒的非必需区(例如多角体蛋白基因)中，并置于 AcNPV 启动子(例如多角体蛋白启动子)的控制之下。在哺乳动物宿主细胞中，可以利用多种基于病毒的表达系统(例如腺病毒表达系统)。

如上所述，可以选择调节所插入序列表达或者以所需特定方式修饰和加工所述基因产物的宿主细胞株。蛋白产物的这类修饰(例如糖基化)和加工(例如切割)可能对于所述蛋白的功能是重要的。

对于重组抗体长期高产量的生产，优选稳定的表达。例如，可以如下产生稳定表达目的抗体的细胞系：用含有目的抗体的核苷酸序列和选择标记(例如新霉素或潮霉素)的核苷酸序列的表达载体转染所述细胞，并根据所述选择标记的表达进行选择。这类工程细胞系可能在与所述抗体分子直接或间接相互作用的化合物的筛选和评价方面特别有用。

所述抗体分子的表达水平可以通过载体扩增来提高(有关综述，参见 Bebbington 和 Hentschel, *The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning*, 第 3 卷(Academic Press, New York, 1987))。当表达抗体的载体系统中的标记可扩增时，宿主细胞培养物中存在的抑制剂水平的增加，将增加所述标记基因的拷贝数。由于所述扩增的区与所述抗体基因相连，所以所述抗体的产生也将增加(Crouse 等, 1983, *Mol. Cell. Biol.* 3:257)。

可以用两种本发明表达载体共转染所述宿主细胞，第一种载体编

码重链来源的多肽，而第二种载体编码轻链来源的多肽。所述两种载体可以含有相同的选择标记，所述选择标记能够使重链多肽和轻链多肽的表达相等。或者，可以使用单一载体，所述载体既编码重链多肽，也编码轻链多肽。在这种情况下，轻链应该置于重链之前，以避免有毒性的游离重链过量(Proudfoot, 1986, Nature 322:52; Kohler, 1980, Proc. Natl. Acad. Sci. USA 77:2197)。重链和轻链的编码序列可以包括 cDNA 或基因组 DNA。

一旦重组表达了本发明的抗体分子，则可以通过用于纯化抗体分子的本领域已知的任何方法来纯化，所述方法例如通过色谱(例如离子交换色谱、亲和色谱如具有 A 蛋白或特定抗原的亲和色谱以及大小柱色谱)、离心、差别溶解性、或通过用于蛋白质纯化的任何其它标准技术来纯化。

或者，通过利用对所表达的融合蛋白特异性的抗体，可以容易地纯化任何融合蛋白。例如，Janknecht 等所述的系统，可供用于在人细胞系中表达的非变性融合蛋白的快速纯化(Janknecht 等, 1991, Proc. Natl. Acad. Sci. USA 88:8972-897)。在这种系统中，将目的基因亚克隆到痘苗病毒重组质粒(vaccinia recombination plasmid)中，使得所述基因的可读框在转录上与由 6 个组氨酸残基组成的氨基末端标志融合。所述标志用作所述融合蛋白的基质结合结构域。将得自用重组痘苗病毒感染的细胞的提取物，上样到  $\text{Ni}^{2+}$  氮川乙酸-琼脂糖柱上，用含咪唑的缓冲液选择性地洗脱含组氨酸标志的蛋白。

### 缀合抗体

在一个优选实施方案中，抗 BCMP 抗体或其片段与一个诊断或治疗部分缀合。所述抗体可以用于诊断或给定治疗方案疗效的测定。通过将所述抗体与可检测物质偶联，可以有利于检测。可检测物质的实例包括各种酶、辅基、荧光物质、发光物质、生物发光物质、放射性核素、发射正电子的金属(用于正电子发射断层扫描)和非放射性顺磁

性金属离子。有关可以与抗体缀合作按照本发明的诊断剂的金属离子，一般参见美国专利第 4,741,900 号。合适的酶包括辣根过氧化物酶、碱性磷酸酶、 $\beta$ -半乳糖苷酶或乙酰胆碱酯酶；合适的辅基包括链霉抗生物素蛋白、抗生物素蛋白和生物素；合适的荧光物质包括伞形酮、荧光素、异硫氰酸荧光素、罗丹明、二氯三嗪胺荧光素、丹磺酰氯和藻红蛋白；合适的发光物质包括鲁米诺；合适的生物发光物质包括荧光素酶、荧光素和水母发光蛋白；合适的放射性核素包括  $^{125}\text{I}$ 、 $^{131}\text{I}$ 、 $^{111}\text{In}$  和  $^{99}\text{Tc}$ 。

抗 BCMP 抗体或其片段可以与治疗药或药物部分缀合，以改变给定生物应答。所述治疗药或药物部分不应解释为限于典型的化学治疗药。例如，所述药物部分可以是一种具有所需生物活性的蛋白质或多肽。这类蛋白可以包括例如毒素，例如相思豆毒蛋白、蓖麻毒蛋白 A、假单胞菌外毒素或白喉毒素；蛋白质，例如肿瘤坏死因子、 $\alpha$ -干扰素、 $\beta$ -干扰素、神经生长因子、血小板衍生生长因子、组织纤溶酶原激活物、血栓形成剂(thrombotic agent)或抗血管生成剂例如制管张素或 endostatin；或生物反应调节物，例如淋巴因子、白介素-1 (IL-1)、白介素-2 (IL-2)、白介素-6 (IL-6)、粒细胞-巨噬细胞集落刺激因子(GM-CSF)、粒细胞集落刺激因子(G-CSF)、神经生长因子(NGF)或其它生长因子。

用于将这类治疗部分与抗体缀合的技术是众所周知的，参见例如 Arnon 等，“用于癌症治疗中药物的免疫靶向的单克隆抗体”，*Monoclonal Antibodies And Cancer Therapy*, Reisfeld 等(编著)，第 243-56 页(Alan R. Liss, Inc. 1985)；Hellstrom 等，“用于给药的抗体”*Controlled Drug Delivery* (第 2 版), Robinson 等(编著)，第 623-53 页(Marcel Dekker, Inc. 1987)；Thorpe，“癌症治疗中细胞毒性剂的抗体载体：综述”，*Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera 等(编著)，第 475-506 页(1985)；“放射性标记抗体在癌症治疗中的治疗应用的分析、结果和未来前景”，*Monoclonal Antibodies For Cancer*

Detection And Therapy, Baldwin 等(编著), 第 303-16 页(Academic Press 1985), 和 Thorpe 等, “抗体-毒素缀合物的制备和细胞毒性特性”, Immunol. Rev., 62:119-58 (1982)。

或者, 可以将抗体与第二抗体缀合, 形成抗体异源缀合物, 如 Segal 在美国专利第 4,676,980 号中所述。

具有或不具有与其缀合的治疗部分的抗体, 可以用作单独给予或与细胞毒性因子和/或细胞因子联合给予的治疗剂。

### 乳腺癌的诊断

按照本发明, 得自怀疑患有或已知患有乳腺癌的受治疗者的乳腺组织、血清、血浆或尿试验样品, 可以用于诊断或监测。在一个实施方案中, 与对照样品(得自一个或多个无乳腺癌的受治疗者)或预先确定的参比范围相比, 试验样品中一种或多种 BCMP 丰度的改变指示存在乳腺癌; 正如以上详述的, 在表 1 和表 2 中鉴定了适用于此目的的 BCMP。在另一实施方案中, 与对照样品或预先确定的参比范围相比, 试验样品中一种或多种 BCMP 的相对丰度指示存在一种亚型的乳腺癌(例如家族性或散发性乳腺癌)。在再一实施方案中, 与对照样品或预先确定的参比范围相比, 试验样品中一种或多种 BCMP 的相对丰度指示乳腺癌的严重程度(例如转移的可能性)。在任一上述方法中, 本文表 1 和表 2 中所述的一种或多种 BCMP 的检测可以任选地与以上表 3 中鉴定的一种或多种 BCMP 和/或其它乳腺癌生物标记的检测相结合。本领域中的任何合适方法都可以用来测定 BCMP 的水平, 包括但不限于本文所述的 the Preferred Technology、激酶测定、检测和/或显现所述 BCMP 的免疫测定(例如蛋白质印迹法、免疫沉淀后接十二烷基硫酸钠聚丙烯酰胺凝胶电泳、免疫细胞化学等)。在 BCMP 具有已知功能时, 可以使用针对该功能的测定来测量 BCMP 的表达。在再一实施方案中, 与对照样品或预先确定的参比范围相比, 试验样品中表 1 和表 2 所鉴定的编码一种或多种 BCMP 的 mRNA 的丰度的改变指示存在乳腺

癌。任何合适的杂交测定都可以用来通过检测和/或显现编码所述 BCMP 的 mRNA 来测定 BCMP 的表达(例如 RNA 印迹分析、斑点印迹法、原位杂交等)。任选的是,也可以检测以上表 3 中鉴定的一种或多种 BCMP。

在本发明的另一实施方案中,与 BCMP 特异性结合的标记抗体、其衍生物和类似物可以用于诊断目的,以检测、诊断或监测乳腺癌。最好是,在动物中、更优选在哺乳动物中、最优选在人类中检测乳腺癌。

### 筛选测定

本发明提供与 BCMP 结合或对 BCMP 的表达或活性具有刺激或抑制效应的因子(例如候选化合物或试验化合物)的鉴定方法。本发明也提供与 BCMP 相关多肽或 BCMP 融合蛋白结合或对 BCMP 相关多肽或 BCMP 融合蛋白的表达或活性具有刺激或抑制效应的因子、候选化合物或试验化合物的鉴定方法。因子、候选化合物或试验化合物的实例包括但不限于核酸(例如 DNA 和 RNA)、糖类、脂质、蛋白质、肽、肽模拟物、小分子和其它药物。可以采用本领域已知的组合文库法方法的多种方法中的任一种获得因子,所述文库法包括:生物学文库;空间可寻址平行固相或溶液相文库;需要去卷积(deconvolution)的合成文库法;“一微珠一化合物”文库法;和利用亲和色谱选择的合成文库法。所述生物学文库法限于肽文库,而其它四种方法适用于肽、非肽寡聚物或化合物的小分子文库(Lam, 1997, Anticancer Drug Des. 12:145; 美国专利第 5,738,996 号和美国专利第 5,807,683 号,每个所述文献通过引用全部结合到本文中)。

分子文库合成方法的实例可以在本领域中找到,例如 DeWitt 等, 1993, Proc. Natl. Acad. Sci. USA 90:6909; Erb 等, 1994, Proc. Natl. Acad. Sci. USA 91:11422; Zuckermann 等, 1994, J. Med. Chem. 37:2678; Cho 等, 1993, Science 261:1303; Carrell 等, 1994, Angew. Chem. Int. Ed. Engl.

33:2059; Carell 等, 1994, *Angew. Chem. Int. Ed. Engl.* 33:2061; 和 Gallop 等, 1994, *J. Med. Chem.* 37:1233, 每个文献通过引用全部结合到本文中。

化合物文库可以存在, 例如存在于溶液中(例如 Houghten, 1992, *Bio/Techniques* 13:412-421)或存在于微珠上(Lam, 1991, *Nature* 354:82-84)、芯片(Fodor, 1993, *Nature* 364:555-556)、细菌(美国专利第 5,223,409 号)、孢子(spore)(专利第 5,571,698 号; 第 5,403,484 号; 和第 5,223,409 号)、质粒(Cull 等, 1992, *Proc. Natl. Acad. Sci. USA* 89:1865-1869)或噬菌体(Scott 和 Smith, 1990, *Science* 249:386-390; Devlin, 1990, *Science* 249:404-406; Cwirla 等, 1990, *Proc. Natl. Acad. Sci. USA* 87:6378-6382; 和 Felici, 1991, *J. Mol. Biol.* 222:301-310), 每个文献通过引用全部结合到本文中。

在一个实施方案中, 在一个基于细胞的测定系统中, 鉴定与 BCMP、BCMP 的片段(例如功能活性片段)、BCMP 相关多肽、BCMP 相关多肽片段或 BCMP 融合蛋白相互作用(即与其结合)的因子。按照该实施方案, 使表达 BCMP、BCMP 的片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白的细胞与候选化合物或对照化合物接触, 测定所述候选化合物与所述 BCMP 相互作用的能力。如果需要, 这种测定可以用来筛选多种候选化合物(例如文库)。所述细胞例如可以来源于原核生物(例如大肠杆菌)或者来源于真核生物(例如酵母或哺乳动物)。此外, 所述细胞可以是内源表达所述 BCMP、所述 BCMP 的片段、BCMP 相关多肽、所述 BCMP 相关多肽的片段或 BCMP 融合蛋白, 或者可以对所述细胞进行遗传工程改造, 以表达所述 BCMP、所述 BCMP 的片段、BCMP 相关多肽、所述 BCMP 相关多肽的片段或 BCMP 融合蛋白。在某些情况下, 可以用放射性标记(例如  $^{32}\text{P}$ 、 $^{35}\text{S}$  和  $^{125}\text{I}$ )或荧光标记(例如异硫氰酸荧光素、罗丹明、藻红蛋白、藻蓝蛋白、别藻蓝蛋白、邻苯二醛或荧光胺), 标记所述 BCMP、所述 BCMP 的片段、BCMP 相关多肽、所述 BCMP 相关多肽的片段或 BCMP 融合蛋白

或所述候选化合物，使得可以检测 BCMP 和候选化合物之间的相互作用。所述候选化合物与 BCMP、BCMP 的片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白直接或间接相互作用的能力，可以通过本领域技术人员已知的方法来测定。例如，候选化合物和 BCMP、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白之间的相互作用，可以通过流式细胞术、闪烁测定、免疫沉淀或蛋白质印迹分析来测定。

在另一实施方案中，在一个无细胞测定系统中，鉴定与 BCMP、BCMP 的片段(例如功能活性片段)、BCMP 相关多肽、BCMP 相关多肽片段或 BCMP 融合蛋白相互作用(即与其结合)的因子。按照该实施方案，使天然或重组 BCMP 或其片段、或天然或重组 BCMP 相关多肽或其片段、或 BCMP 融合蛋白或其片段与候选化合物或对照化合物接触，测定所述候选化合物与所述 BCMP 或 BCMP 相关多肽或 BCMP 融合蛋白相互作用的能力。如果需要，这种测定可以用来筛选多种候选化合物(例如文库)。最好是首先例如通过使所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白与特异性识别其并与其结合的固定化抗体接触，或者通过使所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白的纯化制剂与设计用以结合蛋白质的表面接触，将所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白固定化。可以将所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白部分或完全纯化(例如部分或完全不含其它多肽)，或者它们是细胞裂解液的部分。此外，所述 BCMP、BCMP 片段、BCMP 相关多肽或 BCMP 相关多肽的片段可以是包含所述 BCMP 或者其生物活性部分或 BCMP 相关多肽和一个结构域例如谷胱甘肽-S-转移酶的融合蛋白。或者，可以用本领域技术人员众所周知的技术(例如生物素化试剂盒, Pierce Chemicals; Rockford, IL)，将所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP

融合蛋白生物素化。所述候选化合物与 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白相互作用的能力，可以通过本领域技术人员已知的方法测定。

在另一实施方案中，用一种基于细胞的测定系统来鉴定与负责 BCMP 的产生或降解或者负责 BCMP 的翻译后修饰的蛋白(例如酶)或其生物活性部分结合或者调节其活性的因子。在初次筛选中，使多种化合物(例如文库)与天然表达或重组表达以下 BCMP 的细胞接触：(i) BCMP、BCMP 同种型、BCMP 同系物、BCMP 相关多肽、BCMP 融合蛋白或任何上述 BCMP 的生物活性片段；和(ii) 负责加工所述 BCMP、BCMP 同种型、BCMP 同系物、BCMP 相关多肽、BCMP 融合蛋白或它们的片段的蛋白质，以鉴定调节所述 BCMP、BCMP 同种型、BCMP 同系物、BCMP 相关多肽、BCMP 融合蛋白或它们的片段的产生、降解或翻译后修饰的化合物。如果需要，可以在第二次筛选中，针对天然表达或重组表达所述特定目的 BCMP 的细胞，分析在所述初次筛选中鉴定的化合物。所述候选化合物调节 BCMP、同种型、同系物、BCMP 相关多肽或 BCMP 融合蛋白的产生、降解或翻译后修饰的能力，可以通过本领域技术人员已知的方法测定，所述方法包括但不限于流式细胞术、闪烁测定、免疫沉淀和蛋白质印迹法。

在另一实施方案中，在一种竞争性结合测定中，鉴定与 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽片段或 BCMP 融合蛋白竞争性相互作用(即与其结合)的因子。按照该实施方案，使表达 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽的片段或 BCMP 融合蛋白的细胞与候选化合物或已知与所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽片段或 BCMP 融合蛋白相互作用的化合物接触；然后测定所述候选化合物与所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽片段或 BCMP 融合蛋白优先相互作用的能力。或者，在一个无细胞测定系统中，通过使 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽片段或 BCMP 融合蛋白与候

选化合物和已知与所述 BCMP、BCMP 相关多肽或 BCMP 融合蛋白相互作用的化合物接触，鉴定优先与所述 BCMP、BCMP 片段、BCMP 相关多肽、BCMP 相关多肽片段或 BCMP 融合蛋白相互作用(即与其结合)的因子。如上所述，所述候选化合物与 BCMP、BCMP 片段、BCMP 相关多肽或 BCMP 相关多肽片段相互作用的能力，可以通过本领域技术人员已知的方法测定。这些测定无论是基于细胞的测定，还是无细胞测定，都可以用来筛选多种候选化合物(例如一种文库)。

在另一实施方案中，通过使表达 BCMP 或 BCMP 相关多肽的细胞(例如原核来源或真核来源的细胞)与候选化合物或对照化合物(即磷酸缓冲盐溶液(PBS))接触，并且测定所述 BCMP、BCMP 相关多肽或 BCMP 融合蛋白、编码所述 BCMP 的 mRNA 或编码所述 BCMP 相关多肽的 mRNA 的表达，鉴定调节(即正调节或负调节)所述 BCMP 或 BCMP 相关多肽表达或活性的因子。将在存在所述候选化合物情况下选定的 BCMP、BCMP 相关多肽、编码所述 BCMP 的 mRNA 或编码所述 BCMP 相关多肽的 mRNA 的表达水平与在缺乏所述候选化合物的情况下(例如在存在对照化合物时)所述 BCMP、BCMP 相关多肽、编码所述 BCMP 的 mRNA 或编码所述 BCMP 相关多肽的 mRNA 的表达水平进行比较。则根据这种比较，所述候选化合物被鉴定为所述 BCMP 或所述 BCMP 相关多肽表达的调节剂。例如，所述 BCMP 或 mRNA 在存在所述候选化合物时的表达显著高于无所述候选化合物时的表达，则所述候选化合物被鉴定为所述 BCMP 或 mRNA 表达的刺激物。或者，当所述 BCMP 或 mRNA 在存在所述候选化合物时的表达显著低于无所述候选化合物时的表达，则所述候选化合物被鉴定为所述 BCMP 或 mRNA 表达的抑制剂。BCMP 或编码其的 mRNA 的表达水平，可以通过本领域技术人员已知的方法测定。例如，mRNA 的表达可以通过 RNA 印迹分析或 RT-PCR 来评价，而蛋白水平可以通过蛋白质印迹分析来评价。

在另一实施方案中，通过使含有 BCMP 或 BCMP 相关多肽的制

剂或表达所述 BCMP 或 BCMP 相关多肽的细胞(例如原核细胞或真核细胞)与试验化合物或对照化合物接触,并且测定所述试验化合物调节(例如刺激或抑制)所述 BCMP 或 BCMP 相关多肽的活性的能力,鉴定调节所述 BCMP 或 BCMP 相关多肽活性的因子。通过测定所述 BCMP 或 BCMP 相关多肽的细胞信号转导途径(例如胞内  $\text{Ca}^{2+}$ 、二酰基甘油、IP3 等)的诱导、测定所述靶对合适底物的催化活性或酶活性、测定报道基因(例如对 BCMP 或 BCMP 相关多肽应答并且与编码可检测标记例如荧光素酶的核酸有效连接的调节元件)的诱导、或者测定细胞应答例如细胞分化或细胞增殖,可以评价所述 BCMP 或 BCMP 相关多肽的活性。根据本发明的描述,可以用本领域技术人员已知的技术来测量这些活性(参见例如美国专利第 5,401,639 号,其通过引用结合到本文中)。然后,通过将所述候选化合物的效应与对照化合物的效应进行比较,将这些候选化合物鉴定为 BCMP 或 BCMP 相关多肽活性的调节剂。合适的对照化合物包括磷酸缓冲盐溶液(PBS)和生理盐水(NS)。

在另一实施方案中,在动物模型中鉴定调节(即正调节或负调节)BCMP 或 BCMP 相关多肽的表达、活性或者表达和活性两者的因子。合适动物的实例包括但不限于小鼠、大鼠、兔子、猴、豚鼠、犬和猫。所用的动物最好代表一种乳腺癌模型(例如人乳腺癌细胞系例如 MDA-MB-345 在缺乏雌激素的严重联合免疫缺陷(SCID)小鼠中的异种移植物, Eccles 等, 1994 Cell Biophysics 24/25, 279)。这些可以用来测试调节 BCMP 水平的化合物,因为在这些模型中表现出的病状与乳腺癌的病状相似。按照该实施方案,将所述试验化合物或对照化合物给予(例如口服、直肠或胃肠外给予,例如腹膜内或静脉内给予)合适的动物,并测定对所述 BCMP 或 BCMP 相关多肽的表达、活性或者表达和活性两者的影响。可以通过以上概述的方法,评价 BCMP 或 BCMP 相关多肽表达的改变。

在又一实施方案中,在双杂种测定或三杂种测定中,BCMP 或 BCMP 相关多肽用作“饵蛋白”,以鉴定与 BCMP 或 BCMP 相关多肽

结合或相互作用的其它蛋白(参见例如美国专利第 5,283,317 号; Zervos 等(1993) Cell 72:223-232; Madura 等(1993) J. Biol. Chem. 268:12046-12054; Bartel 等(1993) Bio/Techniques 14:920-924; Iwabuchi 等(1993) Oncogene 8:1693-1696; 和 PCT 公开说明书 WO 94/10300)。正如本领域技术人员会认识到的, 这类结合蛋白也可能作为例如涉及本发明 BCMP 的信号途径的上游或下游元件参与本发明 BCMP 的信号传播。

本发明还提供通过上述筛选测定鉴定的新型因子及其在本文所述治疗方法的用途。另外, 本发明也提供与一种或多种本发明 BCMP 相互作用或调节其活性的因子在乳腺癌治疗药物生产方面的用途。

### BCMP 的治疗用途

本发明可供用于通过给予治疗化合物治疗或预防各种疾病和障碍。这类化合物包括但不限于: BCMP、BCMP 类似物、BCMP 相关多肽和它们的衍生物(包括片段); 针对上述 BCMP、BCMP 类似物、BCMP 相关多肽和它们的衍生物的抗体; 编码 BCMP、BCMP 类似物、BCMP 相关多肽和它们的片段的核酸; BCMP 或 BCMP 相关多肽编码基因的反义核酸; 和 BCMP 或 BCMP 相关多肽编码基因的调节剂(例如激动剂和拮抗剂)。本发明的一个重要特征是鉴定涉及乳腺癌的 BCMP 编码基因。通过给予一种提升患乳腺癌受治疗者的乳腺组织中被降低的一种或多种 BCMP 的功能或表达的治疗化合物, 或者通过给予降低患乳腺癌受治疗者的乳腺组织中被增加的一种或多种 BCMP 的功能或表达的治疗化合物, 可以治疗(例如缓解症状或延迟发作或发展)或预防乳腺癌。

在一个实施方案中, 单独给予或者与一种或多种其它治疗化合物或疗法联合给予分别与一种 BCMP 特异性结合的一种或多种抗体。这类治疗化合物或疗法的实例包括但不限于紫杉酚、环磷酰胺、他莫昔芬和多柔比星。

最好是, 生物产物例如抗体对于接受给药的受治疗者而言是同种

异体的(allogeneic)。在一个优选实施方案中,将人 BCMP 或人 BCMP 相关多肽、编码人 BCMP 或人 BCMP 相关多肽的核苷酸序列、或针对人 BCMP 或人 BCMP 相关多肽的抗体给予人类受治疗者以用于治疗(例如缓解症状或延迟发作或发展)或预防。

### 乳腺癌的治疗和预防

通过给予怀疑或已知患有乳腺癌或者有发生乳腺癌危险的受治疗者一种调节(即提高或降低)在患乳腺癌受治疗者的乳腺组织中相对于无乳腺癌受治疗者乳腺组织而言差别存在的一种或多种 BCMP 水平或活性(即功能)的化合物,治疗或预防乳腺癌。在一个实施方案中,通过给予怀疑或已知患有乳腺癌或者有发生乳腺癌危险的受治疗者一种正调节(即提高)在患乳腺癌受治疗者的乳腺组织中被降低的一种或多种 BCMP 水平或活性(即功能)的化合物,治疗或预防乳腺癌。在另一实施方案中,给予负调节在患乳腺癌受治疗者的乳腺组织中被增加的一种或多种 BCMP 水平或活性(即功能)的化合物。这种化合物的实例包括但不限于:BCMP、BCMP 片段和 BCMP 相关多肽;编码 BCMP、BCMP 片段和 BCMP 相关多肽的核酸(例如用于基因治疗);具有酶活性的那些 BCMP 或 BCMP 相关多肽、已知调节该酶活性的化合物或分子。可以用体外测定,鉴定可以使用的其它化合物例如 BCMP 激动剂。

通过给予怀疑或已知患有乳腺癌或者有发生乳腺癌危险的受治疗者一种负调节在患乳腺癌受治疗者的乳腺组织中被增加的一种或多种 BCMP 水平或活性的化合物,也可治疗或预防乳腺癌。在另一实施方案中,给予正调节在患乳腺癌受治疗者的乳腺组织中被降低的一种或多种 BCMP 水平或活性的化合物。这种化合物的实例包括但不限于:BCMP 反义寡核苷酸、核酶、针对 BCMP 的抗体以及抑制 BCMP 酶活性的化合物。可以用体外测定鉴定其它有用的化合物,例如 BCMP 拮抗剂和小分子 BCMP 拮抗剂。

在一个优选实施方案中,使治疗或预防适应各个受治疗者的需

要。因此，在具体实施方案中，将提升一种或多种 BCMP 水平或功能的化合物，治疗性或预防性地给予怀疑或已知患有乳腺癌的、体内所述一种或多种 BCMP 的水平或功能缺乏或相对于对照或正常参比范围降低的受治疗者。在其它实施方案中，将提升一种或多种 BCMP 水平或功能的化合物，治疗性或预防性地给予怀疑或已知患有乳腺癌的、体内所述一种或多种 BCMP 的水平或功能相对于对照或正常参比范围增加的受治疗者。在其它实施方案中，将降低一种或多种 BCMP 水平或功能的化合物，治疗性或预防性地给予怀疑或已知患有乳腺癌的、体内所述一种或多种 BCMP 的水平或功能相对于对照或正常参比范围增加的受治疗者。在其它实施方案中，将降低一种或多种 BCMP 水平或功能的化合物，治疗性或预防性地给予怀疑或已知患有乳腺癌的、体内所述一种或多种 BCMP 的水平或功能相对于对照或正常参比范围降低的受治疗者。例如通过获得样品(例如乳腺组织、血液或尿或者组织样品例如活检组织样品)并且体外分析所述 BCMP 的水平或活性、或者编码所述 BCMP 的 mRNA 的水平、或者上述的任何组合，可以容易地检测出由于给予这类化合物所致的 BCMP 功能或水平的改变。这类测定可以在给予本文所述化合物之前和之后进行。

本发明的化合物包括但不限于将所述 BCMP 分布型朝正常分布型恢复的任何化合物，例如小有机分子、蛋白质、肽、抗体、核酸等，条件是这类化合物或治疗不包括紫杉酚、环磷酰胺、他莫昔芬和多柔比星。本发明的化合物可以与任何其它化合物联合给予，所述其它化合物包括紫杉酚、环磷酰胺、他莫昔芬和多柔比星。

### 疫苗治疗

BCMP 可用作抗原性物质，可以用于生产供乳腺癌治疗或预防用的疫苗。这类物质可以是“抗原性”和/或“免疫原性”物质。一般而言，“抗原性”是指蛋白质能够用来在受治疗者体内产生抗体或实际上能够诱发抗体应答。“免疫原性”是指所述蛋白质能够在受治疗者

体内引发保护性免疫应答。因此，在后一种情况下，所述蛋白质可能不仅能够产生抗体应答，而且能够产生基于非抗体的免疫应答。

技术人员会认识到，所述 BCMP 的同系物或衍生物也可用作抗原性/免疫原性物质。因此，本发明包括例如包括一种或多种添加、缺失、取代等的蛋白质。另外，一种氨基酸有可能用相似“类型”的另一种氨基酸取代。例如，一种疏水性氨基酸用另一种疏水性氨基酸取代。人们可以使用程序例如 CLUSTAL 程序来比较氨基酸序列。该程序比较氨基酸序列，通过在任一个序列中适当插入空位发现最佳比对。有可能计算最佳比对的氨基酸同一性或相似性(同一性加上氨基酸类型的保守)。如 BLASTx 的程序将对最长的相似序列段进行比对，并赋予所述拟合(fit)一个分值。因此，在发现几个相似性区时有可能获得比较，每个区具有一个不同的分值。在本发明中考虑了两种类型的分析。

在同系物和衍生物的情况下，与本文所述的蛋白质的同一性程度的重要性低于所述同系物或衍生物应该保留其抗原性和/或免疫原性的重要性。然而，适宜提供与本文所述蛋白或多肽具有至少 60%相似性(如上文所讨论的)的同系物或衍生物。优选提供具有至少 70%相似性、更优选至少 80%相似性的同系物或衍生物。最优选提供具有至少 90%或甚至 95%相似性的同系物或衍生物。

在另一方法中，所述同系物或衍生物可以是融合蛋白，加入使得更易于纯化(例如通过有效标记所需蛋白或多肽)的部分。可能必需除去所述“标记”，或可能是所述融合蛋白本身保留足够的有用的抗原性。

众所周知，有可能筛选出抗原性蛋白或多肽以鉴定表位区，即负责所述蛋白或多肽的抗原性或免疫原性的那些区。本领域技术人员众所周知的方法可以用来测试片段和/或同系物和/或衍生物的抗原性。因此，本发明的片段应该包括一个或多个这类表位区或与这类区足够相似，以保留其抗原性/免疫原性特性。因此，关于按照本发明的片段，同一性的程度也许不重要，因为它们可以与本文所述的一种蛋白或多

肽、同系物或衍生物的一个特定部分 100%相同。再者，关键的问题是所述片段保留其所来源蛋白的抗原性/免疫原性特性。

对同系物、衍生物和片段而言重要的是，它们具有至少一定程度的其所来源的蛋白或多肽的抗原性/免疫原性。因此，在本发明的另一方面，提供 BCMP 的抗原性/免疫原性片段或其同系物或衍生物。

所述 BCMP 或其抗原性片段可以单独提供，或者作为纯化或分离的制剂提供。它们可以作为一种或多种其它本发明蛋白或其抗原性片段的混合物的部分提供。因此，在另一方面，本发明提供一种抗原组合物，所述抗原组合物包含一种或多种本发明 BCMP 和/或一种或多种其抗原性片段。这种组合物可以用于乳腺癌的检测和/或诊断。

第六方面，本发明提供一种乳腺癌的检测和/或诊断方法，所述方法包括：

使一种本发明的抗原性 BCMP 或其抗原性片段或抗原组合物与受检样品接触；并且

检测抗乳腺癌抗体的存在。

具体地说，本发明的蛋白、其抗原性片段或抗原组合物可以用来检测 IgA、IgM 或 IgG 抗体。受检样品适宜为生物样品，例如血液或唾液样品。

另一方面，本发明提供本发明抗原性 BCMP、其抗原性片段或抗原组合物在乳腺癌检测和/或诊断方面的用途。所述检测和/或诊断最好在体外进行。

本发明的抗原性 BCMP、其抗原性片段或抗原组合物可以作为试剂盒提供，以供用于乳腺癌的体外检测和/或诊断。因此，在再一方面，本发明提供用于乳腺癌检测和/或诊断的试剂盒，所述试剂盒包含本发明的抗原性 BCMP、其抗原性片段或抗原组合物。

另外，本发明的抗原性 BCMP、其抗原性片段或抗原组合物可以用来诱发针对乳腺癌的免疫应答。因此，在又一方面，本发明提供本发明的抗原性 BCMP、其抗原性片段或抗原组合物在药物中的应用。

另一方面，本发明提供能够诱发受治疗者体内免疫应答的组合物，所述组合物包含本发明的 BCMP、其抗原性片段或抗原组合物。所述组合物适宜为疫苗组合物，任选包含一种或多种合适的佐剂。这种疫苗组合物可以或者是预防性或者是治疗性疫苗组合物。

本发明的疫苗组合物可以包括一种或多种佐剂。本领域众所周知的实例包括无机凝胶例如氢氧化铝和油包水乳液例如弗氏不完全佐剂。其它有用的佐剂是技术人员众所周知的。

在其它方面，本发明提供：

(a) 本发明 BCMP、其抗原性片段或抗原组合物在制备免疫原性组合物(最好是疫苗)方面的用途；

(b) 这种免疫原性组合物在诱发受治疗者体内免疫应答方面的用途；和

(c) 治疗或预防受治疗者乳腺癌的方法或为受治疗者接种抗乳腺癌的疫苗的方法，所述方法包括下述步骤：给予所述受治疗者有效量的本发明 BCMP、至少一种其抗原性片段或抗原组合物，最好是作为疫苗给予。

在一个具体实施方案中，一种或多种选自表 1 和表 2 的 BCMP 或 BCMP 肽片段的制剂用作治疗乳腺癌的疫苗。这类制剂可以包括佐剂或其它载体。

在另一实施方案中，包含与一种编码选自表 1 和表 2 的一种或多种 BCMP 或 BCMP 肽片段的核苷酸序列互补的 10 个或更多个连续核苷酸的寡核苷酸制剂，用作治疗乳腺癌的疫苗。这种制剂可以包括佐剂或其它载体。

### 基因治疗

在一个具体实施方案中，给予编码 BCMP、BCMP 片段、BCMP 相关多肽或 BCMP 相关多肽片段的核酸，以借助基因治疗增进 BCMP 功能。基因治疗是指给予受治疗者一种已表达或可表达核酸。在这一

实施方案中，所述核酸产生其编码的、通过增进 BCMP 功能介导治疗效应的多肽。

本领域可利用的任何基因治疗方法都可以按照本发明使用。以下描述示例性的方法。

有关基因治疗方法的综述，参见 Goldspiel 等，1993, *Clinical Pharmacy* 12:488-505; Wu 和 Wu, 1991, *Biotherapy* 3:87-95; Tolstoshev, 1993, *Ann. Rev. Pharmacol. Toxicol.* 32:573-596; Mulligan, 1993, *Science* 260:926-932; 和 Morgan 和 Anderson, 1993, *Ann. Rev. Biochem.* 62:191-217; 1993 年 5 月, *TIBTECH* 11(5):155-215。可以使用的重组 DNA 技术领域众所周知的方法描述于 Ausubel 等(编著), 1993, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY; 和 Kriegler, 1990, *Gene Transfer and Expression, A Laboratory Manual*, Stockton Press, NY。

在一个优选方面，所述化合物包含编码 BCMP 或其片段或其嵌合蛋白的核酸，所述核酸是在合适宿主中表达 BCMP 或其片段或其嵌合蛋白的表达载体的部分。具体地说，这种核酸具有与所述 BCMP 编码区有效连接的启动子，所述启动子是诱导型或组成型(任选地为组织特异性)启动子。在另一具体实施方案中，使用一种核酸分子，其中所述 BCMP 编码序列和任何其它所需序列邻接启动在所述基因组中所需位点同源重组的区域，因而保证所述 BCMP 核酸的染色体内表达(Koller 和 Smithies, 1989, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 ;Zijlstra 等, 1989, *Nature* 342:435-438)。

可以直接将所述核酸传递至受治疗者体内，在这种情况下，所述受治疗者直接暴露于所述核酸或携带核酸的载体；这种方法称为体内基因治疗。或者，可以间接地将所述核酸传递至所述受治疗者体内，在这种情况下，首先用所述核酸体外转化细胞，然后将细胞移植到所述受治疗者体内；这种方法称为离体基因治疗。

在一个具体实施方案中，体内直接给予所述核酸，其中所述核酸表达，产生所编码的产物。这可以通过本领域已知的多种方法中任一

种来完成,例如通过将其作为合适核酸表达载体的一部分来构建,然后给予所述核酸使其成为胞内的,例如通过用缺陷型或减毒反转录病毒或其它病毒载体感染(参见美国专利第 4,980,286 号);通过直接注射裸 DNA;通过利用微粒轰击(例如,基因枪; Biolistic, Dupont);通过包上脂质、细胞表面受体或转染剂;通过包囊于脂质体、微粒或微囊中;通过给予与已知进入细胞核的肽连接的所述核酸;或通过给予与经历受体介导胞吞的配体连接的所述核酸(参见例如 Wu 和 Wu, 1987, J. Biol. Chem. 262:4429-4432),所述配体可以用来靶向特异性表达所述受体的靶细胞。在另一实施方案中,可以形成核酸-配体复合物,其中所述配体包含病毒融合肽以破坏内体,使得所述核酸避免被溶酶体降解。在再一实施方案中,通过靶向特异性受体,可以使所述核酸在体内定向,以利于细胞特异性摄取和表达(参见例如 1992 年 4 月 16 日公开的 PCT 公开说明书 WO 92/06180 (Wu 等); 1992 年 12 月 23 日公开的 WO 92/22635 (Wilson 等); 1992 年 11 月 26 日公开的 WO92/20316 (Findeis 等); 1993 年 7 月 22 日公开的 WO93/14188 (Clarke 等); 1993 年 10 月 14 日公开的 WO 93/20221 (Young))。或者,可以将所述核酸导入胞内,并且通过同源重组掺入到宿主细胞 DNA 中以供表达(Koller 和 Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra 等, 1989, Nature 342:435-438)。

在一个具体实施方案中,使用含有编码 BCMP 的核酸的病毒载体。例如,可以使用反转录病毒载体(参见 Miller 等, 1993, Meth. Enzymol. 217:581-599)。这些反转录病毒载体已经过修饰,缺失了对于病毒基因组包装和整合到宿主细胞 DNA 中非必需的反转录病毒序列。将待用于基因治疗的编码所述 BCMP 的核酸克隆到所述载体中,这有助于将所述基因传递到受治疗者体内。有关反转录病毒载体的更详细的细节可以在 Boesen 等, 1994, Biotherapy 6:291-302 中找到,所述文献描述了应用反转录病毒载体将 *mdr1* 基因传递到造血干细胞中,以便使所述干细胞更抗化学治疗。描述在基因治疗中应用反转录病毒载体的其它参考

文献是: Clowes 等, 1994, *J. Clin. Invest.* 93:644-651; Kiem 等, 1994, *Blood* 83:1467-1473; Salmons 和 Gunzberg, 1993, *Human Gene Therapy* 4:129-141; 和 Grossman 和 Wilson, 1993, *Curr. Opin. in Genetics and Devel.* 3:110-114。

腺病毒是可以用于基因治疗的其它病毒载体。对于将基因传递至呼吸上皮而言, 腺病毒是尤其有吸引力的载体。腺病毒正常感染呼吸道上皮, 在此它们引起轻度疾病。基于腺病毒的传递系统的其它靶是肝、中枢神经系统、内皮细胞和肌肉。腺病毒的优点是能够感染非分裂细胞。Kozarsky 和 Wilson, 1993, *Current Opinoin in Genetics and Development* 3:499-503 对基于腺病毒的基因治疗进行了综述。Bout 等, 1994, *Human Gene Therapy* 5:3-10 证明了应用腺病毒载体将基因转移至猕猴呼吸上皮。在基因治疗中应用腺病毒的其它例子可以在以下文献中找到: Rosenfeld 等, 1991, *Science* 252:431-434; Rosenfeld 等, 1992, *Cell* 68:143-155; Mastrangeli 等, 1993, *J. Clin. Invest.* 91:225-234; PCT 公开说明书 WO94/12649; 和 Wang 等, 1995, *Gene Therapy* 2:775-783。

已经提出将腺伴随病毒(AAV)用于基因治疗(Walsh 等, 1993, *Proc. Soc. Exp. Biol. Med.* 204:289-300; 美国专利第 5,436,146 号)。

基因治疗的另一方法涉及通过例如电穿孔、脂质转染、磷酸钙介导的转染或病毒感染的方法将基因转移到组织培养物的细胞中。通常, 所述转移方法包括将选择标记转移到细胞中。然后将所述细胞置于选择之下, 以分离摄入并表达所转移基因的细胞。然后将那些细胞给予受治疗者。

在该实施方案中, 将所述核酸引入细胞中, 然后体内给予所产生的重组细胞。这类引入可以通过本领域已知的任何方法进行, 包括但不限于转染、电穿孔、微注射、用含所述核酸序列的病毒载体或噬菌体载体感染、细胞融合、染色体介导的基因转移、微细胞介导的基因转移、原生质球融合等。用于将外源基因导入细胞的许多技术是本领域已知的(参见例如 Loeffler 和 Behr, 1993, *Meth. Enzymol.* 217:599-618;

Cohen 等, 1993, Meth. Enzymol. 217:618-644; Cline, 1985, Pharmac. Ther. 29:69-92), 并且可以按照本发明使用, 条件是不破坏受体细胞必需的发育功能和生理功能。所述技术应该保证将所述核酸稳定地转移到细胞中, 使得所述核酸被所述细胞表达, 并且最好是可遗传的且可为其细胞子代表达。

可以将所产生的重组细胞通过本领域已知的各种方法传递给受治疗者。在一个优选实施方案中, 例如皮下注射上皮细胞。在另一实施方案中, 重组皮肤细胞可以作为皮肤移植物应用于所述受治疗者。最好静脉内给予重组血细胞(例如造血干细胞或祖细胞)。设想使用的细胞量取决于所需的效应、受治疗者的病症等, 并且可以由本领域技术人员确定。

可以导入核酸以用于基因治疗的细胞包括任何所需的、可得到的细胞类型, 包括但不限于神经元细胞、神经胶质细胞(例如少突胶质细胞或星形胶质细胞)、上皮细胞、内皮细胞、角质细胞、成纤维细胞、肌细胞、肝细胞; 血细胞, 例如 T 淋巴细胞、B 淋巴细胞、单核细胞、巨噬细胞、嗜中性粒细胞、嗜酸性粒细胞、巨核细胞、粒细胞; 各种干细胞或祖细胞, 特别是造血干细胞或祖细胞, 例如得自骨髓、脐带血、外周血或胎肝的干细胞或祖细胞。

在一个优选实施方案中, 用于基因治疗的细胞是所治疗的受治疗者的自体细胞。

在一个将重组细胞用于基因治疗的实施方案中, 将编码 BCMP 的核酸导入细胞中, 使得所述核酸可被所述细胞或其子代表达, 然后体内给予重组细胞以实施治疗。在一个具体实施方案中, 使用干细胞或祖细胞。可以分离并且体外保持的任何干细胞或祖细胞都可以按照本发明的该实施方案使用(参见例如 1994 年 4 月 28 日公开的 PCT 公开说明书 WO 94/08598; Stemple 和 Anderson, 1992, Cell 71:973-985; Rheinwald, 1980, Meth. Cell Bio. 21A:229; 和 Pittelkow 和 Scott, 1986, Mayo Clinic Proc. 61:771)。

在一个具体实施方案中，待导入以用于基因治疗的核酸包含与所述编码区有效连接的诱导型启动子，使得所述核酸的表达可通过控制转录的合适诱导物的存在与否来控制。

直接注射编码 BCMP 的 DNA 也可以按照例如美国专利第 5,589,466 号中所述的技术来进行。这些技术涉及注射“裸 DNA”，即无脂质体、细胞或除合适载体外的任何其它物质情况下的分离的 DNA 分子。注射编码蛋白质并与合适启动子有效连接的 DNA，导致在注射部位附近的细胞中产生所述蛋白质，并且引发受治疗者体内针对所注射 DNA 编码的蛋白质的免疫应答。在一个优选实施方案中，将包含(a)编码 BCMP 的 DNA 和(b)启动子的裸 DNA 注射到受治疗者体内，以诱发针对所述 BCMP 的免疫应答。

### 抑制 BCMP 以治疗乳腺癌

在本发明的一个实施方案中，通过给予一种拮抗(抑制)在患乳腺癌受治疗者的乳腺组织中与无乳腺癌受治疗者的乳腺组织相比被增加的一种或多种 BCMP 水平和/或功能的化合物，治疗或预防乳腺癌。

可用于此目的的化合物包括但不限于抗 BCMP 抗体(和含有其结合区的片段和衍生物)、BCMP 反义核酸或核酶核酸、以及可以用来通过同源重组“敲除”内源 BCMP 功能的、编码功能异常 BCMP 的核酸(参见例如 Capecchi, 1989, *Science* 244:1288-1292)。抑制 BCMP 功能的其它化合物可以通过利用已知的体外测定来鉴定，所述体外测定例如对试验化合物抑制 BCMP 与另一蛋白或结合配偶体结合的能力、或抑制已知 BCMP 功能的测定。最好在体外或在细胞培养物中分析这类抑制，但也可以使用遗传分析。也可以在给予所述化合物之前和之后，用 the Preferred Technology 来测定所述 BCMP 的水平。最好是，如下文更详细描述，利用合适的体外或体内测定来测定特定化合物的效应，以及确定其给予是否是适用于受影响组织的治疗。

在一个具体实施方案中，将抑制 BCMP 功能的化合物治疗性或预

防性给予与无乳腺癌的受治疗者的乳腺组织或预定参比范围相比检测出体内所述 BCMP 的乳腺组织水平或功能活性增加(例如高于正常水平或所需水平)的受治疗者。可以如上所概述的,用本领域的标准方法,测量 BCMP 水平或功能的增加。优选的 BCMP 抑制剂组合物包括小分子,即 1000 道尔顿或更小的分子。这类小分子可以通过本文所述的筛选方法来鉴定。

### BCMP 的反义调节

在一个具体实施方案中,利用 BCMP 反义核酸抑制 BCMP 的表达。本发明提供包含至少 6 个核苷酸的核酸的治疗性或预防性用途,其中所述核酸为编码 BCMP 或其部分的基因或 cDNA 的反义核苷酸。本文所用的 BCMP “反义”核酸是指借助一定的序列互补性能够与编码 BCMP 的 RNA (最好是 mRNA)的一部分杂交的核酸。所述反义核酸可以与编码 BCMP 的 mRNA 的编码区和/或非编码区互补。这类反义核酸可用作抑制 BCMP 表达的化合物,并且可以用于乳腺癌的治疗或预防。

本发明的反义核酸是双链或单链寡核苷酸、RNA 或 DNA 或其修饰物或衍生物,可以直接给予细胞,或通过外源所导入的序列的转录在胞内产生。

本发明还提供药用组合物,所述药用组合物包含药学上可接受的载体中的有效量的本发明 BCMP 反义核酸,如下文所述。

在另一个实施方案中,本发明提供用于抑制原核细胞或真核细胞中 BCMP 核酸序列表达的方法,所述方法包括为所述细胞提供有效量的包含本发明 BCMP 反义核酸的组合物。

以下详细描述 BCMP 反义核酸及其用途。

### BCMP 反义核酸

所述 BCMP 反义核酸具有至少 6 个核苷酸,最好是范围为 6 至约

50个寡核苷酸的寡核苷酸。在具体方面,所述寡核苷酸至少有10个核苷酸、至少15个核苷酸、至少100个核苷酸或至少200个核苷酸。所述寡核苷酸可以是DNA或RNA或嵌合混合物或衍生物或其修饰形式,可以是单链或双链的。所述寡核苷酸可以在碱基部分、糖部分或磷酸骨架上加以修饰。所述寡核苷酸可以包括其它附加的基团,例如肽;促进穿越细胞膜的转运的因子(参见例如 Letsinger 等,1989, Proc. Natl. Acad. Sci. USA 86:6553-6556; Lemaitre 等,1987, Proc. Natl. Acad. Sci. 84:648-652; 1988年12月15日公开的PCT公开说明书WO 88/09810)或血脑屏障(参见例如1988年4月25日公开的PCT公开说明书WO 89/10134);杂交触发的切割剂(参见例如 Krol 等,1988, BioTechniques 6:958-976)或嵌入剂(参见例如 Zon, 1988, Pharm. Res. 5:539-549)。

在本发明的一个优选方面,提供一种BCMP反义寡核苷酸,最好是单链DNA。所述寡核苷酸可以在其结构的任何位置用本领域一般已知的取代基来修饰。

所述BCMP反义寡核苷酸可以包含至少一种以下经修饰的碱基部分:5-氟尿嘧啶、5-溴尿嘧啶、5-氯尿嘧啶、5-碘尿嘧啶、次黄嘌呤、黄嘌呤、4-乙酰胞嘧啶、5-(羧基羟甲基)尿嘧啶、5-羧甲基氨基甲基-2-硫代尿苷、5-羧甲基氨基甲基尿嘧啶、二氢尿嘧啶、 $\beta$ -D-半乳糖基Q核苷(Queosine)、肌苷、N6-异戊烯基腺嘌呤、1-甲基鸟嘌呤、1-甲基肌苷、2,2-二甲基鸟嘌呤、2-甲基腺嘌呤、2-甲基鸟嘌呤、3-甲基胞嘧啶、5-甲基胞嘧啶、N6-腺嘌呤、7-甲基鸟嘌呤、5-甲基氨基甲基尿嘧啶、5-甲氧基氨基甲基-2-硫尿嘧啶、 $\beta$ -D-甘露糖基Q核苷、5-甲氧羰基甲基尿嘧啶、5-甲氧基尿嘧啶、4-甲硫基-N6-异戊烯基腺嘌呤、尿嘧啶-5-氧乙酸(v)、wybutoxosine、假尿嘧啶、Q核苷、2-硫代胞嘧啶、5-甲基-2-硫尿嘧啶、2-硫尿嘧啶、4-硫尿嘧啶、5-甲基尿嘧啶、尿嘧啶-5-氧乙酸甲酯、尿嘧啶-5-氧乙酸(v)、5-甲基-2-硫尿嘧啶、3-(3-氨基-3-N-2-羧丙基)尿嘧啶、(acp3)w、2,6-二氨基嘌呤和其它碱基类似物。

在另一实施方案中，所述寡核苷酸包含至少一个经修饰的糖部分，例如以下糖部分之一：阿拉伯糖、2-氟阿拉伯糖、木酮糖和己糖。

在再一实施方案中，所述寡核苷酸包含至少一种以下经修饰的磷酸骨架：硫代磷酸酯、二硫代磷酸酯、氨基硫代磷酸酯、氨基磷酸酯、二氨基磷酸酯、甲基磷酸酯、烷基磷酸三酯、formacetal 或 formacetal 类似物。

在又一实施方案中，所述寡核苷酸是 $\alpha$ -端基异构寡核苷酸。 $\alpha$ -端基异构寡核苷酸与互补 RNA 形成特殊的双链杂交体，其中与通常的 $\beta$ -单位相反，所述链相互平行(Gautier 等，1987, *Nucl. Acids Res.* 15:6625-6641)。

可以将所述寡核苷酸与另一种分子(例如肽)、杂交触发交联剂、转运因子或杂交触发切割因子缀合。

本发明的寡核苷酸可以通过本领域已知的标准方法合成，例如利用自动 DNA 合成仪(例如市场上可得自 Biosearch, Applied Biosystems 等的自动 DNA 合成仪)来合成。例如，可以用 Stein 等的方法(1988, *Nucl. Acids Res.* 16:3209)合成硫代磷酸寡核苷酸、可以利用控制的多孔玻璃聚合物支持体制备甲基磷酸酯寡核苷酸(Sarin 等，1988, *Proc. Natl. Acad. Sci. USA* 85:7448-7451)。

在一个具体实施方案中，通过从外源序列转录，在胞内产生本发明的 BCMP 反义核酸。例如，可以将载体导入体内，使得所述载体被细胞摄入，所述载体或其部分在所述细胞内转录，产生本发明的反义核酸(RNA)。这种载体可能含有编码所述 BCMP 反义核酸的序列。这种载体可以保持为附加型，或者整合到染色体上，只要它可以转录产生所需的反义 RNA。这类载体可以通过本领域标准的重组 DNA 技术来构建。载体可以是质粒、病毒或本领域已知的用于在哺乳动物细胞内复制和表达的其它载体。可以通过本领域已知的在哺乳动物细胞(优选人类细胞)中起作用的任何启动子，表达编码所述 BCMP 反义 RNA 的序列。这类启动子可以是诱导型或组成型的。以上概述了这类启动

子的实例。

本发明的反义核酸包含与 BCMP 编码基因(最好是编码 BCMP 的人类基因)的 RNA 转录物的至少一部分互补的序列。然而,虽然优选绝对的互补性,但并不需要绝对的互补性。本文所指的“与 RNA 的至少一部分互补的”序列是指具有足够的互补性能够在严格条件下(例如高严格条件包括在 7%十二烷基硫酸钠(SDS)、1 mM EDTA 中于 65°C 杂交并且在 0.1 × SSC/0.1% SDS 中 68°C 洗涤;或中等严格条件包括在 0.2 × SSC/0.1% SDS 中于 42°C 洗涤)与所述 RNA 杂交形成稳定双链体的序列;在双链 BCMP 反义核酸的情况下,可能必须测试所述双链体 DNA 的一条链,或分析三链体的形成。杂交的能力既取决于互补的程度,也取决于所述反义核酸的长度。一般而言,所述杂交核酸越长,可能含有的与编码 BCMP 的 RNA 错配的碱基则越多,但仍形成稳定的双链体(或三联体,情况同样如此)。本领域技术人员可以利用测定所述杂交复合体解链温度的标准方法,来确定可耐受的错配程度。

### BCMP 反义核酸的治疗用途

当所述靶 BCMP 在怀疑患有或患有乳腺癌的受治疗者的乳腺组织中过量表达时,所述 BCMP 反义核酸可以用来治疗或预防乳腺癌。在一个优选实施方案中,使用单链 DNA 反义 BCMP 寡核苷酸。

表达或过量表达编码 BCMP 的 RNA 的细胞类型,可以采用本领域已知的各种方法来鉴定。这类细胞类型包括但不限于白细胞(例如嗜中性粒细胞、巨噬细胞、单核细胞)和居留细胞(例如星形胶质细胞、神经胶质细胞、神经元细胞和室管膜细胞)。这类方法包括但不限于与 BCMP 特异性核酸杂交(例如通过 RNA 印迹杂交、斑点印迹杂交、原位杂交)、观测来自所述细胞类型的 RNA 在体外翻译成 BCMP 的能力、免疫测定等。在一个优选方面,在治疗之前,可以例如通过免疫细胞化学或原位杂交,分析得自受治疗者的原代组织中的 BCMP 表达。

可以将包含药学上可接受的载体中的有效量的 BCMP 反义核酸的本发明药用组合物，给予患有乳腺癌的受治疗者。

在乳腺癌治疗方面有效的 BCMP 反义核酸的量，可以通过标准临床技术来确定。

在一个具体实施方案中，通过脂质体、微粒或微囊给予包含一种或多种 BCMP 反义核酸的药用组合物。在本发明的各个实施方案中，这类组合物可以用来达到所述 BCMP 反义核酸的缓释。在一个具体实施方案中，可能理想的是使用通过针对特定的可鉴定肿瘤抗原的抗体靶向的脂质体(Leonetti 等, 1990, Proc. Natl. Acad. Sci. USA 87:2448-2451; Renneisen 等, 1990, J. Biol. Chem. 265:16337-16342)。

#### 抑制性核酶和三股螺旋法

在另一实施方案中，通过利用编码所述 BCMP 的基因序列结合熟知的基因“敲除”、核酶或三股螺旋法降低 BCMP 的基因表达，来降低 BCMP 水平或 BCMP 活性，可以缓解乳腺癌的症状。在这种方法中，用核酶或三股螺旋分子来调节编码所述 BCMP 的基因的活性、表达或合成，因而缓解乳腺癌的症状。可以设计这类分子，使其降低或抑制突变型或非突变型靶基因的表达。用于生产和使用这类分子的技术是本领域技术人员众所周知的。

可以使用设计用于催化切割编码 BCMP 的基因 mRNA 转录物的核酶分子，防止靶基因 mRNA 的翻译，因而防止所述基因产物的表达。(参见例如 1990 年 10 月 4 日公开的 PCT 国际公开说明书 WO90/11346; Sarver 等, 1990, Science 247:1222-1225)。

核酶是能够催化 RNA 特异性切割的酶性 RNA 分子。(有关综述，参见 Rossi, 1994, Current Biology 4, 469-471)。核酶的作用机制涉及核酶分子与互补靶 RNA 的序列特异性杂交，后接内切核酸水解(endonucleolytic)切割事件。核酶分子的组成必须包括与靶基因 mRNA 互补的一个或多个序列，并且必须包括众所周知的负责 mRNA 切割的

催化序列。关于这种序列，参见例如美国专利第 5,093,246 号，所述专利通过引用全部结合到本文中。

虽然可以用在位点特异性识别序列切割 mRNA 的核酶，来破坏编码 BCMP 的 mRNA，但优选使用锤头型核酶。锤头型核酶在与靶 mRNA 形成互补碱基对的侧翼区所指示的位置切割 mRNA。唯一的要求是，靶 mRNA 具有两个碱基的以下序列：5'-UG-3'。锤头型核酶的构建和产生是本领域众所周知的，并且在以下文献中有更为全面的描述：Myers, 1995, *Molecular Biology and Biotechnology: A Comprehensive Desk Reference*, VCH Publishers, New York, (尤其参见图 4、第 833 页)，以及 Haseloff 和 Gerlach, 1988, *Nature*, 334, 585-591，每个所述文献通过引用全部结合到本文中。

最好对所述核酶进行工程改造，使得所述切割识别位点位于编码所述 BCMP 的 mRNA 的 5'端附近，即提高效率并且将非功能性 mRNA 转录物的胞内积累减至最小。

本发明的核酶也包括 RNA 内切核糖核酸酶(下文称为“Cech 型核酶”)，例如在嗜热四膜虫(*Tetrahymena thermophila*)中天然存在的 RNA 内切核糖核酸酶(称为 IVS 或 L-19 IVS RNA)和 Thomas Cech 及其同事广泛描述的 RNA 内切核糖核酸酶(Zaug 等, 1984, *Science*, 224, 574-578; Zaug 和 Cech, 1986, *Science*, 231, 470-475; Zaug 等, 1986, *Nature*, 324, 429-433; 由 University Patents Inc.公开的国际专利申请 WO 88/04300; Been 和 Cech, 1986, *Cell*, 47, 207-216)。Cech 型核酶具有一个 8 碱基对活性部位，所述活性部位与靶 RNA 序列杂交，此后发生靶 RNA 切割。本发明包括靶向所述 BCMP 编码基因中存在的 8 碱基对活性部位序列的那些 Cech 型核酶。

同反义方法中一样，所述核酶可以由经修饰的寡核苷酸组成(例如以改进稳定性、靶向等)，并且应该传递给在体内表达所述 BCMP 的细胞。一种优选的传递方法包括使用“编码”在强组成型 pol III 和 pol II 启动子控制之下的所述核酶的 DNA 构建体，使得转染细胞将产生足量

的所述核酶，以破坏编码所述 BCMP 的内源 mRNA，并且抑制翻译。因为核酶与反义分子不同，是催化性的，所以为达到功效所要求的胞内浓度较低。

内源 BCMP 的表达也可以通过采用定向同源重组使编码所述 BCMP 的基因或这种基因的启动子失活或“敲除”来降低(例如参见 Smithies 等, 1985, *Nature* 317:230-234; Thomas 和 Capecchi, 1987, *Cell* 51:503-512; Thompson 等, 1989, *Cell* 5:313-321; 和 Zijlstra 等, 1989, *Nature* 342:435-438, 每个所述文献通过引用全部结合到本文中)。例如, 可以用邻接与所述内源基因(所述 BCMP 编码基因的或者编码区或者调节区)同源的 DNA、编码非功能性 BCMP 的突变基因(或完全不相关的 DNA 序列)以及加上或不加上选择标记和/或负选择标记, 来转染体内表达所述靶基因的细胞。通过定向同源重组插入所述 DNA 构建体, 导致所述靶基因的失活。这类方法特别适合于农场, 在农场中可以用对 ES (胚胎干细胞)细胞的修饰来产生具有无活性靶基因的动物后代(例如参见 Thomas 和 Capecchi, 1987, 和 Thompson, 1989, 参见上文)。然而, 可以修改这种方法以用于人类, 条件是用合适的病毒载体将所述重组 DNA 构建体直接体内给予或靶向所需的部位。

或者, 编码 BCMP 的基因的内源表达的降低可以如下进行: 靶向与所述基因的调节区(即所述基因的启动子和/或增强子)互补的脱氧核糖核苷酸序列, 形成三股螺旋结构, 该结构阻止编码所述 BCMP 的基因在体内靶细胞内转录。(一般参见 Helene, 1991, *Anticancer Drug Des.*, 6(6), 569-584; Helene 等, 1992, *Ann. N.Y. Acad. Sci.*, 660, 27-36; 和 Maher, 1992, *Bioassays* 14(12), 807-815)。

用于形成三股螺旋以抑制转录的核酸分子应该是单链的, 并且由脱氧核糖核苷酸组成。必须设计这些寡核苷酸的碱基组成, 以促进通过 Hoogsteen 碱基配对原则形成三股螺旋, 这一般需要在双链体的一条链上存在相当大的或者嘌呤或者嘧啶序列段。核苷酸序列可以基于嘧啶, 这将产生跨越所得三股螺旋的三条相连链的 TAT 和 CGC<sup>+</sup>三联

体。富含嘧啶的分子提供与所述双链体中一条链的富含嘌呤区以与该链平行方向互补的碱基。另外，可以选择富含嘌呤(例如含有 G 残基的序列段)的核酸分子。这些分子将与富含 GC 对、其中大多数所述嘌呤残基位于所靶向双链体中一条链上的 DNA 双链体形成三股螺旋，产生跨越所述三联体中三条链的 GGC 三联体。

或者，通过构建所谓的“Z形(switchback)”核酸分子，可以增加靶向形成三股螺旋的潜在序列。Z型分子可以以交替的 5'-3'、3'-5'方式合成，使得它们首先与双链体的一条链进行碱基配对，然后与另一条链进行碱基配对，使得无需要求相当大的或者嘌呤或者嘧啶序列段存在于双链体中一条链上。

在其中使用本文所述的反义分子、核酶分子或三股螺旋分子抑制突变基因表达的情况下，有可能所述技术可以如此有效地降低或抑制 BCMP 正常基因的等位基因产生的 mRNA 的转录(三股螺旋)或翻译(反义分子、核酶)，使得可能产生其中 BCMP 的存在浓度可能低于正常表型的必需浓度的情况。在这种情况下，为了确保维持 BCMP 编码基因的基本正常水平或活性，可以用基因治疗来将编码并表达所述 BCMP 的核酸分子导入细胞中，其中所述核酸分子表现出正常的基因活性，并且不含无论对所使用的反义治疗、核酶治疗还是三股螺旋治疗都敏感的序列。或者，在所述基因编码胞外蛋白的情况下，可以共同给予正常 BCMP，以便维持 BCMP 活性所必需的水平。

本发明的反义 RNA 和 DNA、核酶和三股螺旋分子，可以用如上所述的本领域已知的用于合成 DNA 和 RNA 分子的任何方法来制备。这些包括本领域众所周知的用于化学合成寡脱氧核糖核苷酸和寡核糖核苷酸的技术，例如固相亚磷酸胺化学合成法。或者，通过编码所述反义 RNA 分子的 DNA 序列的体外和体内转录，可以产生 RNA 分子。可以将这类 DNA 序列加入到各种各样的加入了合适 RNA 聚合酶启动子(例如 T7 或 SP6 聚合酶启动子)的载体中。或者，根据所用的启动子，可以将组成型或诱导型合成反义 RNA 的反义 cDNA 构建体稳定导入

到细胞系中。

### 治疗或预防化合物的测定

本发明也提供用于药物发现以鉴定或证实化合物治疗或预防乳腺癌的功绩的测定。可以分析试验化合物将乳腺癌受治疗者的 BCMP 水平向无乳腺癌受治疗者的水平恢复的能力，或者在乳腺癌实验动物模型中产生相似变化的能力。能够将乳腺癌受治疗者的 BCMP 水平向无乳腺癌受治疗者的水平恢复或在乳腺癌实验动物模型中产生相似变化的化合物，可以用作可供进一步药物发现的前导化合物，或者用于治疗。BCMP 的表达可以通过以下测定来分析：the Preferred Technology、免疫测定、凝胶电泳然后显现、BCMP 活性的检测、或本领域技术人员已知或本文所述的任何其它方法。在其中 BCMP 的丰度可以用作临床疾病的替代标记的情况下，在临床监测或药物开发方面，这类测定可以用来筛选候选药物。

在各种具体实施方案中，可以用涉及受治疗者疾病的细胞类型的代表细胞来进行体外测定，以确定所述化合物对这类细胞类型具有所需效应。

在进行人体试验之前，可以在合适的动物模型系统中测试用于治疗的化合物，所述动物模型包括但不限于大鼠、小鼠、鸡、牛、猴子、兔子等。对于体内试验，在给予人体之前，可以使用本领域已知的任何动物模型系统。乳腺癌的动物模型的实例包括但不限于人乳腺癌细胞系(例如 MDA-MB-435)在缺乏雌激素的严重联合免疫缺陷(SCID)小鼠中的异种移植物(Eccles 等, 1994, Cell Biophysics 24/25, 279)。这些可以用来测试调节 BCMP 水平的化合物，因为在这些动物模型中表现出的病状与乳腺癌的病状相似。根据本文公开的内容，对于本领域技术人员显而易见的是，可以用编码一种或多种 BCMP 的一种或多种基因的“敲除”突变，产生转基因动物。基因的“敲除”突变是一种使得突变基因不表达或者以异常形式表达或以低水平表达、使得与所述基

因产物相关的活性几乎没有或完全缺乏的突变。所述转基因动物最好是哺乳动物，更优选所述转基因动物是小鼠。

在一个实施方案中，在表达 BCMP 的非人类动物(例如小鼠、大鼠、猴、兔子和豚鼠)、优选乳腺癌的非人类动物模型中，鉴定调节所述 BCMP 表达的试验化合物。按照该实施方案，将试验化合物或对照化合物给予所述动物，测定所述试验化合物对一种或多种 BCMP 表达的影响。通过比较用试验化合物治疗的一只动物或一组动物中所选一种或多种 BCMP (或其编码 mRNA)的水平与用对照化合物治疗的一只动物或一组动物中的所述 BCMP 或 mRNA 的水平，可以鉴定出改变所述 BCMP (或多种 BCMP)表达的试验化合物。可以用本领域技术人员已知的技术来测定所述 mRNA 和蛋白质的水平，例如用原位杂交法。可以处死或不处死所述动物，以分析试验化合物的效应。

在另一实施方案中，在表达 BCMP 的非人类动物(例如小鼠、大鼠、猴、兔子和豚鼠)、优选乳腺癌的非人类动物模型中，鉴定调节所述 BCMP 或其生物活性部分的活性的试验化合物。按照该实施方案，将试验化合物或对照化合物给予所述动物，测定所述试验化合物对所述 BCMP 活性的影响。通过分析用对照化合物治疗的动物和用试验化合物治疗的动物，可以鉴定出改变所述 BCMP (或多种 BCMP)活性的试验化合物。通过检测所述 BCMP 的细胞第二信使(例如胞内  $Ca^{2+}$ 、二酰基甘油、IP3 等)的诱导、检测所述 BCMP 或其结合配偶体的催化活性或酶活性、检测报道基因(例如对本发明 BCMP 应答并且与编码可检测标记例如荧光素酶或绿色荧光蛋白的核酸有效连接的调节元件)的诱导、或者检测细胞应答(例如细胞分化或细胞增殖)，可以评价所述 BCMP 的活性。可以用本领域技术人员已知的技术来检测 BCMP 活性的改变(参见例如美国专利第 5,401,639 号，其通过引用结合到本文中)。

在另一实施方案中，在患有乳腺癌的人类受治疗者、最好是患有重症乳腺癌的人类受治疗者中，鉴定调节 BCMP (或多种 BCMP)的水

平或表达的试验化合物。按照该实施方案,将所述试验化合物或对照化合物给予所述人类受治疗者,并且通过分析生物样品(例如乳腺组织、血清、血浆或尿)中所述 BCMP 或编码所述 BCMP 的 mRNA 的表达,测定试验化合物对所述 BCMP 表达的影响。通过比较用对照化合物治疗的一位受治疗者或一组受治疗者体内 BCMP 或编码所述 BCMP 的 mRNA 的水平与用试验化合物治疗的一位受治疗者或一组受治疗者体内的所述水平,可以鉴定出改变所述 BCMP 表达的试验化合物。或者,通过比较在给予试验化合物之前和之后的一位受治疗者或一组受治疗者体内 BCMP 或编码所述 BCMP 的 mRNA 的水平,可以鉴定出所述 BCMP 表达的改变。可以用本领域技术人员已知的技术,获得所述生物样品,并分析所述 mRNA 或蛋白质的表达。例如,可以用本文所述的 the Preferred Technology 来评价 BCMP 水平的改变。

在再一实施方案中,在患有乳腺癌的人类受治疗者(最好是患有重症乳腺癌的人类受治疗者)中,鉴定调节 BCMP (或多种 BCMP)活性的试验化合物。在该实施方案中,将试验化合物或对照化合物给予所述人类受治疗者,测定试验化合物对 BCMP 活性的影响。通过比较得自用对照化合物治疗的受治疗者的生物样品与得自用所述试验化合物治疗的受治疗者的生物样品,可以鉴定出改变所述 BCMP 活性的试验化合物。或者,通过比较在给予试验化合物之前和之后的一位受治疗者或一组受治疗者体内 BCMP 的活性,可以鉴定出 BCMP 活性的改变。通过检测生物样品(例如乳腺组织、血清、血浆或尿)中所述 BCMP 的细胞信号转导途径(例如胞内  $\text{Ca}^{2+}$ 、二酰基甘油、IP3 等)的诱导、所述 BCMP 或其结合配偶体的催化活性或酶活性、或者细胞应答(例如细胞分化或细胞增殖),可以评价所述 BCMP 的活性。可以用本领域技术人员已知的技术,检测 BCMP 的第二信使诱导的改变、或细胞应答的改变。例如,可以用 RT-PCR 检测细胞第二信使诱导的改变。

在一个优选实施方案中,选择将 BCMP 的水平或表达朝在对照受治疗者(例如无乳腺癌的人)中检测到的水平改变的试验化合物,以供

进一步试验或治疗应用。在另一优选实施方案中，选择将 BCMP 活性朝在对照受治疗者(例如无乳腺癌的人)中检测到的活性改变的试验化合物，以供进一步试验或治疗应用。

在另一实施方案中，在患有乳腺癌的人类受治疗者、最好是患有重症乳腺癌的人类受治疗者中，鉴定降低一个或多个与乳腺癌相关症状的严重程度的试验化合物。按照该实施方案，将试验化合物或对照化合物给予所述受治疗者，并且测定试验化合物对乳腺癌的一个或多个症状的影响。通过比较用对照化合物治疗的受治疗者与用试验化合物治疗的受治疗者，可以鉴定出缓解一个或多个症状的试验化合物。可以用熟悉乳腺癌的医师已知的技术，确定试验化合物是否缓解与乳腺癌相关的一个或多个症状。例如，减小患有乳腺癌的受治疗者体内所带肿瘤的试验化合物，将有益于患有乳腺癌的受治疗者。

在一个优选实施方案中，选择在患有乳腺癌的病人中降低一个或多个与乳腺癌相关症状的严重程度的试验化合物，以供进一步试验或治疗应用。

### 治疗和预防组合物及其应用

本发明提供治疗(和预防)方法，所述方法包括给予受治疗者有效量的本发明化合物。在一个优选方面，所述化合物是基本纯化的(例如基本上无限制其效应或产生不想要的副作用的物质)。所述受治疗者最好是动物，包括但不限于诸如牛、猪、马、鸡、猫、狗等的动物，优选哺乳动物，最优选人类。在一个具体实施方案中，非人类哺乳动物是所述受治疗者。

当所述化合物包含核酸时可以使用制剂和给药方法如上所述；以下描述其它合适的制剂和给药途径。

各种传递系统是已知的，可以用来给予本发明的化合物，例如在脂质体、微粒、微囊中包囊、能够表达所述化合物的重组细胞、受体介导胞吞作用(参见例如 Wu 和 Wu, 1987, J. Biol. Chem. 262:4429-

4432)、作为反转录病毒或其它载体的部分的核酸的构建等。导入方法可以是肠溶或胃肠外途径,包括但不限于皮内、肌内、腹膜内、静脉内、皮下、鼻腔、硬膜外和口服途径。所述化合物可以通过任何方便的途径给予,例如通过输注或大剂量注射、通过上皮或粘膜内衬(例如口腔粘膜、直肠和肠粘膜等)吸收,并且可以与其它生物活性剂一起给予。可以系统给药或局部给药。另外,可能理想的是通过任何合适的途径(包括脑室内注射和鞘内注射)将本发明的药用组合物导入中枢神经系统;通过例如与贮器(例如 Ommaya 贮器)连接的脑室内插管,可以便于进行脑室内注射。也可以利用肺部给药,例如利用吸入器或雾化器和具有气雾化剂的制剂。

在一个具体实施方案中,可能理想的是将本发明的药用组合物局部给予需要治疗的区域;这可以例如通过例如但不限于下述方法达到:在手术期间局部输注、局部应用、例如通过注射、借助插管或利用植入物,所述植入物是多孔、非多孔或凝胶状材料,包括膜例如 sialastic 膜或纤维。在一个实施方案中,可以通过直接注射到乳腺组织或恶性肿瘤或肿瘤或肿瘤前组织的部位(或肿瘤前(former)部位)进行给药。

在另一实施方案中,所述化合物可以在泡囊、特别是在脂质体中给药(参见 Langer, 1990, Science 249:1527-1533; Treat 等, Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein 和 Fidler (编著), Liss, New York, 第 353-365 页(1989); Lopez-Berestein, 出处同上, 第 317-327; 一般参见同一文献)。

在再一实施方案中,所述化合物在控释系统中给予。在一个实施方案中,可以使用一种泵(参见 Langer, 参见上文; Sefton, 1987, CRC Crit. Ref. Biomed. Eng. 14:201; Buchwald 等, 1980, Surgery 88:507; Saudek 等, 1989, N. Engl. J. Med. 321:574)。在另一实施方案中,可以使用聚合材料(参见 Medical Applications of Controlled Release, Langer 和 Wise (编著), CRC Pres., Boca Raton, Florida (1974); Controlled Drug

Bioavailability, Drug Product Design and Performance, Smolen 和 Ball (编著), Wiley, New York (1984); Ranger 和 Peppas, 1983, J. Macromol. Sci. Rev. Macromol. Chem. 23:61; 也参见 Levy 等, 1985, Science 228:190; During 等, 1989, Ann. Neurol. 25:351; Howard 等, 1989, J. Neurosurg. 71:105)。在又一实施方案中, 可以将控释系统置于治疗靶(例如胸部)附近, 从而仅需要系统剂量的一部分(参见例如 Goodson, Medical Applications of Controlled Release, 参见上文, 第 2 卷, 第 115-138 页 (1984))。

在 Langer 的综述(1990, Science 249:1527-1533)中论述了其它控释系统。

在本发明化合物为编码蛋白的核酸的一个具体实施方案中, 通过将所述核酸作为合适核酸表达载体的一部分来构建, 并且例如利用反转录病毒载体(参见美国专利第 4,980,286 号)或通过直接注射或利用微粒轰击(例如基因枪; Biolistic, Dupont), 或包上脂质或细胞表面受体或转染剂, 或通过将其与已知进入细胞核的同源框样肽(参见例如 Joliot 等, 1991, Proc. Natl. Acad. Sci. USA 88:1864-1868)连接给予等, 给予所述载体使其成为胞内的所述核酸可以体内给予, 以促进其编码蛋白的表达。或者, 可以将核酸导入细胞内, 并通过同源重组掺入到宿主细胞 DNA 中以供表达。

本发明也提供药用组合物。这类组合物包含治疗有效量的化合物和药学上可接受的载体。在一个具体实施方案中, 术语“药学上可接受的”是指由联邦政府或州政府的管理机构批准的或在美国药典或用于动物(更优选用于人类)的其它一般承认的药典中列出的。术语“载体”是指与所述治疗药一起给予的稀释剂、辅料、赋形剂或载体。这类药用载体可以是无菌液体, 例如水和油, 包括石油来源、动物来源、植物来源或合成来源的油, 例如花生油、豆油、矿物油、芝麻油等。当静脉内给予所述药用组合物时, 水是优选的载体。盐水溶液和葡萄糖水溶液和甘油溶液也可以作为液体载体使用, 特别是对于注射液而

言。合适的药用赋形剂包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、稻米、面粉、白垩粉、硅胶、硬脂酸钠、甘油一硬脂酸酯、滑石粉、氯化钠、脱脂奶粉、甘油、丙二醇、水、乙醇等。如果需要，所述组合物也可以含有少量润湿剂或乳化剂、或 pH 缓冲剂。这些组合物可以采用溶液剂、混悬剂、乳剂、片剂、丸剂、胶囊剂、散剂、缓释制剂等的形式。所述组合物可以用传统粘合剂和载体例如甘油三酯配制为栓剂。口服制剂可以包括标准载体例如药用级甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维素、碳酸镁等。在 E.W. Martin 的“Remington’s Pharmaceutical Sciences”中描述了合适的药用载体的实例。这类组合物将含有治疗有效量的所述化合物、最好是纯化形式的所述化合物与适量的载体，以便提供适当给予所述受治疗者的形式。所述制剂应该适合给药模式。

在一个优选实施方案中，所述组合物按照常规方法配制为适用于静脉内给予人类的药用组合物。通常，用于静脉内给药的组合物是无菌等渗水性缓冲液中的溶液。必要时，所述组合物也可以包括增溶剂和局部麻醉药例如利多卡因，以缓解注射部位的疼痛。一般而言，所述成分或者分开供应，或者在单位剂型中混合在一起，例如作为冻干粉或密封容器(例如安瓿或指示有效成分数量的小药囊)中的无水浓缩物。当所述组合物通过输注给予时，它可以用含有无菌药用级水或盐水的输液瓶调配。当所述组合物通过注射给予时，可以提供注射用无菌水或盐水的安瓿，使得可以在给药前混合所述成分。

本发明的化合物可以以中性形式或盐形式配制。药学上可接受的盐包括用游离氨基形成的盐例如衍生于盐酸、磷酸、乙酸、草酸、酒石酸等的盐以及与游离羧基形成的盐，例如由氢氧化钠、氢氧化钾、氢氧化铵、氢氧化钙、氢氧化铁、异丙胺、三乙胺、2-乙氨基乙醇、组氨酸、普鲁卡因等衍生的盐。

有效治疗乳腺癌的本发明化合物的量可以通过标准临床技术来确定。另外，可以任选地用体外测定，帮助确定最适剂量范围。所述

制剂中的精确用量也取决于给药途径以及所述疾病或疾患的严重程度，应该根据医师的判断和每个受治疗者的状况来决定。然而，静脉内给药的合适剂量范围一般约为 20-500 微克活性化合物/kg 体重。鼻腔给药的合适剂量范围一般约为 0.01 pg/kg 体重至 1 mg/kg 体重。有效量可以从得自体外或动物模型试验系统的剂量反应曲线来外推。

栓剂一般含有范围为 0.5%-10% (重量) 的有效成分；口服制剂最好含有 10%-95% 的有效成分。

本发明也提供药用包装或试剂盒，所述包装或试剂盒包含一个或多个容器，所述容器装有本发明药用组合物的一种或多种成分。这种容器可以任选地伴有通知，所述通知是管理生产、药物或生物制品的使用或销售的政府机构规定的形式，所述通知反映出(a)所述机构批准生产、使用或销售用于人类给药，(b)使用说明，或这两者。

本发明每个方面的优选特征就每个其它方面而言在细节上作必要的修改。本文提及的现有技术文件在法律允许的最大程度上结合到本文中。

### **实施例 1: 在乳腺癌细胞系中表达的膜蛋白的鉴定**

采用以下参考方案，通过 SDS-PAGE 分离乳腺癌细胞系膜中的蛋白质，并对其进行分析。

#### **材料和方法**

##### **1a - 粘附乳腺癌细胞系的粗制分级分离**

将人乳腺癌细胞系 MDA-MB-468 (ATCC:HTB-132)、T47D (ATCC:HTB-133) 和 BT20 (ATCC:HTB-19) 在不同组织培养条件下培养 (对于 BT20 细胞系: 培养基为改进的 Eagle 氏培养基、10% 胎牛血清、非必需氨基酸混合物、2mM 谷氨酰胺、1% 青霉素 + 链霉素。对于 T47D 和 MDA-MB-468 细胞系: 培养基为 DMF12、10% 胎牛血清、2mM 谷氨酰胺、1% 青霉素 + 链霉素。细胞于 37°C、5% CO<sub>2</sub> 中在上述培养基

中培养)。按照下述方案裂解细胞并进行分级分离。

将 5 ml 冷匀浆缓冲液(50 mM TrisHCl, 250 mM 蔗糖, 1mM EDTA pH 7.4)与抗蛋白酶(Sigma 蛋白酶抑制剂混合物 P2714)一起加入 10 个含有  $10^8$  乳腺癌细胞(约 10mg 蛋白)的 15 cm<sup>2</sup> 培养皿的系列中。将平板中的细胞刮到冷匀浆缓冲液中, 转移到 5 ml 螺帽小玻璃瓶(在冰上进行)中。用带有平底探头的 MSE Soniprep 150 对所得样品以 5 微米的振幅超声处理(在冰上)10 秒。然后, 将样品倾析到 15 ml falcon 管中, 以 1000g 于 4℃ 离心 5 分钟。将所得的上清液转移到 11 × 60 mm 透明超离心管中, 小心不要扰动沉淀(不溶性蛋白), 向管中加匀浆缓冲液至几乎加满。然后将所述管以 100 000 × g 于 4℃ 离心 1 小时, 然后将其置于冰上, 除去上清液(细胞溶胶)。

沉淀(膜部分)用冰冷的 PBS 温和洗涤 3 次, 重悬于含有抗蛋白酶的洗涤缓冲液(50 mM TrisHCl, 1mM EDTA pH 7.4)中。然后在玻璃匀浆器温和研磨(5-10 个冲程), 转移到干净的超离心管中, 以 100 000 × g 于 4℃ 离心 1 小时。然后重复这最后一步, 用 0.5 M NaCl 洗涤, 以除去周边蛋白。得自  $10^8$  乳腺癌细胞的经洗涤膜蛋白的产量约为 1-2 mg 蛋白。

#### 1b - 用去垢剂 Tx114 提取经洗涤的膜蛋白

这种提取提供三个潜在部分: 去垢剂不溶性蛋白、膜周边蛋白和膜内在蛋白。

如以上 1a 中所述制备粗制细胞膜, 用 Dounce 匀浆器通过匀浆将其重悬于 50mM Tris HCl、1mM EDTA、1.5% Triton X114、蛋白酶抑制剂 pH 7.4 中。

将 Triton X114 (Tx114)膜混合物涡旋混合, 在冰上孵育 30 分钟。此后, 将 Tx114 膜混合物以 13000 g 于 4℃ 离心 10 分钟, 小心从任何去垢剂不溶性蛋白沉淀中提取上清液。

将 Tx114 上清液在水浴中温热至 37℃ 达 3 分钟, 然后以 13000 g

离心 3 分钟。除去上层水层，将下层去垢剂相重悬于 1ml Tris HCl、0.2mM EDTA pH 7.4 中。将这一提取步骤再重复 2 次。

水相代表亲水性膜蛋白，去垢剂相代表疏水性膜蛋白。

通过以下 1d 中所述的氯仿-甲醇提取，从水层和去垢剂相中提取蛋白。得自  $10^8$  细胞的亲水性膜蛋白的产量约为 0.5 mg，疏水性膜蛋白的产量约为 0.1-0.2 mg。最后，将所述膜蛋白溶于约 30 微升 1D 裂解缓冲液中(1-2  $\mu\text{g}/\mu\text{l}$ )。

#### 1c - 用去垢剂毛地黄皂苷提取经洗涤的膜蛋白

用含去垢剂的缓冲液如下提取得自 1a 的经洗涤的膜沉淀。

将冰冷的毛地黄皂苷缓冲液(0.01%的 50mM Tris HCl pH 7.4 溶液)加入到所述膜部分中，将所得溶液匀浆 10 个冲程，然后置于冰上 30 分钟，以让蛋白溶解。

然后将样品以  $13000 \times g$  于  $4^\circ\text{C}$  离心 5 分钟，以沉淀任何不溶性蛋白，用如下 1d 所述的氯仿-甲醇提取法提取上清液。得自  $10^8$  癌细胞的毛地黄皂苷提取的蛋白的产量约为 30-50 微克。

#### 1d - 用氯仿/甲醇从水溶液或去垢剂溶液中提取蛋白

将 400 $\mu\text{l}$  甲醇、100 $\mu\text{l}$  氯仿和 300 $\mu\text{l}$  水加入到 100-200 $\mu\text{l}$  得自 1b 和 1c 的去垢剂溶液中。将所得混合物涡旋混合并震荡 60 秒，然后以  $13000 \times g$  于  $20^\circ\text{C}$  离心 10 分钟。

产生两层，所需蛋白在所述两层之间的界面中。小心除去上层，加入 300 $\mu\text{l}$  甲醇，将离心管倒置，然后以  $13000g$  于  $4^\circ\text{C}$  离心 5 分钟，以沉淀蛋白。除去甲醇，让沉淀风干 5-10 分钟。

此后，将沉淀再溶于 1D 裂解缓冲液(63mM Tris.HCl pH 7.4、10% 甘油、2% SDS、0.0025% 溴酚蓝、2% 巯基乙醇)中。

## 1e - 1D 凝胶技术

将蛋白或膜沉淀溶于 1D 样品缓冲液中(1-2 $\mu\text{g}/\mu\text{l}$ )。然后将样品缓冲液和蛋白的混合物加热至 95 $^{\circ}\text{C}$  达 3 分钟。

用浓缩胶和浓缩胶梳子, 按照 Ausubel F.M 等编著, 1989, *Current Protocols in Molecular Biology*, 第 II 卷, Green Publishing Associates, Inc. 和 John Wiley & Sons, Inc., New York, 第 10.2 节(该文献通过引用全部结合到本文中)中所述的程序, 灌制 9-16% 丙烯酰胺梯度凝胶。

用 10 微升移液器吸头将得自去垢剂处理的 30-50 微克蛋白混合物和分子量标准(66、45、31、21、14 kDa)加入到浓缩胶孔中, 将样品以 40mA 电泳 5 小时。

然后撬开板, 将凝胶置于盘中的固定液(10%乙酸、40%乙醇、50%水)中, 震荡过夜。此后, 在引发(primer)溶液(7.5%乙酸(75ml)、0.05% SDS (5ml 10%))中将凝胶震荡引发(prime) 30 分钟。然后将凝胶与荧光染料(7.5%乙酸、0.06% OGS in-house 染料(600 $\mu\text{l}$ ))一起震荡温育 3 小时。Sypro Red (Molecular Probes, Inc., Eugene, Oregon)是一种用于此目的合适染料。一种优选的荧光染料公开于 1999 年 10 月 5 日申请的美国申请第 09/412,168 号中, 该申请通过引用全部结合到本文中。

经荧光染色的凝胶用 Apollo 3 扫描器(Oxford Glycosciences, Oxford, UK)成像, 产生计算机可读输出。这种扫描器由 WO 96/36882 中和 David A. Basiji 的题为“利用内部反射光学和相位灵敏检测的高通量荧光扫描器的开发(全内部反射, 电泳)”, University of Washington (1997), *Dissertation Abstracts International* 的第 58/12-B 卷, 第 6686 页(所述文献的内容通过引用结合到本文中)中描述的扫描器发展而来。该仪器的最新形式包括以下改进: 凝胶在精密引导螺杆驱动系统上传送通过扫描器。优选将玻璃板置于 Basiji 论文中所述的皮带传动系统上, 因为它提供一种准确传送凝胶通过成像光学系统的可再现方法。

在所述扫描器中, 凝胶用三个对准止动件固定, 将玻璃板大致保持在已知位置。借此并且连同上述精密传送系统以及凝胶结合于玻璃

板上的事实，可以预测并记录凝胶的绝对位置。这确保将凝胶上每一特征的准确坐标传送至切割机器人以进行切取。这种切割机器人对于玻璃板有相同的安装配置，以保持位置的准确。

将凝胶保持在适当位置的载体有固有的荧光标记(称为 M1、M2、M3)，它们可以用来校正图象的几何形状，并且是一种质量控制特征，以证实扫描正常进行。

该系统的光学部件是倒置的。激光、反射镜、波导和其它光学部件在被扫描的玻璃板之上。Basiji 论文的实施方案中，这些部件在玻璃板之下。因此，玻璃板固定在扫描器上，凝胶面向下，使得光路仍通过玻璃板。借此，可能离开玻璃板的任何凝胶颗粒都将落在仪器的底部，而不是落到光学器件中。

在扫描凝胶时，将凝胶从染料中取出，用水冲洗并让其短暂风干，并在 Apollo 3 上成像。成像之后，将凝胶密封在含有小体积染色溶液的聚乙烯袋中，然后贮存于 4℃。

通过从在样品旁边电泳的一组已知分子量标记内推，计算表观分子量。

#### 1f- 所选蛋白的回收和分析

通过美国专利第 6,064,754 号第 5.4 节和第 5.6、5.7、5.8 节(所述文献通过引用结合到本文中)中描述的适用于 1D 电泳的方法，由机器人从凝胶中切取蛋白质，对所述方法的机器人切割手如下作出修改：该切割手在所述泳道顶部开始，从该泳道的左边边缘开始切割直径为 1.7mm 的凝胶片，然后切割手向右边移动 2mm，向下移动 0.7mm，再切下另一凝胶片。然后重复这一步。切割手然后返回位于第一次凝胶切割正下方的位置，但向下偏转 2.2mm，重复所述三个斜线切割的模式。在凝胶的全长上继续进行这一模式。

注意：若观测到泳道显著变宽，则也可以向旁边作出校正，即不是返回到位于前次凝胶切割的正下方的位置，切割可以略微向左边(泳

道的左边)和/或右边(泳道的右边)偏转。对凝胶片内含有的蛋白进行加工,以产生胰蛋白酶肽;如 WO98/53323 和 1998 年 6 月 15 日申请的申请第 09/094,996 号中描述的质谱法,测定这些肽的部分氨基酸序列。

对蛋白进行加工,产生胰蛋白酶消化肽。用 PerSeptive Biosystems Voyager-DETM STR 基质辅助激光解吸电离飞行时间(MALDI-TOF)质谱仪通过质谱法分析胰蛋白酶肽,通过串联质谱法(MS/MS),用配有 nanoflowTM 电喷雾 Z-喷雾源的 Micromass 四极时间飞行(Q-TOF)质谱仪(Micromass, Altrincham, U.K.),分析所选的胰蛋白酶肽。对于部分氨基酸测序和 BCMP 的鉴定,用 SEQUEST 搜索程序(Eng 等, 1994, J. Am. Soc. Mass Spectrom. 5:976-989), version v.C.1, 搜索胰蛋白酶肽的无解释串联质谱。数据库鉴定的标准包括:胰蛋白酶的切割特异性;从数据库返回的肽中一组 a、b 和 y 离子的检测、和说明脲基甲基化的所有 Cys 残基的质量增加。搜索的数据库是在由 the National Centre for Biotechnology Information (NCBI)维护的非冗余数据库中由蛋白实体构成的数据库,所述数据库可在 <http://www.ncbi.nlm.nih.gov/> 上接入。在用 SEQUEST 程序通过质谱-质谱相关鉴定蛋白质后,将 MALDI-TOF 质谱中检测到的质量赋予所鉴定的蛋白中的胰蛋白酶消化肽。运用 SEQUEST 程序通过搜索胰蛋白酶消化肽的无解释 MS/MS 谱不能鉴定出氨基酸序列的情况下,采用本领域已知的方法人工解释所述肽的串联质谱(在解释肽离子的低能碎裂质谱的情况下,参见 Gaskell 等, 1992, Rapid Commun. Mass Spectrom. 6:658-662)。

## 结果

这些初步实验鉴定出匹配尚未描述蛋白或生物功能的 cDNA 概念翻译的 14 种 BCMP (表 1)和匹配未曾在乳腺细胞膜中描述的已知蛋白的 37 种 BCMP (表 2)、与 4 种其它 BCMP (表 3)。

## 实施例 2: 通过胰蛋白酶消化蛋白片段的质谱法衍生串联质谱对基因组 DNA 序列中的基因作图

### 质谱法

对于 BCMP81, 通过 MALDI-TOF-MS (Voyager STR, Applied Biosystems, Framingham, MA), 用 337 nm 波长激光进行解吸和分析的反射模式, 分析从 1D 凝胶切取的蛋白的胰蛋白酶消化物。所选质谱 ( $[M+H] = 1557.81$ ) 通过串联质谱法, 用配有纳米喷雾离子源的 QTOF-MS (Micromass UK Ltd, Manchester) 进一步表征。在 MALDI 分析之前, 用 C18 Zip Tips™ (Millipore, Bedford, MA) 将样品脱盐和浓缩。用于串联 MS 的样品用加入 C18 SPE 材料的纳米 LC 系统 (LC Packings, Amsterdam, The Netherlands) 纯化。

人工分析串联质谱 (Biemann K, 通过串联质谱法和高能碰撞诱导解离对肽测序, *Methods Enzymol* 1990; 193:455-79), 以测定部分位置的氨基酸序列以及所述肽片段的 N 末端和 C 末端其余部分的质量 (表 4)。

表 4. 得自 BCMP 81 的串联质谱法分析的氨基酸序列信息

肽先驱离子 m/z	核心序列 <sup>a</sup>	N 末端质量 <sup>b</sup>	C 末端质量 <sup>c</sup>
1557.81	MEQQQQLQQR	241.2	0.00

<sup>a</sup> “核心序列” 是由所述肽的碎片质谱的解释阐明的肽的部分氨基酸序列。

<sup>b</sup> 所述肽的 N 末端质量是核心序列起点和所述肽 N 末端之间的质量。

<sup>c</sup> 所述 C 末端质量是核心序列末端和所述肽 C 末端之间的质量。

发现从串联质谱读出的序列 (MEQQQQLQQR) 存在于登记号为 AL021707 的人基因组 DNA (可得自 <http://www.ncbi.nlm.nih.gov/entrez/>) 的翻译中, 从核苷酸位置 55443 开始 (BLAST, Altschul, Stephen F., Gish, Warren, Miller, Webb, Myers, Eugene W. 和 Lipman, David, J. (1990). Basic 局部比对搜索工具。J. Mol. Biol. 215; 403-410. Gapped BLAST 描述于 Altschul, Stephen F., Madden, Thomas L., Schaffer, Alejandro A., Zhang, Jinghui Zhang, Zheng, Miller, Webb 和 Lipman, David J. (1997).

GappedBLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17); 3389-3402).

对该位置周围序列的进一步检查, 鉴定出可以翻译匹配另一 MALDI 质谱的 DNA 序列。将这一质谱转换为概念氨基酸序列 (LWGLTEMFPER), 其质量匹配可以从所述 DNA 序列(图 1)翻译出的潜在胰蛋白酶片段。因此, 根据读框确定了基因组 DNA 的这一序列段的两个区段, 例如得自该串联质谱的序列紧接所述基因组 DNA 翻译中的终止密码子之后, 表明或者这是该基因的末端, 或者在该位置有一个内含子。然而, 任何潜在的内含子都必须始于该串联质谱 C 末端的 R 残基之后。

将对应于匹配该串联质谱的序列任一侧 2kb 的基因组 DNA 序列, 用于针对已表达序列标志(EST) (<http://www.ncbi.nlm.nih.gov>)的 blast 搜索中。EST AA316462 以 >99% 匹配该基因组序列的多个区, 有助于内含子和外显子边界的鉴定。这一比对鉴定出 4 个外显子, 但该 EST 序列没有足够的质量, 来测定所述完整基因的蛋白序列, 它们也没有好到足以精确定位内含子/外显子边界, 特别是在缺乏任何同源序列的情况下。一旦将 AL021707 中的 54575 (DDEe(ag)/gt)概念剪接为 54805 (e(ag)/LDET)时(这些剪接位点的位置已经由所述 EST 中测定), 最后一个质谱可以匹配基因组序列 (GDAEKPEEELEEDDDEELDETLSE)的翻译。

有一个 CpG 岛始于 AL021707 中的位置 54283。人基因组 DNA 序列中的 CpG 岛通常指示基因可以开始的区域(例如 Kundu TK, Rao MR, 染色质组构中的 CpG 岛和基因表达。 *J. Biochem (Tokyo)* 1999-年 2 月; 125(2):217-22)。一个可读框显然从 AL021707 的核苷酸 54461 以 ATG (甲硫氨酸)开始。根据 AA316462 的 3' EST 序列通过质谱建立的读框, 确定了该基因外显子 4 中的终止密码子。这种质谱、基因组 DNA 序列和 EST 序列的组合, 得出以下给出的要求保护的蛋白序列(图 1)。以下给予了对应的 DNA 序列(图 1)。在图 2 中给出了所述基因组片段

内的所述谱序列、内含子和外显子的位置。

鉴定出一个全长克隆

### 总 RNA 的制备

用 Trizol (Life Technologies), 从得自细胞系 MB-MDA468、BT20、PC30、PC3M、T47D 和 Cal51 的培养物的约  $10^8$  细胞制备总 RNA, 以约  $1\mu\text{g}/\mu\text{l}$  重悬浮。

### cDNA 克隆

用引物 BCMP 81 F (5'cctacagtcattggctgccgc3') 和 BCMP 81 R (5'gagacagctcaaacagcaataac3'), 30ng 乳腺 cDNA (Clontech) 和 PCR 试剂 (Qiagen, UK), 采用循环参数:  $95^{\circ}\text{C}$  15 秒、 $55^{\circ}\text{C}$  1 分钟 40 个循环, 通过 PCR 扩增 BCMP 81 的预测编码区。获得单一的 500bp 产物, 将其从其它 PCR 试剂 (Qiagen, UK) 中经柱纯化, 用引物 BCMP 81 F 和 BCMP 81 R 测序 (ABI377, Oxford University Contract Sequencing)。

已经从乳腺 cDNA 文库中克隆出编码 BCMP 81 的全长 cDNA 片段, 对其测序, 以证实身份和读框。根据蛋白质质量梯 (mass ladder) 确定, 从 1D 凝胶切下的蛋白的分子量在 14kDa 和 19kDa 之间, 概念翻译的全长蛋白的预测质量为 15.3kDa。

虽然在 GenBank 的注释中在所述肽/基因组 DNA 匹配的区域中容易鉴定并注意到 CpG 岛, 但从这些 DNA 序列没有预测出蛋白。所述肽数据已经能够全面描述所述蛋白、被翻译的 mRNA 和编码其的该基因的基因组结构。

### 正常人组织以及乳腺癌和前列腺癌细胞系中 BCMP 81 的 mRNA 分布型

采用实时 RT-PCR, 定量测量正常人组织 mRNA (Clontech) 和乳腺癌细胞系中的 BCMP 81 表达。使用寡聚 dT 引物和 Superscript II RT 试

剂盒(Life technologies), 从 1-5 $\mu$ g 总 RNA 合成 cDNA。在 ABI7700 序列测定系统(PE Biosystems, UK)上, 分析含有 10ng cDNA、10pmoles 有义引物 (5'gagctagatgagaccctgtc3') 和反义引物 (5'gatcataaaggaagtcccaa3')和 SYBR green 序列测定试剂(PE Biosystems, UK)的反应物。PCR 的条件是: 50 $^{\circ}$ C 2 分钟 1 个循环, 95 $^{\circ}$ C 10 分钟 1 个循环, 然后 95 $^{\circ}$ C 15 秒、55 $^{\circ}$ C 1 分钟 40 个循环。以报告染料荧光的增加实时测量 PCR 产物的积累, 用 Sequence Detector 程序 v1.6.3 (PE Biosystems)分析数据。用所述经扩增的 PCR 产物作为模板, 产生原始模板拷贝数与荧光和扩增循环数相关的标准曲线, 用该曲线来计算每种样品中的 BCMP 81 mRNA 的拷贝数。

通过定量 RT-PCR (Heid, C.A., Stevens, J., Livak, K.J.和 Williams, P.M.实时定量 PCR. *Genome Res.* 6, 986-994 (1996))对所述 mRNA 的组织分布进行的分析, 证实了所述蛋白来源(MDA-MB468)以及其作为乳腺癌抗原的潜力, BT20 和 MDA-MB468 细胞系都表现出与正常乳腺相比其表达显著增加(图 3)。该蛋白的公共数据库 EST 形式的微少反映在这些样品中观测到的一般 mRNA 拷贝数少(一般 <200 拷贝/ng cDNA)。

### 实施例 3: 新蛋白 BCMP 11 的克隆

#### 质谱法

从 1D 凝胶切取的蛋白用胰蛋白酶消化, 通过 MALDI-TOF-MS (Voyager STR, Applied Biosystems)用 337 nm 波长激光解吸和分析的反射模式进行分析。BCMP 11 的两个选定质量([M+H] = 1245.65 和 941.47)通过串联质谱法, 用配有纳米喷雾离子源的 QTOF-MS (Micromass UK Ltd.)进一步表征。在 MALDI 分析之前, 用 C18 Zip Tips<sup>TM</sup> (Millipore)将样品脱盐和浓缩。供串联 MS 用的样品用加入 C18 SPE 材料的纳米 LC 系统(LC Packings)纯化。

人工分析串联质谱(Biemann K, 通过串联质谱法和高能碰撞诱导

解离对肽测序, Methods Enzymol 1990; 193:455-79), 以测定部分位置的氨基酸序列以及所述肽片段的 N 末端和 C 末端其余部分的质量(表 5)。

发现得自蛋白 BCMP 11 的三个质谱(一个串联质谱, 表 5, 和一个 MALDI-质谱, 图 4a)匹配得自人肺癌文库的一个 EST (登记号 AI458391)的翻译, 该翻译确定了一个 166 个氨基酸的可读框(ORF)(图 4a)。用输入项 AI458391 对人 EST 数据库 (<http://www.ncbi.nlm.nih.gov/blast>)进行 Blast 搜索, 鉴定出几种重叠 EST, 提供了额外的 5'和 3' UTR 序列。得自 BCMP 11 的第二种肽串联质谱(表 5)在这些 EST 序列中鉴定出一个差异: 预测的密码子 146 编码或者谷氨酰胺或者精氨酸, 取决于翻译哪种 EST。串连质谱清楚地表明, 该精氨酸存在于所述分离蛋白中; 所述谷氨酰胺或者是测序错误或者是多态性。

表 5. 得自 BCMP 11 的串联质谱法分析的氨基酸序列信息

肽先驱离子 m/z	核心序列 <sup>a</sup>	N 末端质量 <sup>b</sup>	C 末端质量 <sup>c</sup>
1245.65	SPDGQYVPR	227.14	0.00
941.47	LYTYEPR	0.00	0.00

<sup>a</sup> “核心序列”是由所述肽的碎片质谱的解释阐明的肽的部分氨基酸序列。

<sup>b</sup> 所述肽的 N 末端质量是核心序列起点和所述肽 N 末端之间的质量。

<sup>c</sup> 所述 C 末端质量是核心序列末端和所述肽 C 末端之间的质量。

通过 PCR, 从 T-47D cDNA 扩增了一个全长克隆(图 4A)。

### 总 RNA 的制备和 cDNA 合成

用 Trizol 试剂(Life Technologies), 按照生产商的说明, 从培养的细胞和组织样品制备总 RNA, 并且以 1 $\mu$ g/ $\mu$ l 的浓度重悬于无 RNA 酶的水中。用 1-5 $\mu$ g 总 RNA 作为 cDNA 合成的模板, 使用寡聚 dT 引物和 Superscript II 反转录试剂盒(Life Technologies)。cDNA 经柱纯化(Qiagen), 以 10ng/ $\mu$ l 的浓度洗脱。

### BCMP 11 cDNA 的克隆

用以下引物：F, 5' GGCCAAGTCAGCTTCTTCTG 3'; R, 5' GTATTTGTCAATGTGCCAGAGG 3', 通过 PCR 从 T-47D cDNA 扩增预测的全长 BCMP 11 ORF。反应物含有 10ng cDNA 和 PCR 试剂 (Qiagen), 采用以下循环参数: 94°C 30 秒、60°C 30 秒 40 个循环。PCR 产物经柱纯化(Qiagen), 克隆到 T/A 载体(Invitrogen)中, 随后证实所述核苷酸序列(University of Oxford, Sequencing Facility, UK)。

预测的 BCMP 11 蛋白与 hAG-2 高度同源, hAG-2 是一种从 MCF-7 乳腺癌细胞系克隆的 cDNA 编码的一种新的人类蛋白(Thompson, D.A. 和 Weigel, R.J. 有爪蟾胶粘腺基因 XAG-2 的人类同系物 hAG-2 在乳腺癌细胞系中与雌激素受体共表达。 *Biochem. Biophys. Res. Commun.* **251**, 111-116 (1998)). hAG-2 是在蟾蜍发育期间在胶粘腺中表达的有爪蟾蛋白 XAG-2 (Aberger, F., Weidinger, G., Grunz, H. 和 Richter, K. 胚胎外胚层的前特化(anterior specification): 非洲蟾蜍属(*Xenopus*)胶粘腺特异性基因 XAG-2 的作用。 *Mech. Dev.* **72**, 115-130 (1998))的人类同系物。hAG-2 和 BCMP 11 蛋白有 71% 相同, N 末端的序列相似性相对低(图 4b)。BCMP 11 同 hAG-2 一样, 预测其是一种具有 N 末端信号序列的胞外蛋白(<http://psort.nibb.ac.jp>)。

### BCMP 11 mRNA 在人组织中的表达

我们采用实时定量 RT-PCR (Heid, C.A., Stevens, J., Livak, K.J. 和 Williams, P.M. 实时定量 PCR。 *Genome Res.* **6**, 986-994 (1996); Morrison, T.B., Weis, J.J. 和 Wittwer, C.T. 在扩增期间通过连续 SYBR Green I 测定量测定低拷贝转录物。 *Biotechniques* **24**, 954-958 (1998)), 以分析 BCMP 11 mRNA 在正常人组织和乳腺癌细胞系中的分布(图 5)。PCR 所用的引物如下: 有义引物, 5' CTGGAGGATTGTCAATACTC 3', 反义引物, 5' GCATAAGGTTTAGCATGATG 3'。在 ABI7700 序列测定系统(PE Biosystems)上, 分析含有如上所述制备的 10ng cDNA、SYBR

green 序列测定试剂(PE Biosystems)以及有义引物和反义引物的反应物。PCR 的条件是: 50°C 2 分钟 1 个循环, 95°C 10 分钟 1 个循环, 然后 95°C 15 秒、55°C 1 分钟 40 个循环。以 SYBR 绿色荧光的增加实时测量 PCR 产物的积累, 用 Sequence Detector 程序 v1.6.3 (PE Biosystems) 分析数据。用所述经扩增的 PCR 产物作为模板, 产生原始模板拷贝数与荧光和扩增循环相关的标准曲线, 用该曲线来计算每种样品中的 BCMP 11 的拷贝数。

在结肠中观测到 BCMP 11 的最高表达水平(2800 拷贝/ng cDNA), 在乳腺、小肠和气管中观测到较低的表达水平(295-435 拷贝/ng cDNA) (图 5)。在该项研究中所用的 BCMP 11 蛋白来源 T-47D 细胞中也检测到 BCMP 11 mRNA (1200 拷贝/ng cDNA), 与正常乳腺组织相比, 在该细胞系中的表达升高。在检验的其它细胞系或正常组织中检测到很少 BCMP 11 mRNA 或没有检测到 BCMP 11 mRNA。

BCMP 11 mRNA 和 hAG-2 mRNA 的组织分布非常相似。这两种基因都显示出在正常人组织中的限制表达模式, 结肠是两种基因的主要表达部位(图 5, Thompson, D.A. 和 Weigel, R.J. 有爪蟾蜍胶粘腺基因 XAG-2 的人类同系物 hAG-2 在乳腺癌细胞系中与雌激素受体共表达。 *Biochem. Biophys. Res. Commun.* **251**, 111-116 (1998))。表明 hAG-2 的表达与 ER 的表达同时发生, 当 MCF7 细胞用雌二醇处理时, 所述细胞中的 hAG-2 mRNA 水平增加(Thompson 和 Weigel, 参见上文)。同样, 在 ER 阳性 T-47D 细胞中检测到 BCMP 11 mRNA, 但在 ER 阴性细胞系 CAL51、BT20 和 MDA-MB-468 中没有观测到表达(图 5)。在 7 号染色体上相互相邻定位的两个高度同源的基因同时正调节(参见 Chromosomal localisation, 下文), 提示一种协同控制机制, 该机制可能不仅与 ER 状态相关, 而且与 ER 指导的治疗例如他莫昔芬相关。

### 染色体定位

用所述 BCMP 11 cDNA 序列(图 4a)对人类基因组数据库进行 Blast

搜索，识别出两个条目 - 分别与 BCMP 11 的 5'端和 3'端具有同一性的 AC004993 和 AC073333。已经将两个克隆作图至染色体 7p21。也可以将完整的 hAG-2 ORF 作图至条目 AC073333 中与 BCMP 11 相同的重叠群上，证明这些高度同源的基因在人类基因组中聚集在一起。

#### **实施例 4: 新蛋白 BCMP 84 的克隆**

##### **质谱法**

从 1D 凝胶切取的蛋白用胰蛋白酶消化，通过 MALDI-TOF-MS (Voyager STR, Applied Biosystems) 用 337 nm 波长激光解吸和分析的反射模式进行分析。BCMP 84 的一个选定质量( $[M+H] = 1667.73$ )通过串联质谱法，用配有纳米喷雾离子源的 QTOF-MS (Micromass UK Ltd.) 进一步表征。在 MALDI 分析之前，用 C18 Zip Tips™ (Millipore) 将样品脱盐和浓缩。供串联 MS 用的样品用加入 C18 SPE 材料的纳米 LC 系统(LC Packings)纯化。

人工分析串联质谱(Biemann K, 通过串联质谱法和高能碰撞诱导解离对肽测序, Methods Enzymol 1990; 193:455-79), 以测定部分位置的氨基酸序列以及所述肽片段的 N 末端和 C 末端其余部分的质量(表 6)。

表 6. 得自 BCMP 84 的串联质谱法分析的氨基酸序列信息

肽先驱离子 m/z	核心序列 <sup>a</sup>	N 末端质量 <sup>b</sup>	C 末端质量 <sup>c</sup>
1667.73	AEDAQEFSDVER	272.19	0.00

<sup>a</sup> “核心序列”是由所述肽的碎片质谱的解释阐明的肽的部分氨基酸序列。

<sup>b</sup> 所述肽的 N 末端质量是核心序列起点和所述肽 N 末端之间的质量。

<sup>c</sup> 所述 C 末端质量是核心序列末端和所述肽 C 末端之间的质量。

发现所述串联氨基酸序列和三个 MALDI 质谱匹配得自人结肠癌细胞系的一个 EST (登记号 AA315020) 的翻译(图 6)。鉴定出重叠 EST, 这建立了一个 104 个氨基酸的完整 ORF。

通过 PCR, 从 MDA-MB-468 cDNA 扩增了一个全长克隆。

### 总 RNA 的制备和 cDNA 合成

用 Trizol 试剂(Life Technologies), 按照生产商的说明, 从培养的细胞和组织样品制备总 RNA, 并且以 1 $\mu$ g/ $\mu$ l 的浓度重悬于无 RNA 酶的水中。用 1-5 $\mu$ g 总 RNA 作为 cDNA 合成的模板, 使用寡聚 dT 引物和 Superscript II 反转录试剂盒(Life Technologies)。cDNA 经柱纯化(Qiagen), 以 10ng/ $\mu$ l 的浓度洗脱。

### BCMP 84 cDNA 的克隆

用以下引物: F, 5' ATAGGACAACAGAACTCTCACC 3'; R, 5' GCTTCAACGGAAC TTTGCAGAG 3', 通过 PCR 从 MDA-MB-468 cDNA 扩增预测的全长 BCMP 84 ORF。反应物含有 10ng cDNA 和 PCR 试剂(Qiagen), 采用以下循环参数: 94 $^{\circ}$ C 30 秒、60 $^{\circ}$ C 30 秒 40 个循环。PCR 产物经柱纯化(Qiagen), 克隆到 T/A 载体(Invitrogen)中, 随后证实所述核苷酸序列(University of Oxford, Sequencing Facility, UK)。

预测的 BCMP 84 蛋白显示出与钙结合蛋白的 S100 家族(例如 S100A13, 登记号 Q99584, 与 BCMP 84 有 36% 的同一性和 67% 的同源性)和最近鉴定的 cDNA (AAG01893) 的相似性, AAG01893 在其长度的大部分长度范围内与 BCMP 84 相同, 已经被命名为 S100A14, 并注释为钙结合蛋白 S100 家族的一个新成员。然而, 没有证明 AAG01893/S100A14 结合钙, 并且该基因和 BCMP 84 缺乏在该蛋白家族成员之间保守的钙结合基序。BCMP 84 和 S100A14 之间的 2 个氨基酸的差异可能是多态性, 匹配我们在 BCMP 84 中发现的个体间变异。对该蛋白序列的分析揭示, 没有可能表明 BCMP 84 特定功能或细胞位置的蛋白基序。

### BCMP 84 mRNA 在人组织中的表达

我们采用实时定量 RT-PCR (Heid, C.A., Stevens, J., Livak, K.J.和 Williams, P.M. 实时定量PCR. *Genome Res.* 6, 986-994 (1996); Morrison, T.B., Weis, J.J.和 Wittwer, C.T. 在扩增期间通过连续 SYBR Green I 监测定量测定低拷贝转录物. *Biotechniques* 24, 954-958 (1998)), 以分析 BCMP 84 mRNA 在正常人组织和乳腺癌细胞系中的分布(图 7)。PCR 所用的引物如下: 有义引物, 5' TCTGTGCACTCTGTCTTGGA 3', 反义引物, 5' TAGCCAGCTCCTCTCTGTT 3'。在 ABI7700 序列测定系统(PE Biosystems)上, 分析含有如上所述制备的 10ng cDNA、SYBR green 序列测定试剂(PE Biosystems)以及有义引物和反义引物的反应物。PCR 的条件是: 50°C 2 分钟 1 个循环, 95°C 10 分钟 1 个循环, 然后 95°C 15 秒、55°C 1 分钟 40 个循环。以 SYBR 绿色荧光的增加实时测量 PCR 产物的积累, 用 Sequence Detector 程序 v1.6.3 (PE Biosystems) 分析数据。用所述经扩增的 PCR 产物作为模板, 产生原始模板拷贝数与荧光和扩增循环相关的标准曲线, 用该曲线来计算每种样品中的 BCMP 84 的拷贝数。

BCMP 84 mRNA 的分布限于少数组织, 在结肠、甲状腺和胸腺中的表达水平最高(260-930 拷贝/ng cDNA), 在包括乳腺在内的其它正常组织中仅检测到低水平的 BCMP 84 信使。在 BT-20 和 MDA-MB-468 细胞中检测到 BCMP 84 mRNA (分别为 300 和 1600 拷贝/ng cDNA), 但在 T-47D 或 CAL51 乳腺癌细胞系中没有检测到 BCMP 84 mRNA。

### 染色体定位

#### 辐射杂种作图

通过得自 3'非翻译区的以下引物对筛选 Genebridge 4 Radiation Hybrid panel (Research Genetics Inc.), 完成 BCMP 84 基因的染色体定位:

有义引物, 5' TCAGCTTCCTTCCCCAGGTC 3'

反义引物, 5' CCCAGCTCCATTATTCA 3'.

扩增 BCMP 84 序列的 PCR 条件是: 用 *Taq* DNA 聚合酶(Qiagen) 和每个反应 25ng DNA, 于 94°C 变性 30 秒, 然后退火和于 55°C 延伸 30 秒(40 个循环)。所述引物从所述阳性杂种细胞系 DNA 和人基因组 DNA 中扩增了预期的 215bp 片段, 但不能从仓鼠基因组 DNA (对照) 中扩增产物。

将辐射杂种作图数据提交给 Whitehead Institute/MIT Centre for Genome research STS 作图服务器 (<http://carbon.wi.mit.edu.8000/cgi-bin/contig/rhmapper.pl>) 以供分析。

辐射杂种作图将 BCMP 84 基因作图至染色体 1q21 上 S100 钙结合蛋白基因簇中的 STS 标记 AFM291xh1 和 AFM220xf8 之间。因而, 虽然 BCMP 84 与其它 S100 家族成员仅显示出有限的同源性, 并且缺乏明显的钙结合结构域, 但显然它与这一蛋白大家族相关。

## 实施例 5: 组织分布分析

### 材料和方法

#### 定量 mRNA 分析

用双链 DNA 染料 SYBR Green I, 采用 Morrison 等, *Biotechniques* (1998) 24:954-8 中所述的方法和试剂, 使用 ABI7700 (PE Biosystems, UK) 序列测定系统, 通过定量实时 PCR, 测定表 1 和表 2 中所述的 BCMP 7、17 和 23 (即推定的分泌蛋白 XAG (AF088867)、BST2 (Q10589) 和 NCAM2 (O15394) 在不同细胞类型中的 mRNA 水平。

概括而言, 通过用纯化(Qiagen, UK) PCR 产物作为模板产生标准曲线, 达到 mRNA 水平的定量。一式三份进行含有  $10^4$ 、 $10^5$ 、 $10^6$  和  $10^7$  模板分子的 PCR, 以测定扩增信号进入 log 线性范围的循环数(阈循环, Ct)。从而标准曲线确定了 Ct 和模板浓度之间的关系, 可以用来根据其相应的 Ct 值得出未知样品中的原始模板浓度。样品以一式两份进行测定, 每份使用 10ng 第一链 cDNA 作为原料。

通过寡聚 dT 引发的反转录(SuperscriptII – Life Technologies, UK), 从由多个人类来源混合的 RNA (Clontech, USA)衍生人组织 cDNA。从通过用 Trizol (Life Technologies, UK)提取的组织培养物样品中, 产生人细胞系 T47D、Cal51、MDA-MB468、PC3、PC3M 和 BT20 的总 RNA。

用来测定推定的分泌蛋白 XAG、BST2 和 NCAM2 mRNA 水平的引物序列是:

#### **XAG**

正向            agataccacagtcaaacctg

反向            gcactcatccaagtgatgaa

#### **BST2**

正向            gatggagtgtcgcaatgtcacc

反向            agacgcgtcctgaagcttatgg

#### **NCAM2**

正向            gatcatagagctgtcgagacc

反向            tgtagtctccttgccctttcc

### **结果**

结果示于图 8-10 中。在所述图中, 注释“x10”是指该样品的数值是所指示数值的 10 倍。

## 序 列 表

<110> 牛津格莱克科学(英国)有限公司(Oxford Glycosciences (UK) Ltd)  
 Boyd, Robert Simon  
 Stamps, Alasdair Craig  
 Terrett, Jonathan Alexander

<120> 人乳腺细胞膜

<130> NJL/P30142WO

<140> PCT/GB01/00748

<141> 2001-02-21

<150> GB 0004576.5

<151> 2000-02-25

<150> GB 0031341.1

<151> 2000-12-21

<160> 110

<170> PatentIn version 3.0

<210> 1

<211> 25

<212> PRT

<213> 人(Homo sapiens)

<400> 1

Gly Asp Ala Glu Lys Pro Glu Glu Glu Leu Glu Glu Asp Asp Asp Glu

1                    5                    10                    15

Glu Leu Asp Glu Thr Leu Ser Glu Arg

20                    25

<210> 2

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 2

Glu Leu Phe Pro Ile Gln Met Glu Gly Val Lys

1                    5                    10

<210> 3

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 3

Phe Val Asn Trp Gln Val Asp Gly Glu Tyr Arg

1                    5                    10

<210> 4

<211> 8

<212> PRT

<213> 人(Homo sapiens)

<400> 4

Gln Gly His Leu Ser Leu Gln Arg

1                    5

<210> 5

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 5

Tyr Gln Glu Thr Phe Asn Val Ile Glu Arg

1                    5                    10

<210> 6

<211> 16

<212> PRT

<213> 人(Homo sapiens)

<400> 6

Ile Asn Pro Asp Gly Ser Gln Ser Val Val Glu Val Pro Tyr Ala Arg

1                    5                    10                    15

<210> 7

<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 7  
Gln Gln Gln Leu Leu Asn Glu Glu Asn Leu Arg  
1                  5                  10

<210> 8  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 8  
Asn Val Asp Leu Ser Thr Phe Tyr Gln Asn Arg  
1                  5                  10

<210> 9  
<211> 10  
<212> PRT  
<213> 人(Homo sapiens)

<400> 9  
Tyr Asp Asp Ala Ile Gln Leu Tyr Asp Arg  
1                  5                  10

<210> 10  
<211> 8  
<212> PRT  
<213> 人(Homo sapiens)

<400> 10  
Val Gly Asp Val Val Glu Phe Arg  
1                  5

<210> 11  
<211> 9  
<212> PRT  
<213> 人(Homo sapiens)

<400> 11

Glu Asp Leu Glu Glu Leu Gln Ala Arg

1                   5

<210> 12

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 12

Leu Glu Asn Gly Glu Ile Glu Thr Ile Ala Arg

1                   5                   10

<210> 13

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 13

Ser Gly Ala Gly Val Pro Ala Val Ile Leu Arg

1                   5                   10

<210> 14

<211> 13

<212> PRT

<213> 人(Homo sapiens)

<400> 14

Ser Asn Pro Glu Asp Gln Ile Leu Tyr Gln Thr Glu Arg

1                   5                   10

<210> 15

<211> 14

<212> PRT

<213> 人(Homo sapiens)

<400> 15

Val Gly Ala Glu Asn Val Ala Ile Val Glu Pro Ser Glu Arg

1                   5                   10

<210> 16

<211> 12

<212> PRT

<213> 人(Homo sapiens)

<400> 16

Asn Gly Asn Gln Ala Phe Asn Glu Asp Asn Leu Lys  
1                    5                    10

<210> 17

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 17

Phe Ile Leu Met Asp Cys Met Glu Gly Arg  
1                    5                    10

<210> 18

<211> 12

<212> PRT

<213> 人(Homo sapiens)

<400> 18

Phe Asp Gly Ile Leu Thr Glu Gly Glu Gly Pro Arg  
1                    5                    10

<210> 19

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 19

Leu Gln Asp Ala Ser Ala Glu Val Glu Arg  
1                    5                    10

<210> 20

<211> 12

<212> PRT

<213> 人(Homo sapiens)

<400> 20

Leu Ser Cys Ala Tyr Ser Gly Phe Ser Ser Pro Arg  
1                    5                    10

---

<210> 21  
<211> 8  
<212> PRT  
<213> 人(Homo sapiens)  
  
<400> 21  
Phe Asp Gln Gly Asp Thr Thr Arg  
1                   5

<210> 22  
<211> 8  
<212> PRT  
<213> 人(Homo sapiens)

<400> 22  
Ile Thr Ala Ser Tyr Glu Asp Arg  
1                   5

<210> 23  
<211> 8  
<212> PRT  
<213> 人(Homo sapiens)

<400> 23  
Ile Pro Glu Asn Asn Pro Val Lys  
1                   5

<210> 24  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 24  
Gln Leu Val Glu Gln Val Glu Gln Ile Gln Lys  
1                   5                   10

<210> 25  
<211> 10  
<212> PRT  
<213> 人(Homo sapiens)

<400> 25

---

Phe Ser Leu Phe Ala Gly Gly Met Leu Arg  
1                    5                    10

<210> 26

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 26

Ile Pro Asp Gln Leu Val Ile Leu Asp Met Lys  
1                    5                    10

<210> 27

<211> 9

<212> PRT

<213> 人(Homo sapiens)

<400> 27

Leu Lys Pro Leu Glu Val Glu Leu Arg  
1                    5

<210> 28

<211> 7

<212> PRT

<213> 人(Homo sapiens)

<400> 28

His Asp Trp Ile Ile Asn Arg  
1                    5

<210> 29

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 29

Ala Tyr Glu Tyr Val Glu Cys Pro Ile Arg  
1                    5                    10

<210> 30

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 30

Asp Tyr Gln Phe Thr Ile Leu Asp Val Arg

1                    5                    10

<210> 31

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 31

Ser Trp Met Gly Tyr Glu Leu Pro Phe Asp Arg

1                    5                    10

<210> 32

<211> 12

<212> PRT

<213> 人(Homo sapiens)

<400> 32

Thr Glu Asp Leu Glu Ala Thr Ser Glu His Phe Lys

1                    5                    10

<210> 33

<211> 9

<212> PRT

<213> 人(Homo sapiens)

<400> 33

Glu Val Val Ser Pro Gln Glu Phe Lys

1                    5

<210> 34

<211> 13

<212> PRT

<213> 人(Homo sapiens)

<400> 34

Asn Ile Ile Asn Ser Asp Gly Gly Pro Tyr Val Cys Arg

1                    5                    10

<210> 35

<211> 12

<212> PRT

<213> 人(Homo sapiens)

<400> 35

Ala Val Asp Gly Phe Thr Phe Thr Glu Gly Asp Lys

1                    5                    10

<210> 36

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 36

Ile Glu Glu Ala Cys Glu Ile Tyr Ala Arg

1                    5                    10

<210> 37

<211> 12

<212> PRT

<213> 人(Homo sapiens)

<400> 37

Ser Pro Gln Gln Val Leu Pro Ser Pro Asp Gly Arg

1                    5                    10

<210> 38

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 38

Glu Leu Ala Pro Leu Phe Glu Glu Leu Arg

1                    5                    10

<210> 39

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 39

---

Ser Thr Ile Gly Val Glu Phe Ala Thr Arg  
1                    5                    10

<210> 40

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 40

Asp Glu Phe Leu Ile Gln Ala Ser Pro Arg  
1                    5                    10

<210> 41

<211> 8

<212> PRT

<213> 人(Homo sapiens)

<400> 41

Asp Asn Glu Gly Ser Trp Phe Arg  
1                    5

<210> 42

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 42

Leu Gln Ile Trp Asp Thr Ala Gly Gln Glu Arg  
1                    5                    10

<210> 43

<211> 8

<212> PRT

<213> 人(Homo sapiens)

<400> 43

Thr Ile Thr Ser Ser Tyr Tyr Arg  
1                    5

<210> 44

<211> 8

<212> PRT

<213> 人(Homo sapiens)

<400> 44

Gln Trp Leu Gln Glu Ile Asp Arg

1                   5

<210> 45

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 45

Thr Gln Ile Asp His Tyr Val Gly Ile Ala Arg

1                   5                   10

<210> 46

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 46

Gln Ile Ala Val Glu Ala Gln Glu Ile Leu Arg

1                   5                   10

<210> 47

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 47

Phe Leu Thr Gln Pro Gln Val Val Ala Arg

1                   5                   10

<210> 48

<211> 13

<212> PRT

<213> 人(Homo sapiens)

<400> 48

Val Ala Glu Asp Glu Ala Glu Ala Ala Ala Ala Lys

1                   5                   10

<210> 49  
<211> 12  
<212> PRT  
<213> 人(Homo sapiens)

<400> 49  
Ile Gln Thr Gln Pro Gly Tyr Ala Asn Thr Leu Arg  
1                   5                   10

<210> 50  
<211> 14  
<212> PRT  
<213> 人(Homo sapiens)

<400> 50  
Val Val Val Thr Val Glu Gln Thr Glu Glu Glu Leu Glu Arg  
1                   5                   10

<210> 51  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 51  
Gln Gly Tyr Val Leu Ser Ser Ile Glu Gly Arg  
1                   5                   10

<210> 52  
<211> 8  
<212> PRT  
<213> 人(Homo sapiens)

<400> 52  
Thr Gln Leu Gln Leu Asp Glu Arg  
1                   5

<210> 53  
<211> 13  
<212> PRT  
<213> 人(Homo sapiens)

<400> 53  
Glu Val Ser Phe Gln Ser Thr Gly Glu Ser Glu Trp Lys  
1                   5                   10

<210> 54  
<211> 7  
<212> PRT  
<213> 人(Homo sapiens)

<400> 54  
Leu Ser Glu Leu Gln Leu Arg  
1                   5

<210> 55  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 55  
Gln Leu Val Glu Gln Val Glu Gln Ile Gln Lys  
1                   5                   10

<210> 56  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 56  
Asp Asn Val Asp Asp Pro Thr Gly Asn Phe Arg  
1                   5                   10

<210> 57  
<211> 10  
<212> PRT  
<213> 人(Homo sapiens)

<400> 57  
Val Glu Glu Val Gly Pro Tyr Thr Tyr Arg  
1                   5                   10

<210> 58  
<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 58

Ala Thr Val Leu Glu Ser Glu Gly Thr Arg

1                   5                   10

<210> 59

<211> 15

<212> PRT

<213> 人(Homo sapiens)

<400> 59

Asp Val Gln Gly Thr Asp Ala Ser Leu Asp Glu Glu Leu Asp Arg

1                   5                   10                   15

<210> 60

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 60

Phe Asn Ala Ser Gln Leu Ile Thr Gln Arg

1                   5                   10

<210> 61

<211> 9

<212> PRT

<213> 人(Homo sapiens)

<400> 61

His Ser Glu Ile Gln Gln Leu Glu Arg

1                   5

<210> 62

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 62

Asn Ile Leu Ser Ser Ala Asp Tyr Val Glu Arg

1                   5                   10

<210> 63  
<211> 10  
<212> PRT  
<213> 人(Homo sapiens)

<400> 63  
Asp Leu Gln Gln Tyr Gln Ser Gln Ala Lys  
1                   5                   10

<210> 64  
<211> 14  
<212> PRT  
<213> 人(Homo sapiens)

<400> 64  
Ala Glu Glu Trp Gly Val Gln Tyr Val Glu Thr Ser Ala Lys  
1                   5                   10

<210> 65  
<211> 8  
<212> PRT  
<213> 人(Homo sapiens)

<400> 65  
Gln Val Pro Val Glu Glu Ala Arg  
1                   5

<210> 66  
<211> 7  
<212> PRT  
<213> 人(Homo sapiens)

<400> 66  
Val Phe Phe Asp Leu Met Arg  
1                   5

<210> 67  
<211> 13  
<212> PRT  
<213> 人(Homo sapiens)

<400> 67  
Asp Leu Glu Cys Val Thr Asn Leu Gln Glu Val Ala Arg  
1                   5                   10

<210> 68  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 68  
Trp Pro Tyr Ala Gly Thr Ala Glu Ala Ile Lys  
1                   5                   10

<210> 69  
<211> 10  
<212> PRT  
<213> 人(Homo sapiens)

<400> 69  
Glu Val Val Glu Ala His Val Asp Gln Lys  
1                   5                   10

<210> 70  
<211> 10  
<212> PRT  
<213> 人(Homo sapiens)

<400> 70  
Tyr Asp Pro Thr Ile Glu Asp Ser Tyr Arg  
1                   5                   10

<210> 71  
<211> 12  
<212> PRT  
<213> 人(Homo sapiens)

<400> 71  
His Leu Ile Pro Ala Ala Asn Thr Gly Glu Ser Lys  
1                   5                   10

<210> 72  
<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 72

Tyr Leu Ala Glu Phe Ala Thr Gly Asn Asp Arg

1                   5                   10

<210> 73

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 73

Glu Gln Asn Ser Pro Ile Tyr Ile Ser Arg

1                   5                   10

<210> 74

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 74

Val Leu Gln Ser Glu Phe Cys Asn Ala Val Arg

1                   5                   10

<210> 75

<211> 10

<212> PRT

<213> 人(Homo sapiens)

<400> 75

Tyr Leu Glu Cys Ser Ala Leu Thr Gln Arg

1                   5                   10

<210> 76

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 76

Ile Met Phe Val Asp Pro Ser Leu Thr Val Arg

1                   5                   10

<210> 77  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 77  
His Leu Ser Pro Asp Gly Gln Tyr Val Pro Arg  
1                   5                   10

<210> 78  
<211> 11  
<212> PRT  
<213> 人(Homo sapiens)

<400> 78  
Asn Leu Ser Pro Asp Gly Gln Tyr Val Pro Arg  
1                   5                   10

<210> 79  
<211> 12  
<212> PRT  
<213> 人(Homo sapiens)

<400> 79  
Leu Gln Met Glu Gln Gln Gln Leu Gln Gln Arg  
1                   5                   10

<210> 80  
<211> 15  
<212> PRT  
<213> 人(Homo sapiens)

<400> 80  
Ser Ala Asn Ala Glu Asp Ala Gln Glu Phe Ser Asp Val Glu Arg  
1                   5                   10                   15

<210> 81  
<211> 10  
<212> PRT  
<213> 人(Homo sapiens)

<400> 81

Met Glu Gln Gln Gln Gln Leu Gln Gln Arg

1                   5                   10

<210> 82

<211> 11

<212> PRT

<213> 人(Homo sapiens)

<400> 82

Leu Trp Gly Leu Thr Glu Met Phe Pro Glu Arg

1                   5                   10

<210> 83

<211> 25

<212> PRT

<213> 人(Homo sapiens)

<400> 83

Gly Asp Ala Glu Lys Pro Glu Glu Glu Leu Glu Glu Asp Asp Asp Glu

1                   5                   10                   15

Glu Leu Asp Glu Thr Leu Ser Glu Arg

20                   25

<210> 84

<211> 9

<212> PRT

<213> 人(Homo sapiens)

<400> 84

Ser Pro Asp Gly Gln Tyr Val Pro Arg

1                   5

<210> 85

<211> 7

<212> PRT

<213> 人(Homo sapiens)

<400> 85

Leu Tyr Thr Tyr Glu Pro Arg

1                   5

<210> 86  
<211> 20  
<212> DNA  
<213> 人工序列

<220>

<221>

<223> 引物

<400> 86

ggccaagtca gcttcttctg

20

<210> 87  
<211> 22  
<212> DNA  
<213> 人工序列

<220>

<221>

<223> 引物

<400> 87

gtatttgca atgtgccaga gg

22

<210> 88  
<211> 20  
<212> DNA  
<213> 人工序列

<220>

<221>

<223> 引物

<400> 88

ctggaggatt gtcaatactc

20

<210> 89  
<211> 20  
<212> DNA  
<213> 人工序列

<220>

<221>

<223> 引物

<400> 89

gcataagggtt tagcatgatg

20

<210> 90

<211> 12

<212> PRT

<213> 人(Homo sapiens)

<400> 90

Ala Glu Asp Ala Gln Glu Phe Ser Asp Val Glu Arg

1

5

10

<210> 91

<211> 22

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 91

ataggacaac agaactctca cc

22

<210> 92

<211> 22

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 92

gcttcaacgg aactttgcag ag

22

<210> 93

<211> 20

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 93

tctgtgcact ctgtcttgga

20

<210> 94

<211> 19

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 94

tagccagctc ctctctggt

19

<210> 95

<211> 20

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 95

tcagcttctt tccccaggtc

20

<210> 96

<211> 17

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 96

cccagctcca ttattca	17
<210> 97	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<221>	
<223> 引物	
<400> 97	
agataccaca gtcaaacctg	20
<210> 98	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<221>	
<223> 引物	
<400> 98	
gcactcatcc aagtgatgaa	20
<210> 99	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<221>	
<223> 引物	
<400> 99	
gatggagtgt cgcaatgtca cc	22
<210> 100	
<211> 22	
<212> DNA	
<213> 人工序列	

<220>

<221>

<223> 引物

<400> 100

agacgcgtcc tgaagcttat gg

22

<210> 101

<211> 22

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 101

gatcatagag ctgtcgcaga cc

22

<210> 102

<211> 22

<212> DNA

<213> 人工序列

<220>

<221>

<223> 引物

<400> 102

tgtagtctcc ttgtcccttt cc

22

<210> 103

<211> 142

<212> PRT

<213> 人(Homo sapiens)

<400> 103

Met Ala Ala Ala Val Ala Ala Ala Gly Ala Gly Glu Pro Gln Ser Pro

1                    5                    10                    15

Asp Glu Leu Leu Pro Lys Gly Asp Ala Glu Lys Pro Glu Glu Glu Leu

20

25

30



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

<210> 106

<211> 614

<212> DNA

<213> 人(Homo sapiens)

<400> 106

ggccaagtca gcttcttctg agagagtctc tagaagacat gatgctacac tcagctttgg 60  
 gtctctgcoct cttactcgtc acagtttctt coaaccttgc cattgcaata aaaaaggaaa 120  
 agaggcoctcc tcagacactc tcaagaggat ggggagatga catcacttgg gtacaaaactt 180  
 atgaagaagg tctcttttat gctcaaaaaa gtaagaagcc attaatggtt attcatcacc 240  
 tggaggattg tcaatactct caagcactaa agaaagtatt tgcccanaat gaagaaatac 300  
 aagaaatggc tcagaataag ttcacatgc taaaccttat gcatgaaacc actgataaga 360  
 atttatcacc tgatgggcaa tatgtgcoctaat gaatcatggt ttagaccct tctttaacag 420  
 ttagagetga catagctgga agatactcta acagattgta cacatatgag cctcgggatt 480  
 taccctatt gatagaaaac atgaagaaag cattaagact tattcagtca gagctataag 540  
 agatgataga aaaaagcctt cacttcaaag aagtcaaatt tcatgaagaa aacctctggc 600  
 acattgacaa atac 614

<210> 107

<211> 166

<212> PRT

<213> 人(Homo sapiens)

<400> 107

Met Met Leu His Ser Ala Leu Gly Leu Cys Leu Leu Leu Val Thr Val

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

<210> 108

<211> 175

<212> PRT

<213> 人(Homo sapiens)

<400> 108

Met Glu Lys Ile Pro Val Ser Ala Phe Leu Leu Leu Val Ala Leu Ser  
 1                    5                    10                    15  
 Tyr Thr Leu Ala Arg Asp Thr Thr Val Lys Pro Gly Ala Lys Lys Asp  
                   20                    25                    30  
 Thr Lys Asp Ser Arg Pro Lys Leu Pro Gln Thr Leu Ser Arg Gly Trp  
                   35                    40                    45  
 Gly Asp Gln Leu Ile Trp Thr Gln Thr Tyr Glu Glu Ala Leu Tyr Lys  
                   50                    55                    60  
 Ser Lys Thr Ser Asn Lys Pro Leu Met Ile Ile His His Leu Asp Glu  
 65                    70                    75                    80  
 Cys Pro His Ser Gln Ala Leu Lys Lys Val Phe Ala Glu Asn Lys Glu  
                   85                    90                    95  
 Ile Gln Lys Leu Ala Glu Gln Phe Val Leu Leu Asn Leu Val Tyr Glu  
                   100                    105                    110

Thr Thr Asp Lys His Leu Ser Pro Asp Gly Gln Tyr Val Pro Arg Ile  
 115 120 125  
 Met Phe Val Asp Pro Ser Leu Thr Val Arg Ala Asp Ile Thr Gly Arg  
 130 135 140  
 Tyr Ser Asn Arg Leu Tyr Ala Tyr Glu Pro Ala Asp Thr Ala Leu Leu  
 145 150 155 160  
 Leu Asp Asn Met Lys Lys Ala Leu Lys Leu Leu Lys Thr Glu Leu  
 165 170 175

<210> 109  
 <211> 491  
 <212> DNA  
 <213> 人(Homo sapiens)

<400> 109  
 ataggacaac agaactctca ccaaggacc agacacagtg agcaccatgg gacagtgtcg 60  
 gtcagccaac gcagaggatg ctcaggaatt cagtgatgtg gagagggcca ttgagaccct 120  
 catcaagaac tttcaccagt actccgtgga ggggtgggaag gagacgctga ccccttctga 180  
 gctacgggac ctggtcaccc agcagctgcc ccatctcatg ccgagcaact gtggcctgga 240  
 agagaaaatt gccaacctgg gcagctgcaa taactctaaa ctggagtctg ggagtttctg 300  
 ggagctgatt ggagaagcgg ccaagagtgt gaagctggag aggcctgtcc gggggcactg 360  
 agaactccct ctggaattct tgggggtgt tggggagaga ctgtgggcct ggagataaaa 420  
 cttgtctcct ctaccaccac cctgtaccct agcctgcacc tgtctctcct tctgcaaagt 480  
 tccgttgaag c 491

<210> 110  
 <211> 104  
 <212> PRT  
 <213> 人(Homo sapiens)

<400> 110  
 Met Gly Gln Cys Arg Ser Ala Asn Ala Glu Asp Ala Gln Glu Phe Ser  
 1 5 10 15  
 Asp Val Glu Arg Ala Ile Glu Thr Leu Ile Lys Asn Phe His Gln Tyr  
 20 25 30  
 Ser Val Glu Gly Gly Lys Glu Thr Leu Thr Pro Ser Glu Leu Arg Asp  
 35 40 45  
 Leu Val Thr Gln Gln Leu Pro His Leu Met Pro Ser Asn Cys Gly Leu  
 50 55 60  
 Glu Glu Lys Ile Ala Asn Leu Gly Ser Cys Asn Asn Ser Lys Leu Glu  
 65 70 75 80  
 Phe Gly Ser Phe Trp Glu Leu Ile Gly Glu Ala Ala Lys Ser Val Lys

85  
Leu Glu Arg Pro Val Arg Gly His  
100

90

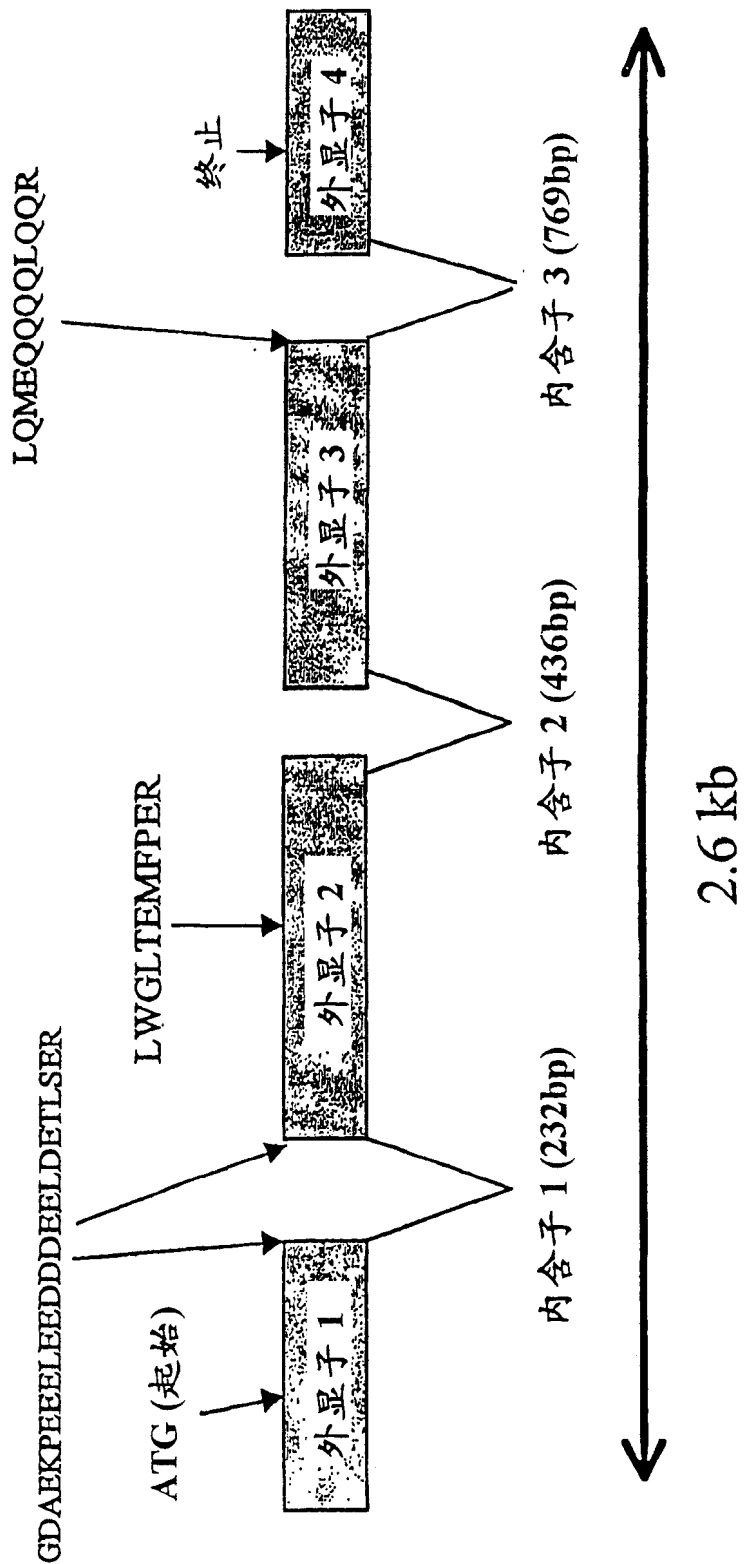
95

## 图 1

MAAAVAAAGA GEPQSPDELL PKGDAEKPEE ELEEDDDEEL DETLSERLWG  
LTEMFPERVR SAAGATFDLS LFVAQKMYRF SRAALWIGTT SFMILVLPVV  
FETEKLOMEQ QQQLQQRQIL LGPNTGLSGG MPGALPSLPG KI#

ATGGCTGCCG CCGTCGCTGC TGCCGGTGCA GGGGAACCCC AGTCCCCGGA  
CGAATTGCTC CCGAAAGGCG ACGCGGAGAA GCCTGAGGAG GAGCTGGAGG  
AGGACGACGA TGAGGAGCTA GATGAGACCC TGTCGGAGAG ACTATGGGGC  
CTGACGGAGA TGTTCCGGA GAGGGTCCGG TCCGCGGCCG GAGCCACTTT  
TGATCTTTCC CTCTTTGTGG CTCAGAAAAT GTACAGGTTT TCCAGGGCAG  
CCTTGTGGAT TGGGACCACT TCCTTTATGA TCCTGGTTCT TCCCGTTGTC  
TTTGAGACGG AGAAGTTGCA AATGGAGCAA CAGCAGCAAC TGCAGCAGCG  
GCAGATACTT CTAGGACCTA ACACAGGGCT CTCAGGAGGA ATGCCAGGGG  
CTCTACCCTC ACTTCCTGGA AAGATCTAG

图 2



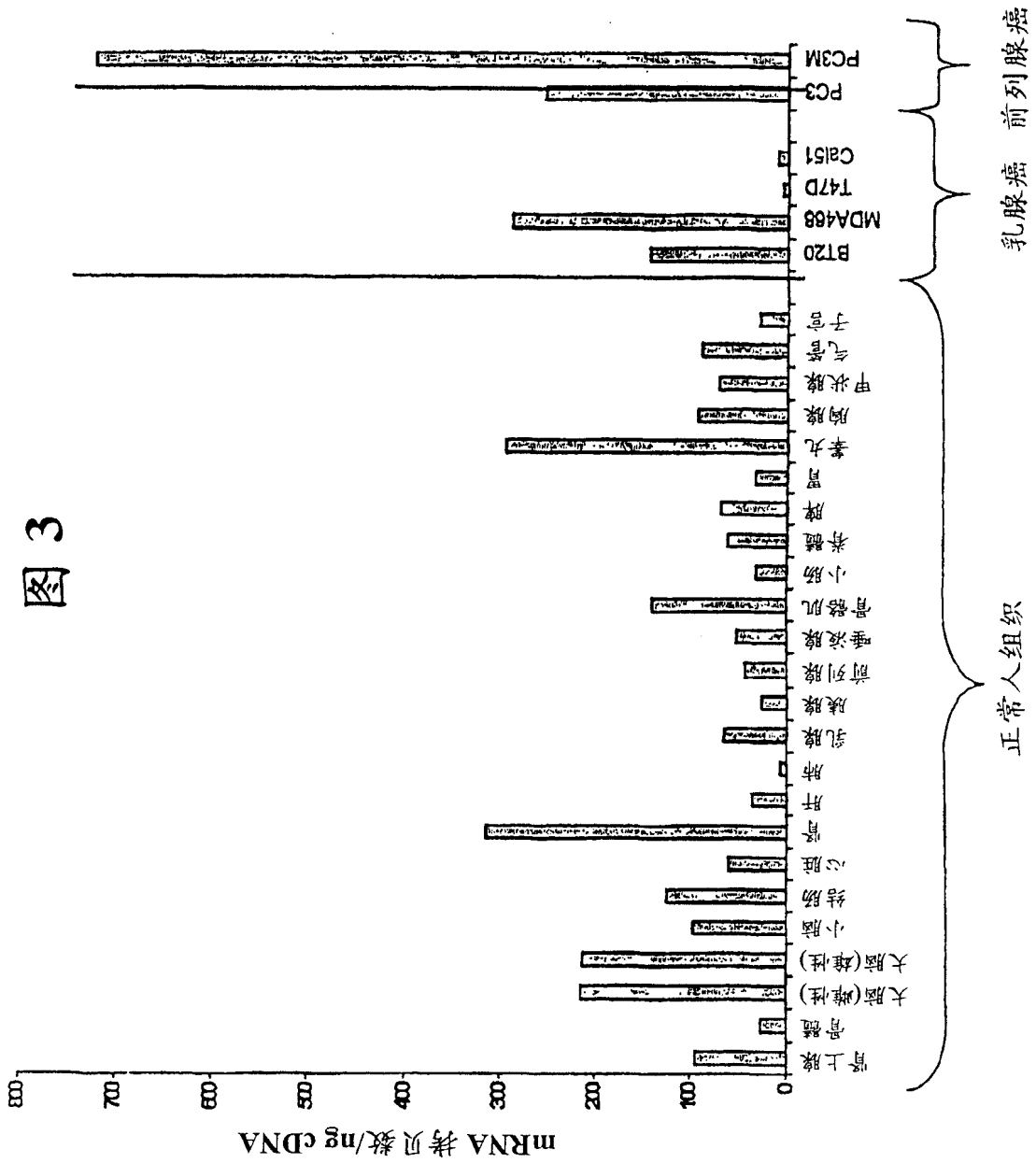
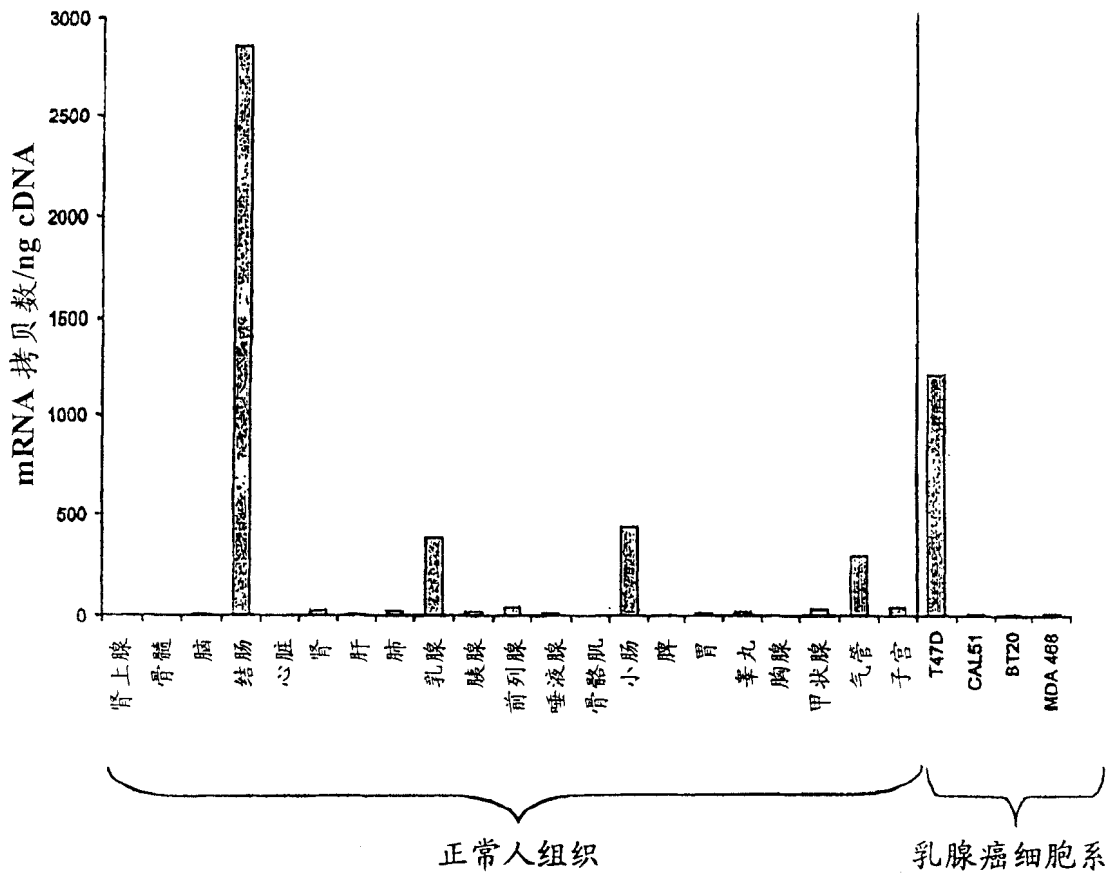


图 3



图 5



## 图 6

ATAGGACAACAGAACTCTACCAAAGGACCAGACACAGTGAGCACCATGGGACAGTGTCTG	60
M G Q C R	5
GTCAGCCAACGCAGAGGATGCTCAGGAATTCAGTGATGTGGAGAGGGCCATTGAGACCCT	120
S A N A E D A Q E F S D V E R <u>A I E T L</u>	25
CATCAAGAACTTTCACCAGTACTCCGTGGAGGGTGGGAAGGAGACGCTGACCCCTTCTGA	180
<u>I K N F H Q Y S V E G G K E T L T P S E</u>	45
GCTACGGGACCTGGTCACCCAGCAGCTGCCCCATCTCATGCCGAGCAACTGTGGCCTGGA	240
<u>L R</u> D L V T Q Q L P H L M P S N C G L E	65
AGAGAAAATTGCCAACCTGGGCAGCTGCAATAACTCTAAACTGGAGTTCGGGAGTTTCTG	300
E K I A N L G S C N N S K L E F G S F W	85
GGAGCTGATYGGAGAAGCGGCCAAGAGTGTGAAGCTGGAGAGGCCTGTCCGGGGGCACTG	360
E L I G E A A K S V K L E R P V R G H *	104
AGAACTCCCTCTGGAATTCCTGGGGGGTGTGGGGAGAGACTGTGGGCCTGGAGATAAAA	420
CTTGTCTCCTCTACCACCACCCTGTACCCTAGCCTGCACCTGTCTCATCTCTGCAAAGT	480
TCCGTTGAAGC	491

图 7

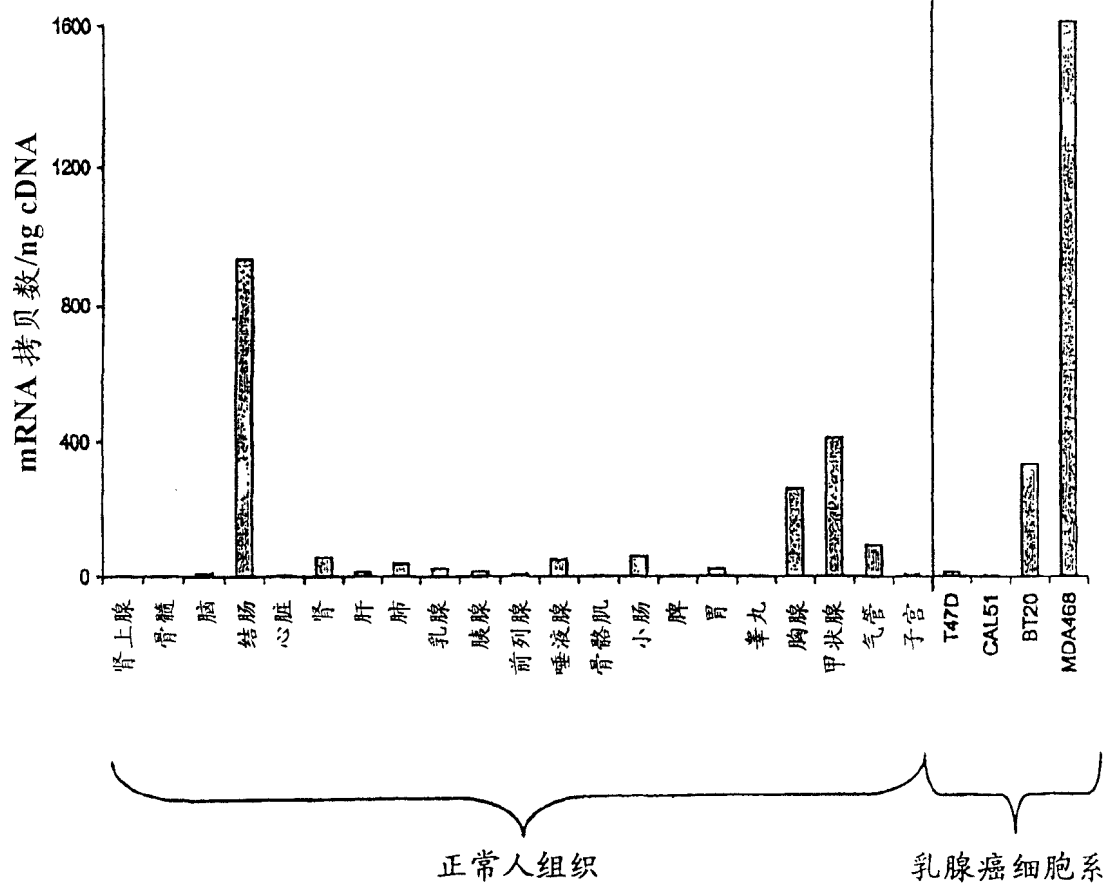
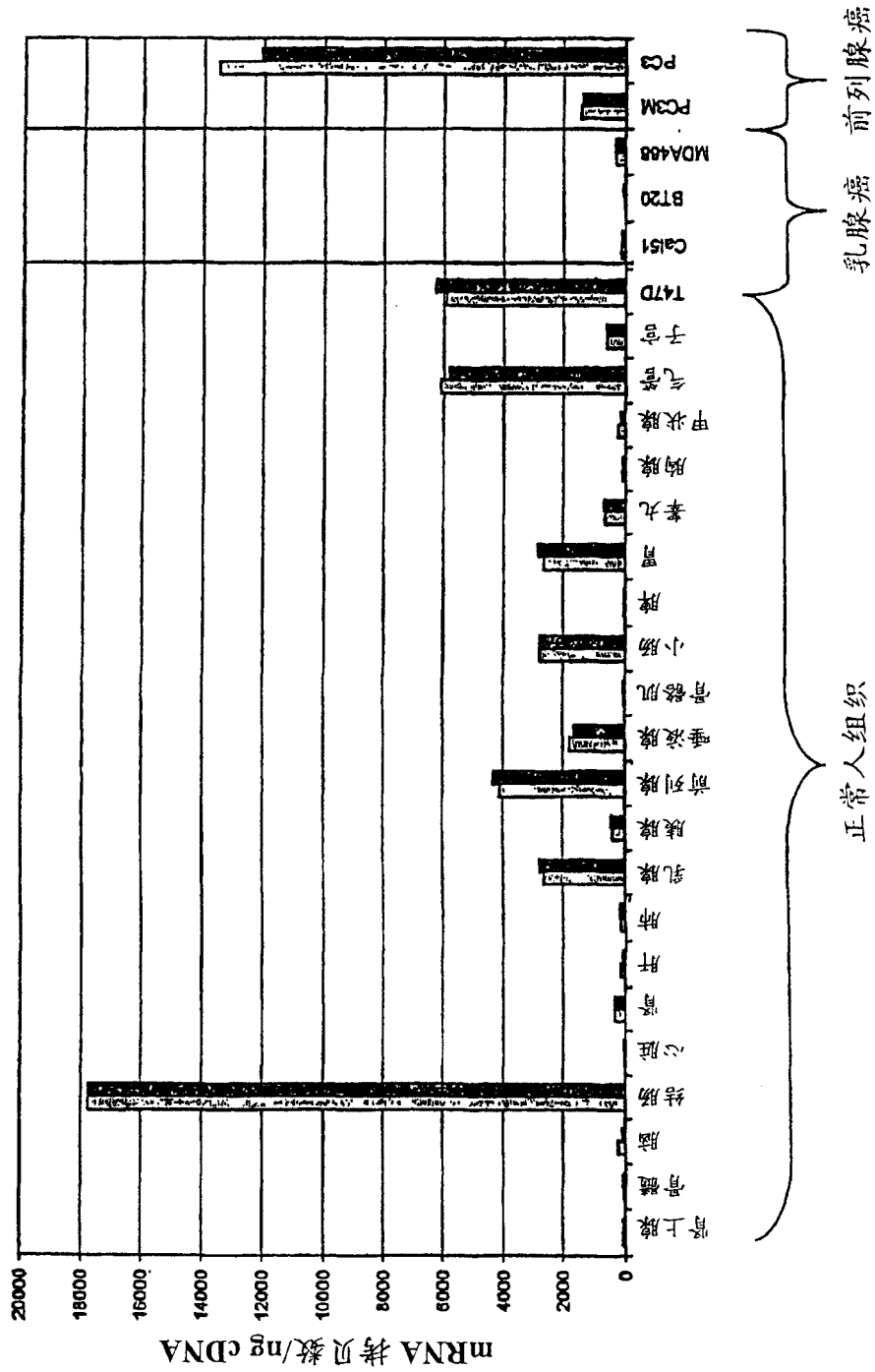


图 8



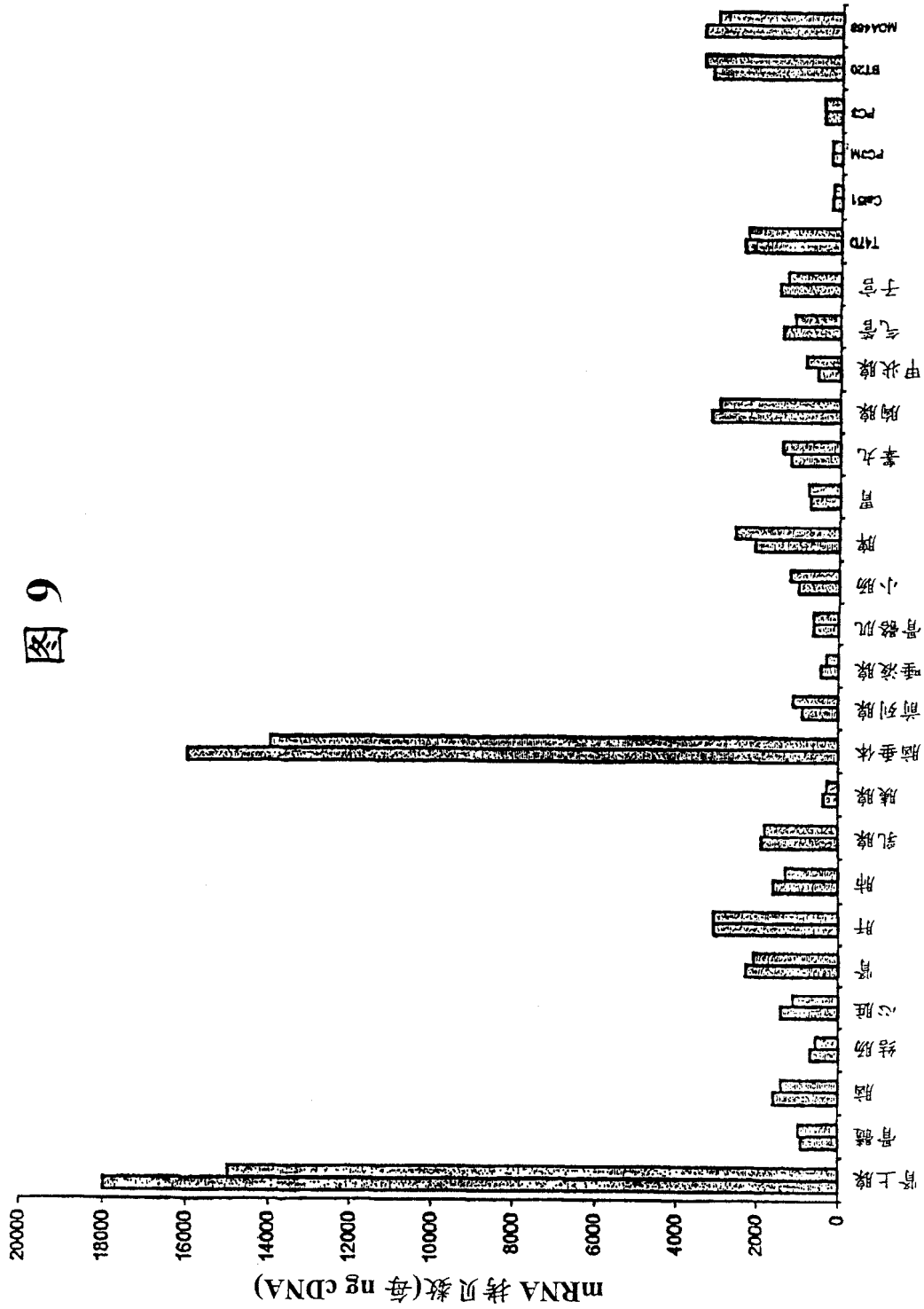
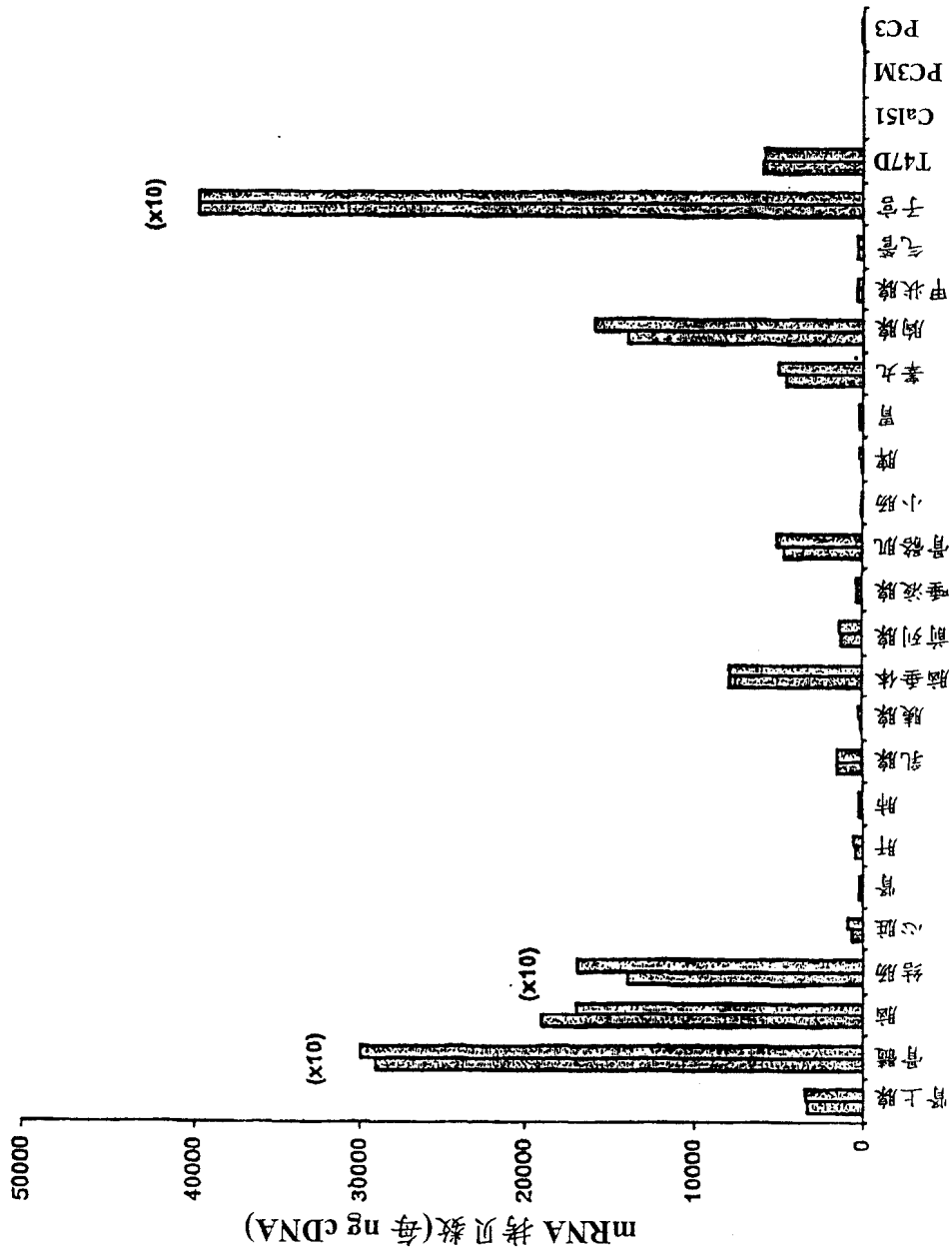


图9

图 10



专利名称(译)	蛋白质		
公开(公告)号	<a href="#">CN100384875C</a>	公开(公告)日	2008-04-30
申请号	CN01808488.5	申请日	2001-02-21
[标]申请(专利权)人(译)	牛津格莱克科学(英国)有限公司		
申请(专利权)人(译)	牛津格莱克科学(英国)有限公司		
当前申请(专利权)人(译)	牛津格莱克科学(英国)有限公司		
[标]发明人	JA特雷特		
发明人	R·S·博伊德 A·C·斯塔姆普斯 J·A·特雷特		
IPC分类号	C07K16/18 A61K38/17 A61K39/00 A61K39/395 G01N33/50 G01N33/53 A61K48/00 C12Q1/68 A61P35/00 A61K31/7088 A61K31/7105 A61K38/00 A61K47/42 A61K49/00 A61K51/10 C07K14/47 C07K16/32 C12M1/00 C12N1/21 C12N15/09 C12N15/12 G01N33/15 G01N33/566 G01N33/574 G01N33/577		
CPC分类号	C07K2319/00 G01N33/57438 G01N33/57415 A61K47/42 G01N33/57419 A61K9/0019 A61K2039/53 G01N2500/00 A61K48/00 A61K2039/505 G01N33/57407 A61K51/1051 A61K39/00 G01N2800/52 G01N33/57434 C07K14/47 A61K49/0002		
审查员(译)	宋智刚		
优先权	2000004576 2000-02-25 GB 2000031341 2000-12-21 GB		
其他公开文献	CN1426307A		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明提供了用于乳腺癌筛选、诊断和预后、监测乳腺癌治疗功效以及药物开发的方法和组合物。

MAAAVAAAGA GEPOSPDELL PKGDAEKPEE ELEEDDDEEL DETLSERLWG  
LTEMPPERVR SAAGATFDLS LFVAQKMYRF SRAALWIGTT SFMILVLPVV  
 FETEKLOMEQ QQQLQQRQIL LGPNTGLSGG MPGALPSLPG KI#

ATGGCTGCCG CCGTCGCTGC TGCCGGTGCA GGGGAACCCC AGTCCCCGGA  
 CGAATTGCTC CCGAAAGGCG ACGCGGAGAA GCCTGAGGAG GAGCTGGAGG  
 AGGACGACGA TGAGGAGCTA GATGAGACCC TGTCGGAGAG ACTATGGGGC  
CTGACGGAGA TGTTTCCGGA GAGGGTCCGG TCCGCGGCCG GAGCCACTTT  
 TGATCTTCC CTCTTTGTGG CTCAGAAAAT GTACAGGTTT TCCAGGGCAG  
 CCTTGTGGAT TGGGACCACT TCCTTTATGA TCCTGGTTCT TCCCCTTGTG  
 TTTGAGACGG AGAAGTTGCA AATGGAGCAA CAGCAGCAAC TGCAGCAGCG  
 GCAGATACTT CTAGGACCTA ACACAGGGCT CTCAGGAGGA ATGCCAGGGG  
 CTCTACCCTC ACTTCCTGGA AAGATCTAG