

[19]中华人民共和国国家知识产权局

[51] Int. Cl⁷

[12] 发明专利申请公开说明书

[21] 申请号 00805597.1

C12N 15/12
C07K 14/705 A61K 38/17
C12Q 1/68 G01N 33/68
C07K 16/30 C12N 15/11
//(A61P35/00)

[43] 公开日 2002年5月29日

[11] 公开号 CN 1351661A

[22] 申请日 2000.3.20 [21] 申请号 00805597.1

[30] 优先权

[32] 1999.3.26 [33] GB [31] 9907113.6

[32] 1999.9.25 [33] GB [31] 9922858.7

[86] 国际申请 PCT/EP00/02478 2000.3.20

[87] 国际公布 WO00/58460 英 2000.10.5

[85] 进入国家阶段日期 2001.9.26

[71] 申请人 史密丝克莱恩比彻姆生物有限公司

地址 比利时里克森萨特

[72] 发明人 C·E·M·布鲁克 J·-P·卡萨特

T·科切 C·维纳尔斯 德巴索尔斯

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 张广育 杨九昌

权利要求书 4 页 说明书 47 页 附图页数 2 页

[54] 发明名称 新化合物

[57] 摘要

本发明公开了 CASB619 多肽和多核苷酸以及通过重组技术生产这种多肽的方法。本发明还公开了在诊断中应用 CASB619 多肽和多核苷酸的方法, 预防和治疗癌症, 特别是卵巢癌和结肠癌, 自身免疫病和相关病症的疫苗。

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

权 利 要 求 书

1. 含有如下氨基酸序列的分离多肽，该氨基酸序列与 SEQ ID NO: 2 的氨基酸序列的全长序列具有至少 70% 的同一性。
2. 权利要求 1 的分离多肽，其中氨基酸序列与 SEQ ID NO: 2 的氨基酸序列具有至少 95% 的同一性。
3. 权利要求 1 的分离多肽，含有 SEQ ID NO: 2 的氨基酸序列。
4. SEQ ID NO: 2 的分离多肽。
5. 含有权利要求 1-4 任一项的多肽的免疫原性片段的多肽，其中所述免疫原性片段的免疫原活性基本上与 SEQ ID NO: 2 的多肽相同。
6. 权利要求 1-5 任一项的多肽，其中所述多肽是更大融合蛋白的一部分。
7. 权利要求 1-6 任一项的多肽，与载体蛋白化学缀合。
8. 编码权利要求 1-6 任一项的多肽的分离多核苷酸。
9. 含有编码与 SEQ ID NO: 2 的氨基酸序列的全长序列具有至少 70% 同一性的多肽的核苷酸序列的分离多核苷酸或与该分离多核苷酸互补的核苷酸序列。
10. 含有与编码 SEQ ID NO: 2 的多肽的核苷酸序列的全部编码区具有至少 70% 同一性的核苷酸序列的分离多核苷酸或与该分离多核苷酸互补的核苷酸序列。
11. 含有与 SEQ ID NO: 1 的核苷酸序列的全长序列具有至少 70% 同一性的多肽的核苷酸序列的分离多核苷酸或与该分离多核苷酸互补的核苷酸序列。
12. 权利要求 8-11 任一项限定的分离多核苷酸，其中同一性为至少 95%。
13. 一种分离的多核苷酸，选自：
 - (a) 含有编码 SEQ ID NO: 2 的多肽的核苷酸序列的多核苷酸；
 - (b) SEQ ID NO: 1 的多核苷酸；和
 - (c) 通过在严紧杂交条件下用标记探针筛选合适的文库获得的多核苷酸，所述探针具有 SEQ ID NO: 1 的序列或其片段，所述多核苷酸编码一种与 SEQ ID NO: 2 的蛋白具有相似免疫原特性的蛋白，或者是与所述多核苷酸互补的核苷酸序列。
14. 含有权利要求 8-13 任一项的分离多核苷酸的表达载体或重

组活微生物。

15. 含有权利要求 14 的表达载体或权利要求 8-13 任一项的分离多核苷酸的宿主细胞。

5 16. 生产权利要求 1-7 的多肽的方法，包括在足以产生所述多肽的条件下培养权利要求 15 的宿主细胞并由培养基中回收该多肽。

17. 含有有效量的权利要求 1-7 任一项的多肽和可药用载体的疫苗。

18. 含有有效量的权利要求 8-13 任一项的多核苷酸和可药用载体的疫苗。

10 19. 含有有效量的抗原呈递细胞和可药用载体的疫苗，所述细胞通过体外负载权利要求 1-7 任一项的多肽而被修饰或在体外被遗传修饰以表达权利要求 1-7 的多肽。

20. 权利要求 17-19 任一项的疫苗，还包括 TH-1 诱导佐剂。

15 21. 权利要求 20 的疫苗，其中 TH-1 诱导佐剂选自：3D-MPL、QS21、QS21 和胆固醇的混合物以及 CpG 寡核苷酸。

22. 对权利要求 1-5 任一项的多肽或免疫片段免疫特异性的抗体。

23. 鉴定刺激或抑制权利要求 1-5 任一项的多肽功能的化合物的筛选方法，包括选自以下的方法：

20 (a) 通过直接或间接与候选化合物结合的标记，测定候选化合物与所述多肽（或负载该多肽的细胞或膜）或其融合蛋白的结合；

(b) 在标记竞争物存在时，测定候选化合物与所述多肽（或负载该多肽的细胞或膜）或其融合蛋白的结合；

25 (c) 使用对负载该多肽的细胞或细胞膜合适的检测系统，测定候选化合物是否导致该多肽激活或抑制产生的信号；

(d) 将候选化合物与含有权利要求 1-7 任一项的多肽的溶液混合形成混合物，测定混合物中的多肽活性，比较混合物与标准品的活性；或

30 (e) 使用如 ELISA 实验测定候选化合物对细胞中编码所述多肽的 mRNA 和所述多肽的产生的影响。

24. 通过免疫预防或治疗治疗受试者的方法，包括将权利要求 1-7 任一项的多肽或权利要求 8-13 任一项的多核苷酸与哺乳动物免

疫系统的细胞体外温育，体外诱导针对权利要求 1-5 任一项的分子的免疫应答，然后将这些活化的免疫细胞再输入哺乳动物中以治疗疾病。

25. 权利要求 24 的方法，其中治疗是针对卵巢癌或结肠癌。

5 26. 权利要求 1-5 的多肽的激动剂或拮抗剂。

27. 一种化合物，其为：

(a) 权利要求 1-5 的多肽的激动剂或拮抗剂；

(b) 权利要求 8-13 的分离多核苷酸；或

10 (c) 调节编码权利要求 1-5 任一项的多肽的核苷酸序列表达的核
酸分子；

其用于治疗。

28. 诊断受试者中与权利要求 1-5 任一项的多肽活性或表达有关的疾病或对疾病的易感性的方法，包括分析来自该受试者的样品中所述多肽的存在或其量。

15 29. 诊断受试者中与权利要求 8-13 任一项的多核苷酸活性或表达有关的疾病或对疾病的易感性的方法，包括分析来自该受试者的样品中所述多核苷酸的存在或其量。

20 30. 诊断受试者中与权利要求 1-5 任一项的多肽活性或表达有关的结肠癌或对结肠癌的易感性的方法，包括分析来自该受试者的样品中所述多肽的存在或其量。

31. 诊断受试者中与权利要求 8-13 任一项的多核苷酸活性或表达有关的疾病或对疾病的易感性的方法，包括分析来自该受试者的样品中所述多核苷酸的存在或其量。

32. 一种分离的多核苷酸，选自：

25 (a) 含有与 SEQ ID NO: 3 的全长序列具有至少 70% 同一性的核苷酸序列的分离多核苷酸；

(b) 含有 SEQ ID NO: 3 的多核苷酸的分离多核苷酸；

(c) SEQ ID NO: 3 的多核苷酸。

30 33. 含有权利要求 14 的表达载体或重组活微生物的活疫苗组合物。

34. 权利要求 8-13 任一项的多核苷酸在制备治疗癌症的药物中的用途。

35. 权利要求 8-13 任一项的多核苷酸在制备治疗结肠癌的药物中的用途。

36. 权利要求 1-7 任一项的多肽在制备治疗癌症的药物中的用途。

5 37. 权利要求 1-7 任一项的多肽在制备治疗结肠癌的药物中的用途。

说明书

新化合物

发明领域

5 本发明涉及多核苷酸，这里指“CASB619 多核苷酸，它们编码的多肽（这里称为“CASB619 多肽”）、重组物质以及它们的制备方法。在另一个方面，本发明涉及使用这种多肽和多核苷酸包括治疗癌症特别是结肠癌和自身免疫病及其它相关病症的方法。再一方面，本发明涉及使用本发明提供的材料鉴定激动剂和拮抗剂/抑制剂的方法，以及用鉴定的化合物治疗与 CASB619 多肽失衡有关的病症的方法。再一方面，本发明涉及检测与 CASB619 多肽活性或水平异常有关的病症的诊断试验。

据信，本发明的多肽和多核苷酸是抗肿瘤的特异性预防或治疗免疫的重要病原体，因为，相对于正常细胞，它们在肿瘤中特异性表达或高度过表达，因此，可以通过抗原特异性免疫机制被寻靶，导致肿瘤细胞的破坏。它们也可以用来诊断肿瘤细胞的出现。此外，它们在某些情况下的异常表达会引起自身免疫的诱导、异常免疫应答，这可以通过使用相同的多肽或多核苷酸进行免疫接种矫正。在这方面，最重要的生物活性是本发明多肽的抗原和免疫原活性。本发明的多肽也可以具有 CASB619 多肽的至少一种其它生物活性，这使之可以作为不同于免疫应答的治疗或预防干预的靶。

在本发明一方面涉及 CASB619 多肽。这种多肽包括含有以下氨基酸序列的分离多肽：与 SEQ ID NO: 2 全长序列至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性的氨基酸序列。这种多肽包括含有 SEQ ID NO: 2 的氨基酸的多肽。

本发明的多肽还包括其氨基酸序列与 SEQ ID NO: 2 全长序列至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性的分离多肽。这种多肽包括 SEQ ID NO: 2 的多肽。

本发明的多肽还包括由包括 SEQ ID NO: 1 所含的序列的多核苷酸编码的分离多肽。

本发明还提供 CASB619 多肽的一种免疫原性片段，即 CASB619 多肽的一个连续的部分，它与含有 SEQ ID NO: 2 的氨基酸序列的多肽具有相同或基本上相同的免疫原性活性。也就是说，该片段（必要时可与一种载体偶联）能够产生能识别 CASB619 多肽的免疫应答。

5 这种免疫原性片段可包括诸如缺少 N-末端前导序列、跨膜区域或 C-末端锚定区域的 CASB619 多肽。在一个优选的方面，根据本发明的 CASB619 的免疫原性片段基本上包括一种多肽的所有胞外结构域，其中该多肽与 SEQ ID NO: 2 在 SEQ ID NO: 2 的整个长度上至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性。

包括 CASB619 表位的肽片段一般包括 SEQ ID NO: 2 的至少 7 个，优选 9 或 10 个连续氨基酸。优选的表位示于 SEQ ID NO: 5 到 SEQ ID NO: 68。

15 包括这些表位的肽是本发明的一个优选方面。与这些表位具有相同特性的模拟表位（mimotope）和含有产生与 CASB619 分子中的表位交叉反应的免疫应答的这些模拟表位的免疫原也是本发明的部分。

因此，本发明包括含有这些表位本身和其模拟表位的分离肽。模拟表位是指与天然 CASB619 表位足够相似从而能够被识别天然分子的抗体识别的实体；（Gheysen, H.M.等，1986，“作为抗原的合成肽”，
20 Wiley, Chichester, Ciba 基金研讨会 119, p130-149; Gheysen, H.M., 1986, Molecular Immunology, 23, 7, 709-715），或者是在与合适载体偶联时能够产生与天然分子交叉反应的抗体的实体。

上述表位的肽模拟表位可通过添加、缺失或取代所选择的氨基酸用于特别的目的。因此，本发明的肽可被修饰以易于与一种蛋白载体
25 连接。例如，对于某些化学结合方法，表位中需要包括一个末端半胱氨酸。此外，对于肽与一种蛋白载体结合，需要包括一个远离肽结合末端的疏水末端，使肽的未结合的游离末端保持与载体蛋白表面的结合。这降低了肽的构象自由度，使肽更有可能呈现与肽在整个天然分子结构中具有构象非常相似的构象。例如，肽可被修改成具有一个
30 N 末端半胱氨酸和一个 C 末端疏水的酰胺化尾巴。另外，可以进行一种或多种氨基酸的 D 立体异构体的添加或取代，来制备一种有用的衍生物，用来例如增强肽的稳定性。本领域技术人员知道，这些修饰肽

或模拟表位可以是全部或部分非肽的模拟表位，其中组成的残基不限于 20 种天然氨基酸。此外，可以使用本领域已知的技术使之环化，从而限制肽的构象，使之类似于肽序列在整个分子中时的形状。环化肽的优选方法包括加入一对半胱氨酸残基以形成二硫桥。

5 并且，本领域技术人员会意识到，本发明的模拟表位或免疫原可以大于上述表位，因而可以包括本文公开的序列。因此，本发明的模拟表位可以通过在一个或两个末端加入一定数目其它天然残基的 N 和 / 或 C 末端延伸来形成。肽模拟表位也可以是天然序列的反向序列，其序列方向相反；或者序列可以全部或至少部分由 D- 立体异构体氨基酸组成（反转序列，*inverso sequence*）。并且，肽序列也可以是反向
10 反转序列，其中序列方向是相反的，氨基酸是 D- 立体异构体形式。这些反向或反向-反转肽的优势是其为非自身的，因此可以克服免疫系统自身的耐受问题。

此外，肽模拟表位可通过诸如噬菌体显示技术（EP 0 552 267 B1）
15 使用能与本发明的多肽结合的抗体来鉴定。这种技术产生大量的模拟天然肽的结构肽序列，这些肽序列因此能够结合抗天然肽的抗体，但不一定与天然多肽具有明显的序列同源性。这种方法的显著优势是可以鉴定具有较强免疫原特性的肽，或可以克服使用天然肽序列时可能出现的潜在自身抗原耐受问题。此外，该技术根据识别的模拟表位
20 序列中的共有化学特性可以鉴定各个天然肽的识别特征。

肽与免疫原性载体的共价偶联可以按本领域众所周知的方式进行。因此，例如对于直接共价偶联，可以利用碳二亚胺、戊二醛或 N - [γ-马来酰丁酰氧]琥珀酰亚胺酯，利用常用商用异源双功能接头，如 CDAP 和 SPDP（按照厂商说明）。偶联反应后，免疫原可以通过透
25 析法、凝胶过滤法和分级分离法等方便地除去并纯化。

用于本发明免疫原的载体是本领域技术人员熟知的。载体的功能是提供细胞因子辅助，以便帮助诱导抗肽的免疫应答。可用于本发明的载体的非限制性实例包括：匙孔血蓝蛋白（KLH）、血清白蛋白如牛血清白蛋白（BSA）、灭活细菌毒素如破伤风或白喉毒素（TT 和 DT）
30 或其重组片段（如 TT 片段 C 的 1 区或 DT 的转运区）或结核菌素的纯化蛋白衍生物（PPD）。或者，模拟表位或表位可以直接与脂质体载体辍合，它还可以包括能够提供 T 细胞辅助的免疫原。优选模拟表位

对载体的比例为 1: 1 到 20: 1, 优选每个载体负载 3-15 个肽。

在本发明的一个实施方案中, 优选的载体是流感嗜血杆菌的蛋白 D (EP 0 594 610 B1)。蛋白 D 是流感嗜血杆菌的 IgD 结合蛋白, 已由 Forsgren 申请专利 (WO 91/18926, 授权号 EP 0 594 610 B1)。在某些
5 情况下, 例如重组免疫原表达系统中, 可能需要使用蛋白 D 的片段, 例如蛋白 D 1/3rd (包括蛋白 D 的 N 末端 100 到 110 个氨基酸(GB 9717953.5))。

呈递本发明肽的另一种优选方法是在重组融合分子中。例如 EP 0 421 635 B 介绍了使用嵌合嗜肝 DNA 病毒核心抗原颗粒以病毒样颗粒
10 呈递外源多肽。这样, 本发明的免疫原可以包括在由乙肝病毒核心抗原组成的嵌合颗粒中呈递的肽。此外, 重组融合蛋白可以包括本发明的模拟表位和载体蛋白, 如流感病毒 NS1。对于组成本发明部分的任何重组表达蛋白, 编码这些免疫原的核酸可以也可以是本发明的一方面。

15 本发明中所用的肽可以通过本领域众所周知的固相方法方便地合成。合适的合成可以利用 T-boc 或 F-boc 方法进行。环化肽可以在全自动仪器中利用众所周知的 F-moc 方法和聚酰胺树脂通过固相法合成。或者, 本领域技术人员会知道人工进行该方法必需的实验室步骤。固相合成的技术和方法可参见“固相肽合成: 实用方法”, E. Atherton
20 和 R.C. Sheppard, 牛津大学出版社 IRL 出版 (1989)。或者, 肽可以通过重组方法制备, 包括在细菌或哺乳动物细胞系中表达编码模拟表位的核酸分子, 然后纯化表达的模拟表位。重组表达肽和蛋白的技术在本领域是已知技术, 可参见 Maniatis, T., Fritsch, E.F.和 Sambrook 等
25 《分子克隆实验室指南, 第二版》Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989)。

本发明的多肽或免疫原性片段可以是“成熟”蛋白形式, 或者可以是一种较大蛋白如前体或融合蛋白的一部分。通常优选包括含有分泌或前导序列、前序列、能协助纯化的序列如多组氨酸残基, 或能在重组制备过程中起稳定作用的额外序列。而且, 还考虑到加入外源性
30 多肽或脂类尾巴或多核苷酸序列以增加最终分子的免疫原性能力。

在一个方面, 本发明涉及遗传工程制备的可溶性融合蛋白, 这种融合蛋白含有本发明的一种多肽或其片段以及各亚类免疫球蛋白

(IgG、IgM、IgA、IgE)的重链或轻链的恒定区的各种部分。优选的是，免疫球蛋白是人 IgG 特别是 IgG1 的恒定区部分，而融合发生在铰链区。在一个特别的实施方案中，Fc 部分可简单地通过引入一个切割序列而去除，这种切割序列可用凝血因子 Xa 切割。而且，本发明
5 涉及通过遗传工程制备这些融合蛋白的方法，以及涉及它们在药物筛选、诊断和治疗中的应用。本发明的一个进一步的方面还涉及编码这种融合蛋白的多核苷酸。融合蛋白技术的例子可在国际专利申请 Nos. WO94/29458 和 WO94/22914 中发现。

蛋白质可通过化学方法接合，或以重组融合蛋白的形式表达，这种融合蛋白形式能以较非融合蛋白高的水平在一种表达系统中制备。
10 融合配体可帮助提供 T 辅助表位(免疫性融合配体)，优选的是能被人识别的 T 辅助表位；或者能帮助蛋白比原来的重组蛋白以较高产量进行表达的 T 辅助表位(表达增强子)。融合配体优选既是一种免疫性融合配体，又是表达增强子配体。

融合配体包括来自流感嗜血杆菌的蛋白 D 和来自流感病毒的非结构蛋白 NS1(凝血素)。另一种融合配体是被称为 LytA 的蛋白。优选使用该分子的 C 末端部分。LytA 来自肺炎链球菌，它合成一种 N-乙酰-L-丙氨酸酰胺酶---LytA 酰胺酶(由 lytA 基因编码{《基因》43(1986)265-272 页})，后者是一种自溶素，它特异性地降解肽聚糖骨架中的
15 某些键。LytA 蛋白的 C 末端区域负责与胆碱或某些胆碱类似物如 DEAE 的亲合性。这种特性已被用于发展能用于融合蛋白表达的大肠杆菌 C-LytA 表达质粒。在其氨基端含有 C-LytA 片段的杂合蛋白的纯化已经有介绍{《生物技术》10 卷(1992)795-798 页}。可以使用 LytA 分子
20 中在 C 末端发现的重复部分，从残基 178 开始，例如残基 188-305。

本发明还包括前面提到的多肽的变体，即通过保守性氨基酸取代与参照物不同的多肽，其中一种残基被另一种具有相似特征的残基取代。典型的这种取代在如下氨基酸之间：Ala、Val、Leu 和 Ile；Ser 和 Thr；酸性残基 Asp 和 Glu；Asn 和 Gln；碱性残基 Lys 和 Arg；或芳族残基 Phe 和 Tyr。特别优选的是其中几个、5-10、1-5、1-3、
25 1-2 或 1 个氨基酸以任何组合方式置换、缺失或增加的变体。

本发明的多肽可以任何合适的方式进行制备。这种多肽包括分离的天然产生的多肽、重组制备的多肽、合成制备的多肽、或用这些方

法的组合制备的多肽。制备这种多肽的方法在本领域是熟知的。

另一方面，本发明涉及 CASB619 多核苷酸。这些多核苷酸包括含有编码以下多肽的核苷酸序列的分离多核苷酸：这些多肽与 SEQ ID NO: 2 全长氨基酸序列有至少 70% 同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性。具有至少 97% 的同一性的多肽是高度优选的，而具有至少 98-99% 同一性的多肽是更加高度优选的，具有至少 99% 同一性的多肽是最高度优选的。这些多核苷酸包括含有编码 SEQ ID NO: 2 的多肽的 SEQ ID NO: 1 所含的核苷酸序列的多核苷酸。

本发明多核苷酸还包括含有与编码 SEQ ID NO: 2 多肽的核苷酸序列的全部编码区具有至少 70% 同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性的核苷酸序列的分离多核苷酸。具有至少 97% 的同一性的多核苷酸是高度优选的，而具有至少 98-99% 同一性的多核苷酸是更加高度优选的，具有至少 99% 同一性的多核苷酸是最高度优选的。

本发明的多核苷酸还包括含有以下核苷酸序列的分离多核苷酸：所述核苷酸序列与 SEQ ID NO: 1 全长序列有至少 70% 同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性。具有至少 97% 的同一性的多核苷酸是高度优选的，而具有至少 98-99% 同一性的多核苷酸是更加高度优选的，具有至少 99% 同一性的多核苷酸是最高度优选的。这些多核苷酸包括含有 SEQ ID NO: 1 多核苷酸的多核苷酸以及 SEQ ID NO: 1 的多核苷酸。这些多核苷酸可以插入合适的质粒或重组微生物载体中，用来免疫（参见例如 Wolff et al., Science 247:1465-1468 (1990); Corr et al., J. Exp. Med. 184:1555-1560 (1996); Doe et al., Proc. Natl. Acad. Sci. 93:8578-8583 (1996)）。本发明还提供与上述多核苷酸互补的多核苷酸。

本发明还提供 CASB619 多核苷酸片段，将其给予受试者时其具有与 SEQ ID NO: 1 的多核苷酸相同的免疫原特性。

本发明也提供编码以上定义的 CASB619 多肽的免疫原性片段的多核苷酸。

SEQ ID NO: 1 的核苷酸序列与人染色体 1 克隆 RP4-641D22 图 p13.1-13.3（入藏号 AL157901）具有同源性。SEQ ID NO: 1 的核苷

酸序列是 cDNA 序列，包括编码 1013 个氨基酸的 SEQ ID NO: 2 的多肽的序列（核苷酸 1 到 3043）。编码 SEQ ID NO: 2 的多肽的核苷酸序列可以与 SEQ ID NO: 1 所含的多肽编码序列相同，或者它可以是不同于 SEQ ID NO: 1 所含序列的序列，由于遗传密码的冗余（简并），它也编码 SEQ ID NO: 2 的多肽。SEQ ID NO: 2 的多肽不与任何已知蛋白相关。

预期本发明优选的多肽和多核苷酸也与其同源多肽和多核苷酸具有类似生物功能/特性。并且，本发明的优选多肽、免疫片段和多核苷酸具有 SEQ ID NO: 1 或 SEQ ID NO: 2 的至少一种活性。

本发明也涉及部分或其它不完整多核苷酸和多肽序列，它们在确定相应全长 SEQ ID NO: 1 和 SEQ ID NO: 2 序列前首先被鉴定。

因此，本发明一方面提供分离的多核苷酸，它：

(a) 包括与 SEQ ID NO: 3 的全长序列至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性的核苷酸序列；

(b) 具有与 SEQ ID NO: 3 的全长序列至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性的核苷酸序列；

(c) SEQ ID NO: 3 的多核苷酸；或

(d) 编码与 SEQ ID NO: 4 的全长氨基酸序列至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性的多肽的核苷酸序列；以及 SEQ ID NO: 3 的多核苷酸。

本发明还提供一种多肽，它：

(a) 包括与 SEQ ID NO: 2 的全长序列至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性的氨基酸序列；

(b) 具有与 SEQ ID NO: 2 的全长氨基酸序列至少有 70% 的同一性，优选至少 80% 的同一性，更优选至少 90% 的同一性，更优选至少 95% 的同一性，最优选至少 97-99% 的同一性的氨基酸序列；

(c) 包括 SEQ ID NO: 4 的氨基酸；和

(d) 是 SEQ ID NO: 4 的多肽；

以及由包括 SEQ ID NO: 3 所含序列的多核苷酸编码的多肽。

SEQ ID NO: 3 的核苷酸序列和其编码的肽序列来自 EST (已表达序列标志) 序列。本领域技术人员知道, 在 EST 序列中不可避免有某些核苷酸序列阅读错误 (见 Adama, M.D. et al., Nature 377 (supp)3, 1995)。因此, SEQ ID NO: 3 的核苷酸序列和由其编码的肽序列在序列准确度方面有同样的内在限制。此外, SEQ ID NO: 3 编码的肽序列包括与最同源或结构相似蛋白相同或密切同源和/或结构极类似 (如保守氨基酸差异) 的区域。

本发明的多核苷酸可以使用标准克隆和筛选技术由人结肠癌细胞 mRNA 衍生 cDNA 文库获得 (例如见 Sambrook 等《分子克隆实验室指南, 第二版》Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989)。本发明的多核苷酸也可以获自基因组 DNA 文库等天然来源或可以使用熟知的商用技术合成。

当本发明的多核苷酸用于重组生产本发明的多肽时, 多核苷酸可以包括成熟多肽的编码序列本身; 或与其它编码序列位于同一读框内的成熟多肽编码序列, 其它编码序列的例子为编码一种前导或分泌序列、前-、原-或前原-蛋白的序列或其它任何肽部分。例如, 可以编码一种能促进融合多肽纯化的标记序列。在本发明这方面的优选实施方案中, 标记序列是一种 6 组氨酸肽, 由 pQE 载体提供 (Qiagen, Inc.), 见 Gentz 等《美国科学院院刊》86: 821-824, (1989) 中的介绍; 或是一种 HA 肽尾巴。该多核苷酸还可以含有非编码 5'和 3'序列, 如转录、非翻译序列、剪接和聚腺苷化信号、核糖体结合位点和序列稳定 mRNA 的序列。

本发明进一步的实施方案包括编码多肽变体的多核苷酸, 所述变体包括 SEQ ID NO: 2 的氨基酸序列, 其中有几个例如 5 到 10、1 到 5、1 到 3、1 到 2 或 1 个氨基酸残基以任何组合方式置换、缺失或增加。

与 SEQ ID NO: 1 所含核苷酸序列相同或足够相同的多核苷酸可以用作 cDNA 和基因组 DNA 的杂交探针或作为核酸扩增 (PCR) 反应的引物, 以分离编码本发明多肽的全长 cDNA 和基因组克隆, 并分离与 SEQ ID NO: 1 具有高序列相似性的其它基因 (包括编码来源于人的共生同源物(paralogs)和来自非人物种的直向同源物(orthologs)和

共生同源物的基因)的 cDNA 和基因组克隆。一般, 这些核苷酸序列与参照物序列 70% 相同, 优选 80% 相同, 更优选 90% 相同, 最优选 95% 相同。探针或引物一般包括至少 15 个核苷酸, 优选至少 30 个核苷酸, 可以有至少 50 个核苷酸。特别优选的探针有 30 到 50 个核苷酸。特别优选的引物有 20 到 25 个核苷酸。具体而言, 来源于同源动物序列的多肽或多核苷酸可以用作免疫原, 获得对人基因交叉反应性免疫应答。

编码本发明多肽的多核苷酸包括来源于非人物种的同源物可以通过包括以下步骤的方法获得: 在严紧杂交条件下用具有 SEQ ID NO: 1 的序列或其片段的标记探针筛选合适文库; 分离含有所述多核苷酸序列全长 cDNA 和基因组克隆。这种杂交技术是本领域技术人员熟知的。优选的严紧杂交条件包括 42°C 在含有如下成分的溶液中温育过夜: 50% 甲酰胺、5x SSC (150 mM NaCl、15 mM 柠檬酸钠)、50 mM 磷酸钠 (pH7.6)、5x Denhardt's 溶液、10% 葡聚糖硫酸酯和 20 微克/ml 的变性剪切鲑精 DNA, 接着在 0.1x SSC 中在约 65°C 时洗涤滤膜。因此, 本发明也包括用具有 SEQ ID NO: 1 的序列或其片段的标记探针在严紧杂交条件下筛选合适的文库获得的多核苷酸。

本领域技术人员会意识到, 在很多情况下, 分离的 cDNA 序列是不完整的, 其中缺少 cDNA 5'末端的多肽编码区。

对本领域的技术人员来说, 有几种方法是可以使用和熟知的, 可以用来获得全长 DNAs 或延伸短的 DNAs, 例如以 cDNA 末端快速扩增 (RACE) 方法为基础的方法 (见诸如 Frohman 等, 《美国科学院院刊》85: 8998-9002, 1988)。这种技术的最近修改, 例如 Marathon™ 技术 (Clontech Laboratories Inc.), 明显简化了对较长 cDNAs 的寻找。在 Marathon™ 技术中, cDNAs 用从一种选择的组织中抽提的 mRNA 制备, 并在每个末端连上一种“接头”序列。然后使用基因特异的和接头特异的寡核苷酸组合进行核酸扩增 (PCR) 来扩增“丢失”的 DNA 的 5'末端。然后使用“巢式”引物重复 PCR 反应, 巢式引物即设计来与扩增产物的内部退火的引物 (通常为与接头序列的 3'退火的一种接头特异性引物和与所选择的基因序列的 5'退火的一种基因特异性引物)。然后用 DNA 测序对该反应的产物进行分析, 通过将产物直接与已有的 DNA 连接产生一个完整的序列, 或者使用设计 5'引物的新序

列信息进行一个单独的全长 PCR，来构建一个全长的 DNA。

5 本发明的重组多肽可以通过本领域熟知的方法由包括表达系统的遗传工程改造宿主细胞制备。因此，另一方面，本发明涉及包括本发明多核苷酸的表达系统，涉及用该表达系统遗传改造的宿主细胞，和通过重组技术生产本发明的多肽。无细胞翻译系统也可用于生产这些蛋白，利用衍生自本发明的 DNA 构建体的 RNA。

10 为进行本发明的多肽的重组制备，宿主细胞可通过遗传工程改造来整合本发明多核苷酸的表达系统或其部分。将多核苷酸导入宿主细胞可通过多种标准实验室手册介绍的方法来完成，例如 Davis 等《分子生物学基本方法》(1986) 和 Sambrook 等《分子克隆实验室指南，第二版》Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989)。优选的这种方法包括例如磷酸钙转染、DEAE-葡聚糖介导的转染、转位(transvection)、显微注射、阳离子脂介导的转染、电穿孔、转导、划痕接种、轰击导入或感染。

15 优选本发明的蛋白与反式硫氧还蛋白 (TIT) 共表达。与顺式共表达相比，反式硫氧还蛋白共表达是优选的，因为它可以使抗原中无硫氧还蛋白，无需蛋白酶。硫氧还蛋白共表达使得本发明蛋白容易溶解。硫氧还蛋白共表达也对蛋白纯化率、纯化蛋白溶解性和质量有显著影响。

20 合适宿主的代表性例子包括细菌细胞如链球菌、葡萄球菌、大肠杆菌、链霉菌、枯草芽孢杆菌细胞；真菌细胞如酵母细胞和曲霉细胞；昆虫细胞如果蝇 S2 和草地夜蛾 Sf9 的细胞；动物细胞如 CHO、COS、HeLa、C127、3T3、BHK、293、CV-1 和 Bowes 黑素瘤细胞；以及植物细胞。

25 可以使用多种表达系统，例如染色体、附加体和病毒衍生系统，例如由细菌质粒、细菌噬菌体、转座子、酵母附加体、插入元件、酵母染色体元件、病毒如杆状病毒、乳多空病毒如 SV40、痘苗病毒、腺病毒、禽痘病毒、伪狂犬病毒和逆转录病毒衍生的载体或由它们的组合物衍生的载体，例如由质粒和噬菌体遗传元件衍生的载体，如粘粒和噬菌粒。表达系统可含有调节以及引起表达的控制区域。一般来
30 说，任何适合于维持、增殖或表达多核苷酸在宿主中产生多肽的系统或载体都可使用。合适的核苷酸序列可通过任何熟知的和常规的技术

插入到表达系统中，例如 Sambrook 等《分子克隆实验室指南》(上文)中列出的技术。可将合适的分泌信号整合到目标多肽中，使翻译蛋白分泌到内质网腔中、周质空间或胞外环境中。这些信号对多肽来说可以是内源性的，或者也可以是异源性的。

5 表达系统也可是一种重组的活微生物，例如一种病毒或细菌。目标基因可插入到活的重组病毒或细菌的基因组中。接种和体内感染这种活的载体将导致抗原的体内表达和免疫应答诱导。用于此目的病毒和细菌是诸如痘病毒(例如痘苗病毒、禽痘病毒、金丝雀禽痘病毒)、甲病毒(辛德毕斯病毒、塞姆利基森林病毒、委内瑞拉马脑脊髓炎病毒)、腺病毒、腺伴随病毒、微小 RNA 病毒(脊髓灰质炎病毒、鼻病毒)、10 疱疹病毒(水痘带状疱疹病毒等)、李斯特氏菌、沙门氏菌、志贺氏菌、BCG。这些病毒和细菌可以是毒性的、或用各种方法减毒来获得一种活疫苗。这种活疫苗也是本发明的一个部分。

15 本发明的多肽可通过熟知的技术从重组细胞培养物中回收和纯化，这些技术包括硫酸铵或乙醇沉淀、酸抽提、阴离子或阳离子交换层析、磷酸纤维素层析、疏水相互作用层析、亲和层析、羟基磷灰石层析和凝集素层析。最优选将金属离子亲和层析(IMAC)用于纯化。当多肽在胞内合成、分离和/或纯化过程中变性时，可使用熟知的用于蛋白重新折叠的技术再生活性构象。

20 本发明的另一重要方面涉及在哺乳动物中诱导、强化或调节免疫应答的方法，包括用本发明的多肽或多核苷酸或其片段接种哺乳动物，其量足以产生预防或治疗癌症和自身免疫病及相关病症的抗体和/或 T 细胞免疫应答。本发明另一方面涉及在哺乳动物中诱导、强化或调节免疫应答的方法，包括通过指导编码本发明多肽的多核苷酸体内25 表达的载体或细胞给予本发明的多肽，从而诱导免疫应答，产生预防或治疗哺乳动物所述疾病的免疫应答。

30 本发明的一个进一步的方面涉及一种免疫/疫苗制剂(组合物)，将其导入哺乳动物宿主时，诱导、强化或调节哺乳动物中针对本发明多肽的免疫应答，其中组合物包括本发明的多肽或多核苷酸或本文定义的其免疫片段。疫苗制剂还可以包括合适的载体。因为多肽可以在胃中降解，优选将其经肠道外途径给药(如皮下、肌肉、静脉内或皮内注射)。适于胃肠外给药的制剂包括水性和非水性无菌注射液，其

可含有抗氧化剂、缓冲剂、抗菌剂和使制剂与接受者血液等渗的溶质；以及可包含悬浮剂或增稠剂的水性和非水性无菌悬液。制剂可置于单剂量或多剂量的容器内，例如密封的安瓶和小瓶，并可贮存于冷冻干燥的环境，只需在使用前加入无菌的液体载体。

5 本发明另一方面涉及体外诱导针对本发明多肽或多核苷酸或其片段或含有本发明多肽或多核苷酸的分子的免疫应答，使用哺乳动物免疫系统细胞，将这些活化免疫细胞再输入哺乳动物以治疗疾病。免疫系统细胞的活化可以通过在各种免疫调节分子存在或不存在时，体外温育本发明的完整多肽或多核苷酸或含有本发明多肽或多核苷酸的分子实现。本发明另一方面涉及通过给予抗原呈递细胞来免疫哺乳动物，该抗原呈递细胞通过体外负荷本发明多肽或其部分或含有本发明多肽的分子被修饰，并以免疫原性途径体内给药。或者，抗原呈递细胞可以用含有本发明多核苷酸或片段或含有本发明多核苷酸的载体体外转染以表达相应多肽，并以免疫原方式体内给药。

15 本发明的疫苗制剂也可包括佐剂系统，用来增强制剂的免疫原性。优选佐剂系统优先引发 TH1 型应答。

免疫应答大体上可分为两个典型的种类，即体液或细胞介导的免疫应答（一般分别用起保护作用的抗体和细胞效应机制来区分）。这些应答种类被称为 TH1 型应答（细胞介导的应答）和 TH2 型免疫应答（体液应答）。

20 典型的 TH1 型免疫应答的特征是产生抗原特异的单元型限制的细胞毒 T 淋巴细胞和自然杀伤型细胞应答。在小鼠中，TH1 型应答的特征通常是产生 IgG2a 亚型的抗体，而在人中，它们的对应物是 IgG1 型抗体。TH2 型免疫应答的特征是产生广谱的免疫球蛋白同种型，在小鼠中包括 IgG1、IgA 和 IgM。

25 可以想象得到，这两种类型的免疫应答之后的驱动力是细胞因子。高水平的 TH1 型细胞因子优先诱导针对所给抗原的细胞介导的免疫应答，而高水平的 TH2 型细胞因子优先诱导针对抗原的体液免疫应答。

30 TH1 和 TH2 型免疫应答的区分不是绝对的。事实上，一个个体可以支持一种被描述为 TH1 优势或 TH2 优势的免疫应答。但是，通常方便地按照 Mosmann 和 Coffman 在鼠 CD4+ve T 细胞克隆中的介绍

来考虑细胞因子的家族 (Mosmann, T.R. 和 Coffman, R.L. (1989), TH1 和 TH2 细胞: 淋巴因子的不同分泌模式导致不同的功能特性。《免疫学年鉴》7 卷, 145-173 页)。通常, TH1 型应答与 T 淋巴细胞产生 INF- γ 和 IL-2 细胞因子相关。其他通常直接与 TH1 型免疫应答诱导有关的
5 细胞因子如 IL-12 不由 T 细胞产生。相反, TH2 型应答与 IL-4、IL-5、IL-6 和 IL-13 的分泌有关。

已知特定的疫苗佐剂特别适合于刺激 TH1 或 TH2 型细胞因子应答。疫苗接种或感染后免疫应答的 TH1: TH2 平衡的最佳指标通常包括在体外用抗原再刺激后直接测量 T 淋巴细胞产生的 TH1 或 TH2 细
10 胞因子, 和/或测量抗原特异的抗体应答的 IgG1: IgG2a 比率。

因此, TH1 型佐剂是在体外用抗原再刺激时优先刺激分离的 T 细胞群体产生高水平的 TH1 型细胞因子和促进 CD8+ 细胞毒 T 淋巴细胞和与 TH1 型同种型相关的抗原特异的免疫球蛋白应答产生的佐剂。

能够优先刺激 TH1 细胞应答的佐剂在国际专利申请
15 No. WO94/00153 和 WO95/17209 中介绍。

3 De-O-酰化单磷酸类脂 A (3D-MPL) 是一种这样的佐剂。这可由 GB2220211 (Ribi) 知道。它在化学上是 3 De-O 酰化单磷酸类脂 A 与 4、5 或 6 酰基链的混合物, 并由 Ribi Immunochem, Montana 制造。3 De-O-酰化单磷酸类脂 A 的一种优选形式在欧洲专利 0 689 454
20 (SmithKline Beecham Biologicals SA) 中公开。

3D-MPL 颗粒优选小到足以通过一个 0.22 μm 微孔滤膜进行过滤除菌 (欧洲专利 0 689 454)。

3D-MPL 以每剂 10 μg -100 μg , 优选是 20-25 μg 的范围存在, 而抗原通常以每剂 2-50 μg 的范围存在。

25 另一个优选的佐剂包含 QS21, 它是一种来自 Quillaja Saponaria Molina 的茎的 Hplc 纯化的无毒组分。它可选地用来与 3 De-O-酰化单磷酸类脂 A (3D-MPL) 混合, 并可选地与一种载体一起使用。

制备 QS21 的方法在美国专利 No. 5,057,540 中公开。

30 含有 QS21 的非反应性的佐剂制剂在以前已有介绍 (WO 96/33739)。含有 QS21 和胆甾醇的这种制剂在与抗原一起配制时已显示是成功的 TH1 刺激佐剂。

优先刺激 TH1 型细胞应答的进一步的佐剂包括免疫调节性的寡核

苷酸，例如非甲基化的 CpG 序列，如 WO 96/02555 中的公开。

不同 TH1 刺激性佐剂如上文所介绍的佐剂的组合也被认为是能提供一种优先刺激 TH1 型细胞应答的佐剂。例如，QS21 可与 3D-MPL 组合配制。OS21: 3D-MPL 的比率通常为 1: 10 至 10: 1，优选为 1: 5 至 5: 1，并且通常为大约 1: 1。优选的最佳组合范围是 2.5: 1 至 1: 1 的 3D-MPL: QS21。

根据本发明，疫苗组合物还优选含有一种载体。载体可以是一种水包油乳剂，或一种铝盐如磷酸铝或氢氧化铝。

水包油乳剂优选包含一种可代谢的油，例如角鲨烯、 α -生育酚和 Tween 80。在一个特别优选的方面，根据本发明，疫苗组合物中的抗原与 QS21 和 3D-MPL 在这样一种乳剂中组合。另外，该水包油乳剂可以含有 span 85 和/或卵磷脂和/或三辛精。

在对人给药时，疫苗中 QS21 和 3D-MPL 的范围是每剂 1 μ g-200 μ g，例如 10-100 μ g，优选是 10 μ g-50 μ g。水包油通常含有 2 至 10% 的角鲨烯、2 至 10% 的 α -生育酚和 0.3 至 3% 的 Tween 80。在以更稳定的乳剂形式提供时，角鲨烯: α -生育酚: Tween 80 的比例优选相同或少于 1。还可包含 1% 水平的 Span85。在某些情况下，本发明的疫苗优选进一步含有一种稳定剂。

非毒性的水包油乳剂优选在一种水载体中含有一种非毒性的油如角鲨烷或角鲨烯、一种乳化剂如 Tween 80。水载体可以是例如磷酸缓冲盐水。

WO 95/17210 中介绍了一种特别强有力的佐剂制剂，它含有在一种水包油乳剂中的 QS21、3D-MPL 和生育酚。

本发明还提供一种多价的疫苗组合物，它含有本发明的疫苗制剂以及其他抗原，特别是对治疗癌症、自身免疫疾病和相关病症有用的抗原。这样一种多价疫苗组合物可包括一种如前所述的 TH1 诱导型佐剂。

本发明也涉及多核苷酸和多肽作为诊断试剂的用途，多核苷酸为衍生自本发明的多核苷酸的引物形式，多肽为本发明的多肽特异的抗体或试剂形式。

血液或组织中能够检测癌发生途径极早期变化的遗传或生化标志的鉴定可以帮助确定患者的最佳治疗。替代肿瘤标志如多核苷酸表达

可用于诊断癌症的不同形式和状态。鉴定本发明多核苷酸的表达水平可以用来区分癌症的阶段并对癌组织的性质分级。区分癌症阶段是监测癌症的进展，根据在活组织检查区域是否存在恶性组织来确定。本发明的多核苷酸通过鉴定癌症侵袭性标志如在体内不同区域的存在，
5 可以帮助完善区分阶段的方法。癌症分级表明肿瘤与其同类正常组织的近似程度，是由细胞形态和其它分化标志来评定的。本发明的多核苷酸可以用来确定肿瘤级别，因为它们可以帮助确定肿瘤细胞的分化状态。

10 诊断实验通过包括确定来自受试者的样品中异常降低或增加的多肽或 mRNA 水平的方法的诊断，提供了诊断或确定对癌症、自身免疫病和相关病症的易感性的方法。该诊断方法已知是分化表达。比较病态组织和正常组织中特定基因的表达。两种组织中多核苷酸相关基因、mRNA 或蛋白在分子量、氨基酸或核苷酸序列或相对丰度方面的差异，表明在怀疑患病的人的组织中基因或调控它的基因的变化。

15 降低或增加的表达可以在 RNA 水平测定。首先从两种组织中分离 PolyA RNA，然后使用含有 Poly A+ mRNA 的 RNA 印迹或任何其它直接或间接 RNA 检测法，通过例如组织切片原位杂交、逆转录酶 PCR 等，可以检测相当于差异表达的本发明多核苷酸的基因编码的 mRNA。与正常组织相比病态组织中给定 RNA 的表达增加或降低，
20 表明转录本和/或表达的蛋白在疾病中起作用。因此，检测到较正常组织高或低水平的相应于 SEQ ID NO: 1 或 3 的 mRNA 表明在患者中存在癌症。

样品中 mRNA 表达水平可以通过产生来自样品的已表达序列标志 (EST) 文库来确定。在文库中 EST 的相对出现可以用来评价在起始样品中基因转录本的相对出现。该 EST 分析实验可以与参考样品的 EST 分析相比较，从而确定目标多核苷酸的相对表达水平。

其它 mRNA 分析可以使用基因表达系列分析 (SAGE) 法 (Velculescu et al., Science (1995) 270-484)、差示法 (例如 US5,776,683) 或依赖于核苷酸相互作用特异性的杂交分析来进行。

30 或者，比较可以在蛋白水平进行。两种组织中的蛋白大小可以使用抗体比较，在来自两种组织的蛋白提取物的蛋白质印迹实验中检测多肽。表达水平和亚细胞定位也可以使用针对相应蛋白的抗体免疫测

定。本领域技术人员熟知其它可以用来测定来自一个宿主的样品中的蛋白如本发明多肽水平的实验技术。与正常组织相比病态组织中多肽表达水平升高或降低表明表达的表可能表达可能与疾病有关。

5 在本发明的实验中，诊断可以通过测定 SEQ ID NO: 1 或 3 所示至少一个序列编码的基因产物的表达水平来确定。病态和正常组织中的 mRNA 或蛋白水平的比较也可以用来追踪疾病进程或消退。

样品中的多种多核苷酸序列可以使用多核苷酸阵列分析。这些可用来揭示基因的差异表达并确定基因功能。例如，SEQ ID NO: 1 或 3 的多核苷酸序列阵列可以用来确定正常和癌变组织中是否有任何多核苷酸表达差异。在一个实施方案中，可以构建含有 SEQ ID NO: 1 或 3 的核苷酸序列或其片段的寡核苷酸探针阵列，以便有效筛选遗传突变。阵列技术是众所周知的，可以广泛应用，可以用来解决分子遗传学中的多种问题，包括基因表达、遗传连锁和遗传可变性（参见例如 M. Chee et al., Science, Vol 274, pp610-613 (1996)）。

15 “诊断”在本发明中包括确定受试者对疾病的易感性，确定受试者是否患病，以及确定患者的预后。

本发明还涉及进行诊断实验的诊断试剂盒，它包括：

(a) 本发明的多核苷酸，优选 SEQ ID NO: 1 或 3 的核苷酸序列或其片段；

20 (b) 与(a)互补的核苷酸序列；

(c) 本发明的多肽，优选 SEQ ID NO: 2 或 4 的多肽或其片段；

或

(d) 本发明多肽的抗体，优选 SEQ ID NO: 2 或 4 的多肽的抗体。

25 本发明的核苷酸序列对染色体定位也有价值。该序列特异性靶向人染色体上特定位置并与其杂交。根据本发明对染色体相关序列的作图是将这些序列与基因相关疾病联系起来的重要第一步。序列被作图到染色体上的准确位置后，序列在染色体上的物理位置可以与遗传图谱数据联系起来。这些数据可参见例如 V. McKusick, Mendelian Inheritance in Man (通过 Jhon Hopkins University Welch 医学图书馆在线获得)。然后已经被作图的澳相同染色体区域的基因和疾病的关系可以通过连锁分析（物理相邻基因的共遗传）鉴定。也可以确定受影响和未受影响个体之间 cDNA 或基因组相邻的差异。

本发明多肽和其片段或类似物或表达该多肽的细胞也可以用作免疫原，产生对本发明多肽免疫特异性的抗体。“免疫特异性”是指抗体对本发明多肽的亲合力明显大于其对现有技术其它相关多肽的亲合力。

5 本发明另一方面提供对本发明多肽或本文中限定的其免疫片段免疫特异性的抗体。优选该抗体是单克隆抗体。

针对本发明多肽的抗体也可以通过常规途径向动物优选是非人动物给予多肽或具有表位的片段、类似物或细胞获得。为了制备单克隆抗体，提供通过连续细胞系培养物产生的抗体的任何技术均可使用。

10 实例包括杂交瘤技术（Kohler,G和 Milstein, C.《自然》256: 495-497（1975））、三瘤(trioma)技术、人 B 细胞杂交瘤技术（Kozbor 等《今日免疫学》4: 72（1983））和 EBV 杂交瘤技术（Cole 等《单克隆抗体和癌症治疗》，77-96 页，Alan R. Liss, Inc., 1985）。

15 制备单链抗体的技术（美国专利 No. 4,946,778）可修改来制备针对本发明的多肽的单链抗体。转基因鼠或其他生物体或动物包括其他哺乳动物也可用来表达人源化抗体。

上述抗体可以用来分离或鉴定表达多肽的克隆或通过亲和层析纯化多肽。本发明的抗体也可以用来防止或治疗癌症，特别是卵巢癌和结肠癌、自身免疫病和相关病症。

20 本发明其它方面涉及在哺乳动物中诱导或调节免疫应答的方法，包括用本发明的多肽接种哺乳动物，其量足以产生保护或缓解疾病症状或进程的抗体和/或 T 细胞免疫应答。本发明的另一方面涉及在哺乳动物中诱导或调节免疫应答的方法，包括通过体内指导本发明多核苷酸表达和编码多肽的载体给予本发明的多肽，以诱导免疫应答产生抗
25 体保护该动物免于该疾病。

因此，应当意识到，本发明提供一种治疗与 CASB619 多肽活性存在、过量或低下有关的异常病症如癌症和自身免疫病特别是卵巢癌和结肠癌的方法。

30 本发明还提供一种筛选化合物鉴定刺激或抑制 CASB619 多肽功能的化合物的方法。一般，激动剂或拮抗剂可以用于上述疾病的治疗和预防目的。可以从多种来源鉴定化合物，例如细胞、无细胞制备物、化学文库和天然产物化合物。这样鉴定的激动剂、拮抗剂或抑制剂根

据情况可以是多肽的天然或修饰底物、配体、受体、酶等；或者可以是其结构或功能模拟物（见 Coligan et al 《当代免疫学方法》 1(2)第 5 章(1991)）。筛选方法是本领域技术人员熟知的。其它筛选方法可参见例如 D. Bennett et al., J Mol Recognition, 8:52-58 (1995); and K. Johanson et al., J Biol Chem, 270(916):9459-9471 (1995)及其中的参考文献。

因此，本发明提供一种筛选方法以鉴定刺激或抑制本发明多肽功能的化合物，包括选自以下的方法：

10 (a) 通过直接或间接与候选化合物结合的标记，测定候选化合物对多肽（或负载多肽的细胞或膜）或其融合蛋白的结合；

(b) 在标记竞争物的存在下测定候选化合物对多肽（或负载多肽的细胞或膜）或其融合蛋白的结合；

(c) 使用对负载多肽的细胞或细胞膜合适的检测系统，测试候选化合物是否导致多肽活化或抑制产生的信号；

15 (d) 将候选化合物与含有权利要求 1 的多肽的溶液混合，形成混合物，测定混合物中多肽的活性，比较混合物与标准品的活性；或

(e) 使用如 ELISA 实验等测定候选化合物对细胞中编码该多肽的 mRNA 和该多肽的产生的影响。

20 本发明的多肽可以用来通过本领域标准的受体结合技术鉴定膜结合或可溶的受体。熟知的筛选方法也可以用来鉴定竞争与本发明多肽受体结合的本发明多肽的激动剂和拮抗剂。

因此，另一方面，本发明涉及鉴定本发明多肽的激动剂、拮抗剂、配体、受体、底物、酶等或降低或提高这些多肽产生的化合物的筛选试剂盒，其包括：

25 (a) 本发明的多肽；

(b) 表达本发明多肽的重组细胞；

(c) 表达本发明多肽的细胞膜；或

(d) 本发明多肽的抗体；

该多肽优选是 SEQ ID NO: 2 的多肽。

30 本领域技术人员容易意识到，本发明的多肽也可以用于基于结构设计多肽激动剂、拮抗剂或抑制剂的方法，具体如下：

(a) 首先确定多肽的三维结构；

(b) 推测激动剂、拮抗剂或抑制剂的类似反应或结合位点的三维结构；

(c) 合成预期与推测的结合或反应位点结合或反应的候选化合物；

5 (d) 测试候选化合物是否确实是激动剂、拮抗剂或抑制剂。

基因治疗也可以用来影响通过受试者相关细胞的 CASB619 多肽的内源产生。基因治疗的概述参见《人类分子遗传学》第 20 章基因治疗和其它基于分子遗传学的治疗途径（及其中的参考文献），T Strachan 和 A P Read, BIOS Scientific Publishers Ltd (1996)。

10 疫苗制剂的一般描述见《药物生物技术》61 卷，疫苗设计 - 亚单位和佐剂途径，Powell and Newman 编，Plenum Press, 1995。疫苗新趋势和发展，Voller 等编，University Park Press, Baltimore, Maryland, USA。包囊在脂质体中可参见例如 Fullerton 美国专利 4,235,877。蛋白辘合到大分子参见例如 Likhite 美国专利 4,372,945 和 Armor 等美国
15 专利 4,474,757。

选定每剂疫苗中蛋白量为在典型疫苗接受者中诱导免疫保护应答而没有明显的副作用。这个量取决于使用的具体免疫原。一般，预期每剂含有 1 - 1000 μ g 蛋白，优选 2 - 100 μ g，最优选 4 - 40 μ g。特定疫苗的最适量可以通过涉及观察抗体滴度和受试者中其它应答的标准研究
20 确定。初始接种后，受试者一般在约 4 周内接受加强免疫。

“分离的”表示“通过人工”改变其天然状态，即如果“分离的”组合物或物质在自然界中存在，则它已经从其原始环境改变或取出或者已经取出并改变。例如，在本文中使用该术语时，在活动物体内存
25 在的多核苷酸或多肽不是“分离的”，但已经与天然状态下共存的物质分开的同样多核苷酸或多肽就是“分离的”。

“多核苷酸”一般指任何聚核糖核苷酸或聚脱氧核糖核苷酸，它可以是修饰或非修饰的 RNA 或 DNA，包括单链和双链区域。

“变体”指不同于参照多核苷酸或多肽但保留了基本特性的多核苷酸或多肽。一般多核苷酸变体的核苷酸序列与参照多核苷酸不同。
30 变体核苷酸序列的变化可能改变或不改变由参照多核苷酸编码的多肽的氨基酸序列。如下所述，核苷酸改变可能导致参照序列编码的多肽中氨基酸的置换、添加、缺失、融合和截短。一般多肽变体的氨基酸

序列与参照多肽不同。通常，差异有限，因此参照多肽和变体的序列在总体上是极其近似的，在许多区域是相同的。变体和参照多肽在氨基酸序列上可以有任何组合形式的一个或多个置换、添加、缺失的区别。置换或插入的氨基酸残基可以是或不是遗传密码编码的氨基酸。

5 多核苷酸或多肽的变体可以是天然存在的，如等位变体，或者是非天然存在的。非天然存在的多核苷酸和多肽变体可以通过诱变技术或直接合成制备。

“同一性”如同本领域所熟知的，是两种或多种多肽序列或两种或多种多核苷酸序列之间的相关性，根据具体情况，可通过序列比较来确定。在本领域，“同一性”还指多肽或多核苷酸序列之间序列相关的程度，根据具体情况，可通过这些序列的链的匹配来确定。“同一性”可使用已知的方法方便地计算，包括但不限于(《计算分子生物学》，Lesk, A.M. 编辑, Oxford University Press, New York, 1988; 《生物计算机：信息和基因组计划》Smith, D.W.编辑, Academic Press, New York, 1993; 《序列数据的计算机分析》I 部分, Griffin, A.M.和 Griffin, H.G.编辑, Humana Press, New Jersey, 1994; 《分子生物学中的序列分析》von Heine, G., Academic Press, 1987; 和《序列分析引物》Gribskov, M.和 Devereux, J. 编辑, M Stockton Press, New York, 1991; 和 Carillo, H.和 Lipman, D., 《SIAM J. Applied Math》48: 1073(1988)。测定同一性的方法被设计来给出检测序列之间的最大匹配。而且，测定同一性的方法在公众可获得的计算机程序中编码。用来测定两种序列之间的同一性的计算机程序方法包括但不限于 GCG 程序包中的 GAP 程序 (Devereux, J.等《核酸研究》12(1): 387(1984)、BLASTP、BLASTN (Altschul,S.F.等《分子生物学杂志》215: 403-410(1990)以及 FASTA (Pearson 和 Lipman 《美国科学院院刊》85: 2444-2448(1988)。BLAST 程序家族可从 NCBI 和其他来源获得(《BLAST 手册》Altschul, S.等, NCBI NLM NIH Bethesda, MD 20894; Altschul, S.等《分子生物学杂志》215: 403-410(1990)。著名的 Smith Waterman 算法也可用来测定同一性。

25 30 优选使用的算法是 FASTA。使用该算法进行多肽或多核苷酸序列比较的优选参数包括以下：

缺口补偿： 12

缺口延伸补偿： 4

单字大小: 2, 最大 6

用其它方法进行多肽序列比较的优选参数包括以下:

1) 算法: Needleman 和 Wunsch, 《分子生物学杂志》48: 443-453 (1970)

5 比较矩阵: BLOSSUM62, 见 Henikoff 和 Henikoff, 《美国科学院院刊》89: 10915-10919 (1992)

缺口补偿: 12

缺口长度补偿: 4

10 使用这些参数的程序可以从 Genetics Computer Group, Madison WI 以“缺口”程序的形式获得。上述参数是多肽比较的缺省参数(对末端缺口无补偿)。

多核苷酸序列比较的优选参数包括以下:

1) 算法: Needleman 和 Wunsch, 《分子生物学杂志》48: 443-453 (1970)

15 比较矩阵: 匹配=+10, 不匹配=0

缺口补偿: 50

缺口长度补偿: 3

来源: Genetics Computer Group, Madison WI 的“缺口”程序。

20 使用这些参数的程序可以从 Genetics Computer Group, Madison WI 以“缺口”程序的形式获得。上述参数是多核苷酸比较的缺省参数。

25 例如, 本发明的多核苷酸序列可以与 SEQ ID NO: 1 的参照序列相同, 即同一性为 100%, 或者与参照序列相比包括一定数量的核苷酸改变。这种改变可以选自至少一个核苷酸缺失、置换(包括转换和颠换)或插入, 其中所述改变可以发生在参照多核苷酸序列的 5'或 3'末端或两个末端之间的任何位置, 分别散布在参照序列的核苷酸之间, 或者以一个或多个毗连群散布在参照序列中。核苷酸改变的数量如下确定: SEQ ID NO: 1 中的总核苷酸数乘以相应同一性百分数(除以 100), 然后将其结果从 SEQ ID NO: 1 的总核苷酸数中减除, 或者:

30
$$n_n \leq x_n - (x_n \cdot y)$$

其中 n_n 是核苷酸改变的数目, x_n 是 SEQ ID NO: 1 中的核苷酸总

数，y 对 70%来说是 0.70，对 80%来说是 0.80，对 85%来说是 0.85，对 90%来说是 0.90，对 95%来说是 0.95，依此类推，其中如 x_n 和 y 的结果不是整数，则将其四舍五入取最近的整数，再从 x_n 中减除。编码 SEQ ID NO: 2 的多肽的多核苷酸的改变可能在该编码序列中产生无义、错义或移码突变，因此改变后的多核苷酸编码的多肽会发生变化。

类似地，本发明的多肽序列可以与 SEQ ID NO: 2 的参照序列相同，即同一性为 100%，或者与参照序列相比包括一定数量的氨基酸改变，因此同一性小于 100%。这种改变可以选自至少一个氨基酸缺失、置换（包括保守和非保守置换）或插入，其中所述改变可以发生在参照多肽序列的氨基或羧基末端或两个末端之间的任何位置，分别散布在参照序列的氨基酸之间，或者以一个或多个毗连群散布在参照序列中。同一性百分数一定时氨基酸改变的数量如下确定：SEQ ID NO: 2 中的总氨基酸数乘以相应同一性百分数（除以 100），然后将其结果从 SEQ ID NO: 2 的总氨基酸数中减除，或者：

15

$$n_a \leq x_a - (x_a \cdot y)$$

其中 n_a 为氨基酸改变数量， x_a 为 SEQ ID NO: 2 的总氨基酸数，y 值在 70% 时为 0.70，在 80% 时为 0.80，在 85% 时为 0.85，依此类推，其中若 x_a 和 y 的结果不是整数，则将其四舍五入取最近的整数，再从 x_a 中减除。

“同系物”是本领域的通用术语，表示与目标序列具有高度序列相关性的多核苷酸或多肽序列。这种相关性可以通过确定以上所述待比较的序列之间的同一性和/或类似性程度来定量。属于该通用术语范围内的术语是“直向同源物”和“共生同源物”，前者是其它物种中多核苷酸或多肽的功能等价物的多核苷酸或多肽，后者是指在相同物种内考虑时的功能类似序列。

30

实施例

实施例 1:

实时 RT-PCR 分析

实时 RT-PCR (U. Gibson, 1996. Genome Research:6, 996) 用来比较来自多个患者的相应的肿瘤和正常结肠组织中候选抗原 mRNA 转录本的丰度。此外, 一系列正常组织中候选基因的 mRNA 水平也通过该方法进行了评价。

5 正常和癌变结肠总 RNA 使用 TriPure 试剂 (Boehringer) 由速冻活检样品中提取。正常组织总 RNA 购自 InVitrogen 或使用 TriPure 试剂由速冻活检样品中提取。DNA 酶处理后使用 oligo-dT 磁性珠 (Dyna) 从总 RNA 中纯化 PolyA+ mRNA。使用 SybrII 染料 (Molecular Probes) 通过分光荧光测定法 (VersaFluor, BioRad) 进行 mRNA 的定
10 量。实时 PCR 扩增的引物采用 Perkin-Elmer Primer Express 软件使用 TaqMan 扩增条件的缺省选择设计。

实时反应根据标准 PCR 方法组合, 每一反应使用 2 ng 纯化 mRNA。实时检测时加入 SybrI 染料 (Molecular Probes) 至最终稀释度为 1/75000。扩增 (40 个循环) 和实时检测在 Perkin-Elmer Biosystems
15 PE7700 系统中使用常规仪器设置进行。Ct 值使用 PE7700 序列测定软件计算。获得每一个患者的两个 Ct 值: 肿瘤 Ct (CtT) 和相应正常结肠 Ct (CtN)。实时 PCR 获得的 Ct 值与靶模板拷贝数是对数-线性相关。因为在常规实验条件下 PCR 扩增的效率与理论扩增效率接近, 由 $2^{(CtN-CtT)}$ 可估计两种组织中相对转录水平 (即肿瘤中 mRNA 过表达
20 倍数)。对 24 个患者的活检样品进行实时 PCR 反应。计算每个患者的 mRNA 过表达水平。然后由该数据集计算候选抗原的平均 mRNA 过表达水平和患者过表达候选抗原的比例。单个值相对同样样品中的肌动蛋白标准化 (比例), 如图 1 所示。该值为 1 相当于与肌动蛋白表
达水平相同。结果以对数标尺示出。

25 代表 28 种不同组织的 48 个正常组织样品也通过同样方法测定。候选抗原的 Ct 值与相同组织样品中获得的肌动蛋白的值比较。结果相对于肌动蛋白标准化, 示于图 2。

结肠癌/正常结肠样品中的实时 PCR 结果总结

结肠癌中 CASB619 过表达的患者 (%)	结肠癌中平均过表达水平 (倍)
16/24 (67%)	208

结论：相对于邻近正常结肠，CASB619 在 67% 结肠癌样品中过表达，平均水平为近 200 倍。正常组织中的表达限于其它消化道组织以及主要的气管和睾丸。

5 实施例 2:

DNA 微阵列

DNA 微阵列用来检查多个样品中大量基因的 mRNA 表达图谱。该信息用来弥补实时 PCR 得到的数据，提供了肿瘤和正常组织中基因表达水平的独立度量。

10 目前产生 DNA 微阵列技术的实例包括 1) Affymetrix “基因芯片 (GeneChip)” 阵列，其中寡核苷酸使用光能无机图形法 (photolithographic process) 在芯片表面合成 2) DNA 斑点技术，其中少量 DNA 溶液通过机器人沉积然后固定在固相表面 (如玻璃)。两种情形下，芯片均以自目标组织 (如正常组织、肿瘤等) 中提取的 cDNA 或 cRNA 杂交并
15 以放射活性或以荧光报告分子标记。标记的材料与芯片杂交，结合在芯片上各个序列的探针的量使用特制的扫描仪确定。实验可以使用单荧光报告分子 (或放射活性) 进行，或者可以用两种报告分子进行。后一种情形下，两种样品中的每一个均以一种报告分子标记。两种标记样品然后竞争与 DNA 芯片上的序列杂交。确定芯片上每个序列的
20 两种荧光信号的比例。该比例用来计算两种样品中转录本的相对丰度。详细方法可获自多种来源，包括 “DNA 微阵列：实用方法, Schena M. Oxford University Press 1999 ” 和 互 联 网 (<http://cmgm.stanford.edu/pbrown/protocols/index.html>, <http://arrayit.com/DNA-Microarray-Protocols/>) 和 特 定 销 售 商 (如
25 Affymetrix)

实施例 3

EST 图谱

对实验抗原组织表达鉴定的补充途径是探测人 “已表达序列标志 (EST)” 数据库。EST 是由特定组织或细胞系提取的多种 mRNA 制备的 cDNA 的小片段。这些数据库目前提供来自数百种 cDNA 组织文库包括各类肿瘤组织和疾病状态的大量 EST (106)。通过检索工具
30

(Blast), 进行 CASB616 序列的比较检索, 以便进一步探索组织表达。

CASB619 的 EST 分布

DbEST 入藏号	ATG lib ID	描述
NCBI:1202616	937	NCI_CGAP_Co2 (结肠绒毛)
NCBI:1202659	937	NCI_CGAP_Co2 (结肠绒毛)
NCBI:1208269	935	NCI_CGAP_Pr4.1 (前列腺上皮内赘生物)
NCBI:1152744	888	NCI_CGAP_Pr6 (前列腺上皮内赘生物)
NCBI:1157532	910	NCI_CGAP_Pr22 (正常前列腺)
NCBI:1178873	882	NCI_CGAP_Co3 (12 个合并的结肠癌)
NCBI:1040601	628	睾丸 NHT
NCBI:1056221	628	睾丸 NHT
NCBI:1298131	910	NCI_CGAP_Pr22 (正常前列腺)
NCBI:976517	715	附睾
NCBI:1618753	417	甲状旁腺瘤
NCBI:1737305	895	NCI_CGAP_Br2 (合并的乳腺癌组织)
NCBI:1738600	895	NCI_CGAP_Br2 (合并的乳腺癌组织)
NCBI:1767617	628	睾丸 NHT
NCBI:1889992	424	牛心
NCBI:1907671	628	睾丸 NHT
NCBI:2074937	1076	NCI_CGAP_Lu5 View stats
NCBI:2147596	1461	NCI_CGAP_Ut1
NCBI:2147632	1461	NCI_CGAP_Ut1
NCBI:2308815	1462	NCI_CGAP_Ut2
NCBI:2441896	1410	NCI_CGAP_Pr28
NCBI:2447648	1463	NCI_CGAP_Ut3
NCBI:2381271	1461	NCI_CGAP_Ut1
NCBI:2582781	1461	NCI_CGAP_Ut1
NCBI:2583713	1461	NCI_CGAP_Ut1
NCBI:2585593	1410	NCI_CGAP_Pr28
NCBI:2587322	1463	NCI_CGAP_Ut3
NCBI:2601136	1461	NCI_CGAP_Ut1
NCBI:2824305	1463	NCI_CGAP_Ut3
NCBI:3007221	1461	NCI_CGAP_Ut1
NCBI:3078572	2301	NCI CGAP Pit1
NCBI:3086989	1461	NCI_CGAP_Ut1
NCBI:3087235	1461	NCI_CGAP_Ut1
NCBI:3045696	1661	NCI_CGAP_Lu19
NCBI:3218161	2467	NCI CGAP Co20
NCBI:3291383	2508	NCI_CGAP_Sub3
NCBI:3028913	2107	BT130
NCBI:3655706	1447	NCI_CGAP_Co14
NCBI:2553046	1728	Soares_Dieckgraefe_结肠_NHCD
NCBI:979345	781	子宫内膜瘤
NCBI:3289738	2508	NCI_CGAP_Sub3
NCBI:614598	464	?
NCBI:978076	781	子宫内膜瘤
NCBI:979060	781	子宫内膜瘤
NCBI:1616799	417	甲状旁腺瘤

总之，与 CASB619 匹配的 93% EST 来源于肿瘤组织、牛组织或正常生殖器官。这表明该基因在该组织的表达更常见。

实施例 4

5 RNA - DNA 印迹分析

通过 Advantage PCR (见上) 扩增限定量的混合肿瘤和相应的正常结肠 cDNA。多个正常组织的信使 RNA 也使用相同的方法扩增。扩增的 cDNA (1 μ g) 在 1.2% 琼脂糖凝胶上电泳，并转移到尼龙膜上。该膜与用候选 TAA cDNA 片段制备的探针杂交 (AlkPhos Direct System)。RNA - DNA 分析提供了转录本大小、剪接变体存在和肿瘤
10 和正常组织中转录本丰度的信息。

实施例 5

RNA 印迹分析

15 RNA 印迹使用 1 μ g polyA+ mRNA 根据标准方法制备。放射活性探针使用 Ready-to-Go 系统 (Pharmacia) 制备。

实施例 6

鉴定全长 cDNA 序列

20 结肠癌 cDNA 文库使用 λ Zap II 系统 (Stratagene) 由 5 微克 polyA+ mRNA 构建。按照提供的方法，但在逆转录步骤使用 SuperscriptII (Life Technologies)。构建 Oligo-dT 引发的和随机引发的文库。筛选每一文库时涂布约 1.5×10^6 独立噬菌体。嗜菌斑转移到尼龙滤膜上，与用 AlkPhos Direct 标记的 cDNA 探针杂交。阳性噬菌体通过化学发光检测。
25 由琼脂平板上切下阳性噬菌体，在 500 微升 SM 缓冲液中洗脱，通过基因特异性 PCR 证实。洗脱的噬菌体通过体内切除转变成单链 M13 噬菌体。噬菌体然后通过感染大肠杆菌转变成双链质粒 DNA。感染的细菌铺板，然后用 cDNA 探针进行第二轮筛选。由阳性细菌克隆纯化质粒 DNA 并对双链测序。当全长基因不能直接由 cDNA 文库
30 获得时，缺失的序列使用 RACE 技术分离 (Marathon Kit, ClonTech)。该方法依赖于逆转录 mRNA 成双链 cDNA，将接头连接到 cDNA 的末

端，并使用基因特异性引物和一个接头寡核苷酸扩增所需 cDNA 的末端。Marathon PCR 产物克隆进质粒 (pCRII-TOPO) 中并测序。

获得的序列 (SEQ ID NO: 1) 具有一个 1013 氨基酸的推测开放读框 (SEQ ID NO: 2)。对推测的蛋白序列进行了细胞定位预测计算 (PSORT: <http://psort.nibb.ac.jp/> and TopPred: http://www.biokemi.su.se/~sever/toppred2/toppred_source.html) 。CASB619 似乎具有肽信号，和一到三个跨膜区 (低置信度预测)。未发现其它基元或结构域。

10 实施例 7:

7.1 肿瘤特异性抗原的表达和纯化

在微生物宿主中的表达或体外转录/翻译系统被用于产生用作疫苗目的的本发明抗原，或产生通过免疫组织化学鉴定天然表达蛋白所需的抗体的快速纯化和生成或进一步纯化所需的蛋白片段或完整蛋白。

15 重组蛋白可在两种微生物宿主大肠杆菌和酵母 (如酿酒酵母或巴斯德毕赤酵母) 中表达。这可以选择对于特定抗原产生最有利特性的表达系统。一般，重组抗原将在大肠杆菌中表达，试剂蛋白在酵母中表达。

表达策略首先涉及重组抗原的一级结构。一般，表达融合配偶体 (EFP) 被置于 N 末端，以提高表达水平，在 N 末端还可以包括用来调节抗原的免疫原特性的区域、免疫融合配偶体 (IFP)。此外，亲和用来帮助进一步纯化的融合配偶体 (AFP) 被置于 C 末端。

当获得了重组菌株，通过评价表达水平鉴定重组产物，通过分析粗提取物的行为预测蛋白的进一步溶解性。

25 在合适的生长培养基中生长和诱导重组蛋白表达后，通过 SDS - PAGE 分析全部提取物。重组蛋白在染色凝胶上显色，使用特异性抗体通过蛋白印迹分析鉴定。

不同型式表达抗原的比较可以选择出最有希望的候选者，它将用于进一步纯化和免疫分析。

30 纯化方案使用重组蛋白中存在 His 亲和尾时的常规方法。在典型的实验中，过滤破碎细胞，将无细胞提取物上样到特异性保留重组蛋白的离子金属亲和层析 (IMAC; Ni⁺⁺NTA, 来自 Qiagen) 上。

保留的蛋白通过磷酸缓冲液中的 0 - 500 mM 咪唑梯度（可能存在变性剂）洗脱。

7.2 抗体制备和免疫组织化学

5 少量相对纯化的蛋白可被用于产生免疫工具，以便

- a) 通过正常或癌症组织切片的免疫组织化学检测表达；
- b) 检测表达并在纯化过程中跟踪蛋白（ELISA/蛋白印迹）；或
- c) 鉴定/定量纯化蛋白（ELISA）。

7.2.1 多克隆抗体：

10 免疫

用 100 μ g 蛋白通过肌肉内途径免疫 2 - 3 只兔子 3 次，每次间隔 3 周，蛋白配制在佐剂 3D - MPL/QS21 中。每次免疫 3 周后取血样，使用作为包被抗原的蛋白按照标准方法通过 ELISA 测定血清中的抗体滴度。

15 ELISA

96 孔微量板（maxisorb Nunc）用 5 μ g 蛋白 4 $^{\circ}$ C 包被过夜。37 $^{\circ}$ C 用 PBS NCS 1% 饱和 1 小时后，加入兔血清的系列稀释液（1/10 开始），37 $^{\circ}$ C 1H 30。PBS Tween 洗涤 3 次后，加入抗兔生物素化抗血清（Amersham）（1/5000）。洗涤平板，加入过氧化物酶偶联的链亲和素（1/5000），37 $^{\circ}$ C 30 分钟。洗涤后，加入 50 μ l TMB (BioRad)7 分钟，然后用硫酸终止反应。在 450nm 测定 OD，通过 SoftmaxPro 计算中点稀释度。

7.2.2 单克隆抗体：

免疫

25 用 5 μ g 纯化蛋白以 3 周的间隔免疫 5 只 BALB/c 小鼠 3 次。第 2 次后 14 天和第 3 次后 1 周取血。以纯化蛋白作为包被抗原通过 Elisa 测定血清。根据这些结果（中点稀释度 > 10000），选择 1 只小鼠用于融合。

融合/HAT 选择

30 根据标准方法使用 PEG 40% 和 DMSO 5% 将脾细胞与 SP2/0 骨髓瘤融合。细胞接种 96 孔板，每孔 2.5×10^4 - 10^5 细胞，在 HAT 培养基上选择抗性克隆。检测杂交瘤的上清中特异性抗体的含量，如结果为阳

性，将杂交瘤进行 2 轮有限稀释。2 轮筛选后，选择 3 个杂交瘤用于腹水生成。

7.2.3 免疫组织化学

当得到抗体后，在正常和癌变组织切片上进行免疫染色，以便确定：

- ◇ 癌变与正常组织中本发明抗原水平或
- ◇ 表达抗原的一定类型癌症的比例
- ◇ 如果其它癌症类型也表达抗原
- ◇ 在癌变组织中包括抗原的细胞的比例

组织样品制备

切开后，组织样品置于 OCH 化合物中的软木盘中，在已在液氮（-160°C）中深度冷冻的异戊烷中速冻。将样品保存在 -70°C 待用。在低温室（-20°C、-30°C）中切出 7-10μm 切片。

染色

组织切片在室温干燥 5 分钟，室温下在丙酮中固定 10 分钟，再次干燥，用 PBS 0.5% BSA 5% 血清饱和。室温下 30 分钟后，使用抗原特异性抗体进行直接或间接染色。直接染色特异性更好，但染色较弱，而间接染色染色更强，但特异性较差。

7.3 针对本发明抗原的人细胞免疫应答分析

本发明抗原的免疫关联可以通过体外引发人 T 细胞来评价。所有 T 淋巴细胞系和树突细胞来源于健康供体（优选 HLA-A2 亚型）PBMC（外周血单核细胞）。HLA-A2.1/Kb 转基因小鼠也用于筛选 HLA-A2.1 肽。

通过每周体外刺激产生和维持新发现的抗原特异性 CD8+ T 细胞系。CD8 细胞系针对抗原和抗原衍生肽的裂解活性和 γ IFN 产生使用标准方法测定。

使用两种策略产生抗原特异性 CD8+ T 细胞：基于肽的途径和基于完整基因的途径。两种途径均需要新发现抗原的全长 cDNA，以准确读框克隆在合适的传递系统中，或者用来预测 HLA 结合肽序列。

基于肽的途径

HLA - A2 结合肽序列太古 Parker 算法 (Parker, K. C., M. A. Bednarek, and J. E. Coligan. 1994 基于单独肽侧链的独立结合将潜在 HLA - A2 结合肽分级的方案。免疫学杂志, 152: 163 和 http://bimas.dcrt.nih.gov/molbio/hla_bind/) 或 Rammensee 方法 (Rammensee, Friede, Stevanovic, MHC 配体和肽基元: 表 1, 免疫遗传学 41, 178 - 228, 1995; Rammensee, Bachmann, Stevanovic: MHC 配体和肽基元。Landes Bioscience 1997, 和 <http://134.2.96.221/scriptsh/laserver.dll/home.htm>)。然后在 HLA - A2.1/Kb 转基因小鼠模型 (Vitiello et al) 中筛选肽。

10 a) 预测的结合 HLA_A0201 基因座的表位

a.1) HLA-A*0201 九聚体

位置	1	2	3	4	5	6	7	8	9	Rammens ee 计分	Parker 计 分°	SEQ ID NO
848	F	L	W	E	S	A	A	A	C		777.681	68
24	R	L	W	R	L	L	L	W	A		521.615	5
761	S	L	A	D	R	L	I	G	V	30	655.875	6
893	S	L	P	E	Q	R	V	T	I	26		7
886	K	L	C	S	G	G	I	S	L	25		8
853	A	A	A	C	P	L	C	S	V	25		9
674	S	A	L	A	N	T	V	T	L	24		10
499	F	V	F	E	T	L	C	S	V	24	976.762	11
129	E	L	P	H	G	F	A	S	L	23		12
973	L	I	F	T	S	K	K	S	L	22		13
936	K	L	E	Y	K	Y	S	K	L	22		14
903	K	T	I	D	F	W	L	K	V	22		15
860	S	V	A	D	Y	H	A	I	V	22		16
830	L	L	L	P	G	T	C	S	D	22		17
675	A	L	A	N	T	V	T	L	A	22		18
503	T	L	C	S	V	N	C	E	L	22		19
169	N	T	D	E	C	T	A	T	L	22		20
81	S	L	P	D	P	V	K	G	T	22		21
980	S	L	F	G	K	I	K	S	F	21		22
918	C	T	A	I	L	L	T	V	L	21		23
867	I	V	S	S	C	V	A	G	I	21		24
710	K	M	S	V	C	T	D	N	V	21		25
259	V	L	V	R	N	I	A	I	T	21		26
234	E	L	N	R	G	N	N	V	L	21		27
175	A	T	L	M	Y	A	V	N	L	21		28
70	V	A	V	P	H	T	P	G	L	21		29
24	R	L	W	R	L	L	L	W	A	21		30
914	S	A	G	T	C	T	A	I	L	20		31
891	G	I	S	L	P	E	Q	R	V	20		32
824	K	T	V	P	G	S	L	L	L	20		33
765	R	L	I	G	V	T	T	D	M	20		34

681	T L A G G P S F T	20		35
539	Y I I E E N T T T	20		36
264	I A I T G V A Y T	20		37
38	V T Q G T G P E L	20		38

°含有该亚序列的分子估计解离半时间

a.2) HLA A02_01 十聚体

位置	1	2	3	4	5	6	7	8	9	10	Rammens ee 计分	Parker 计 分°	SEQ ID NO
980	S	L	F	G	K	I	K	S	F	T		151.648	39
866	A	I	V	S	S	C	V	A	G	I	25		40
852	S	A	A	A	C	P	L	C	S	V	24		41
786	H	L	E	S	L	G	I	P	D	V	24		42
571	K	I	Y	S	I	N	V	T	N	V	24	246.353	43
761	S	L	A	D	R	L	I	G	V	T	23		44
626	I	L	K	A	H	Q	P	Y	G	V	23		5
485	V	M	A	D	T	E	N	K	E	V	23	350.117	46
29	L	L	W	A	G	T	A	F	Q	V	23	5691.997	47
916	G	T	C	T	A	I	L	L	T	V	22		48
778	I	T	S	P	A	E	L	F	H	L	22		49
766	L	I	G	V	T	T	D	M	T	L	22		50
428	T	L	P	T	N	M	E	T	T	V	22		51
350	L	M	Y	K	W	A	K	P	K	I	22		52
972	D	L	I	F	T	S	K	K	S	L	21		53
692	G	L	K	Y	F	H	H	F	T	L	21		54
644	G	T	K	N	N	K	I	H	S	L	21		55
465	S	D	N	D	F	M	I	L	T	L	21		56
260	L	V	R	N	I	A	I	T	G	V	21		57
77	G	L	C	T	S	L	P	D	P	V	21		58
949	T	L	K	D	C	D	L	P	A	A	20		59
795	V	I	F	F	Y	R	S	N	D	V	20		60
733	S	I	T	A	Y	V	C	Q	A	V	20		61
712	S	V	C	T	D	N	V	T	D	L	20		62
702	S	L	C	G	N	Q	G	R	K	M	20		63
432	N	M	E	T	T	V	L	S	G	I	20		64
263	N	I	A	I	T	G	V	A	Y	T	20		65
121	G	I	R	F	D	E	W	D	E	L	20		66

5 °含有该亚序列的分子估计解离半时间

HLA_A0205

位置	1	2	3	4	5	6	7	8	9	Parker 计 分°	SEQ ID NO
499	F	V	F	E	T	L	C	S	V	216	11

HLA A0203

位置	1	2	3	4	5	6	7	8	9	10	Rammens ee 计分	SEQ ID NO
846	F	H	F	L	W	E	S	A	A	A	27	67

简而言之，用佐剂配制的 HLA - A2 肽免疫转基因小鼠，不能诱导 CD8 应答的肽（定义为有效裂解肽脉冲处理的自体同源脾细胞）在人系统中进一步处理。

- 5 人树突细胞（根据 Romani 等培养）用肽脉冲处理并用于刺激 CD8 - 分选的 T 细胞（Facs）。刺激几周后，CD8 细胞系首先在肽脉冲处理的自体同源 BLCL（EBV - B 转化细胞系）上测试。为了确证肽的准确体内加工，CD8 细胞系将在 cDNA 转染的肿瘤细胞（HLA - A2 转染的 LnCaP, Skov3 或 CAMA 肿瘤细胞）上测试。

10 基于全基因的途径

- 用基因枪转染的树突细胞、逆转录病毒转导的 B7.1 转染的成纤维细胞、重组痘病毒（Kim 等）或腺病毒（Butterfield 等）感染的树突细胞引发和刺激 CD8 + T 细胞系。病毒感染细胞可非常有效地呈递抗原肽，因为抗原以高水平表达，但只能使用一次，以避免病毒 T 细胞系的过度生长。

- 15 交替刺激后，CD8 + 细胞系在如上所述的 cDNA 转染的肿瘤细胞系上测试。确定肽特异性和身份以证实免疫有效性。

参考文献

- 20 Vitiello 等（L. Sherman），实验医学杂志 1991，173：1007 - 1015。
Romani 等，实验医学杂志，1994，180：83 - 93。
Kim 等，免疫治疗，1997，20：276 - 286。
Butterfield 等，免疫学杂志，1998，161：5607 - 5613。

- 25 本申请书中引用的所有出版物和参考文献，包括但不限于专利和专利申请，在这里都全文引入作为参考，如同它们各自都特别地和个别地在此全文列出。

序列信息

SEQ ID NO:1

5 atggctgagcctgggcacagccaccatctctccgccagagtcaaggggaagaactgagagggcgcataccccgctgtggcg
gctgctgctctgggctgggaccgccttccaggtgacccagggaaacgggaccggagcttcatgcctgcaaagagtctgagt
accactatgagtacacggcgtgtgacagcacgggttccaggtggagggctgccgtgcccgcataccccgggctgtgcacc
agcctgcctgaccccgtaagggcaccgagtgctccttctcctgcaacgccggggagtttctggatatgaaggaccagtc
atgtaagccatgcgctgagggccgctactccctcggcacagggcattcggtttgatgagtgaggatgagctgccccatggct
10 ttgccagcctctcagccaaatggagctggatgacagtgctgctgagtcaccgggaactgtacttcgccaagtgggtt
ccccggggcgactacatcgctccaacacggacgaatgcacagccacactgatgtacgccgtcaacctgaagcaatctgg
cacggttaacttcgaataactatccagactccagcatcatctttgagttttcgttcagaatgaccagtgccagccca
atgcagatgactccaggtggatgaagaccacagagaaaggatgggaatccacagtgaggagctaaatcgaggcaataat
gtcctctattggagaaccacagccttctcagtatggaccaaagtacccaagcctgtgctggtgagaacattgccataac
15 aggggtggcctacacttcagaatgcttccctgcaaacctggcacgtatgcagacaagcagggctcctcttctgcaaac
tttgcccagccaactcttattcaataaaggagaaacttcttgccaccagtgtgaccctgacaaaatactcagagaaagga
tcttcttctgtaacgtgcgccagcttgacagacaaagattatctacacacacacggcctgcgatgccaacgggaga
gacacaactcatgtacaaaatgggccaagccgaaaatctgtagcggagccttgagggggcagtgaaagctgcctgcctctg
gtgtgaagaccactgcccaccctgcaaccaggttcttcaaaaaccaacaacagcacctgccagccctgccccatattggt
20 tcctactccaatggctcagactgtaccgctgcctgcagggactgaacctgctgtgggatttgatacaaaatgggtggaa
cacgctgcccacaaacatggaaacgaccgttctcagtgaggatcaacttcgagtacaagggcatgacaggctgggaggtgg
ctggatgacatttacacagctgctggagcctcagacaatgacttcatgattctcactctggttgtgccaggatttaga
cctccgcagtcgggtgatggcagacacagagaataaagaggtggccagaatcacatttgctcttgagaccctctgttctgt
gaactgtgagctctacttcatggtgggtgtgaattctaggaccaacactcctgtggagacgtggaaaggttccaaaggca
25 aacagtcctatacctacatcattgaggagaacactaccacgagcttcacctgggccttccagaggaccacttttcatgag
gcaagcaggaagtacaccaatgacgttgccaagatctactccatcaatgtaccaatggtatgaaatggcgtggcctccta
ctgccgtccctgtgccctagaagcctctgatgtgggctcctcctgcacctctgtcctgctggttactatattgaccgag
attcaggaacctgccactcctgccccctaacacaattctgaaagcccaccagccttatgggtgccaggcctgtgtgccc
tgtggtccagggaccaagaacaagaatccactctctgtgctacaatgatgcaacctctcacgcaacactccaaccag
30 gactttcaactacaacttctccgctttggcaaacaccgtcactcttgctggagggccaagcttcaactccaagggttga
aatacttccatcactttaccctcagtcctctgtggaaaccagggtaggaaaatgtctgtgtgcaccgacaatgtcactgac
ctccggattcctgagggtagtcaggggttctccaaatctatcacagcctacgtctgccaggcagtcacatccccccaga
ggtgacaggctacaaggccggggttccctcacagcctgtcagccttgctgatcgacttattggggtgacaacagatatga
ctctggatggaatcacctccccagctgaactttccacctggagtccttgggaataaccggacgtgatcttcttttatagg
35 tccaatgatgtgaccagtcctgcagttctgggagatcaaccaccatccgcgtcaggtgcagtcacagaaaaactgtccc
tggaaagtttgctgctgccaggaacgtgctcagatgggacctgtgatggctgcaacttccacttctgtgggagagcgcgg
ctgcttgcccgcctctgctcagtggtgactaccatgctatcgtcagcagctgtgtggctgggattccagaagactacttac
gtgtggcgagaaccgaagctatgctctggtggcatttctctgcctgagcagagagtcaccatctgcaaaaaccatagattt
ctggctgaaagtgggcatctctgcaggcacctgtactgccatcctgctcaccgtcttgacctgctacttttgaaaaaga
40 atcaaaaactagagtacaagtaactccaagctgggtgatgaatgctactctcaaggactgtgacctgccagcagctgacagc
tgcccatcatggaaggcgaggatgtagaggacacctcatctttaccagcaagaagtcactctttgggaagatcaaatc
atttacctccaagaggactcctgatggatttgactcagtgccgctgaagacatcctcaggaggccagacatggacctgt
gaGAGGCACTGCCTGCCTCACCTGCCTCCTCACCTTGATAGCACCTTTGCAAGCCTGCGGCGATTTGGGTGCCAGCATC
CTGCAACACCCACTGCTGGAAATCTCTTATTGTGGCCTTATCAGATGTTTGAATTTAGATCTTTTTTATAGAGTACC
45 CAAACCCTCCTTTCTGCTTGCCCTCAAACCTGCCAAATATACCCACACTTGTGTTGTAAAAaaaAAAAAAAAAAAAAAAAA

SEQ ID NO:2

MAEPGHSHLSARVRGRTEERRIPRLWRLLWAGTAFQVTQGTGPELHACKESEYHYEYACDSTGSRWRVAVPHTPGLCT
SLPDPVKGTECSFSCNAGEFLDMKDQSCPKCAEGRYSLGTGIRFDEWDELPHGFASLSANMELDDSAESTGNCTSSKWV
PRGDYIASNTDECTATLMYAVNLKQSGTVNFEYYPDSSIIIEFFVQNDQCQPNADDSRWMKTTEKWEFHSVELNRGNN
5 VLYWRTTAFSVWTKVPKVLVRNIAITGVAYTSECFPCPKPGTYADKQSSFCCKLCPANSYSNKGETSCHQCDPKYSEKG
SSSCNVRPACTDKDYFYHTACDANGETQLMYKWKPKICSEDLGAVKLPASGVKTHCPPCNPGFCKTNNSTCQPCPYG
SYSNGSDCTRCPAGTEPAVGFYKWWNTLPTNMETTVLSGINFEYKGMTGWEVAGDHIYTAAGASDNDFMILTLVVPGFR
PPQSVMAADTENKEVARITFVFETLCSVNCELYFMVGVNSRTNTPVETWKGSKGKQSYTYIIIEENTTSFTWAFORTTFHE
ASRKYTNDVAKIYSINVTNMNGVASYCRPCALEASDVGSSCTSCPAGYYIDRDSGTCHSCPNTIILKAHQPYGVQACVP
10 CGPGTKNNKIHSLCYNDCTFSRNTPTRTFNYNFSALANTVTLAGGPSFTSKGLKYFHHFTLSLCGNQGRKMSVCTDNVTD
LRIPEGESGFSKSI TAYVCQAVIIPPEVTGYKAGVSSQPVSLADRLIGVTTDMTLDGITSPAELFHLES LGIPDVIFFYR
SNDVTQSCSSGRSTTIRVRCSPQKTVPGSLLLPGTCSDGTCDCGNFHLWESAAAACPLCSVADYHAI VSSCVAGIQKTTY
VWREPKLCSGGISLPEQRVTICKTIDFWLKVGISAGTCTAILLTVLTCYFWKKNQKLEYKYSKLV MNATLKDCDLPAADS
CAIMEGEDVEDDLIFTSKKS LFGKIKSFTSKRTPDGFDSVPLKTSSGGPDMDL

15

SEQ ID NO:3

TTTTTTAATTTACAAACAAAGTGTGGGTATATTTGGCAGGTTTGAGGCAAGCAGAAAGGAGGGTTTGGGTACTCTA
20 TAAAAAAGATCTGAAATCAAACATCTGATAAGGCCACAATGAAGAGATTTCCAGCAGTGGGTGTTGCAGGATGC
TGGCACCCAAATCGCCGCACGTTGCAAAGGTGCTATGCAAAGGTGAGGAGGCAGGTGAGGCAGGCAGTGCCTCTCAC
AGGTCCATGTCTGGGCCTCCTGAGGATGTCTTCAGCGGCACTGAGTCAAATCCATCAGGAGTCCTCTTGGAGGTAA
ATGATTTGATCTTCCCAAAGAGTGACTTCTTGCTGGTAAAGATGAGGTCGTCTCTACATCCTCGCCTTCCATGAT
GGCGCAGTGT CAGCTGCTGGCAGGTCACAGTCCTTGAGAGTAGCATTCATCACCAGCTTGGAGTACTTGTACTCTA
25 GTTTTTGATTCTTTTTCCAAAAGTAGCAGGTCAAGACGGTGAGCAGGATGGCAGTACAGGTGCCTGCAGAGATGCC
CACTTTCAGCCAGAAATCTATGGTTTTGCAGATGGTACTCTCTGCTCAGGCAGAGAAATGCCACCAGAGCATAGC
TTGGGTTCTCGCCACACGTAAGTAGTCTTCTGGATCCCAGCCACACAGCTGCTGACGATAGCATGGTAG

SEQ ID NO:4

30 YHAI VSSCVAGIQKTTYVWREPKLCSGGISLPEQRVTICKTIDFWLKVGISAGTCTAILLTVLTCYFWK
KNQKLEYKYSKLV MNATLKDCDLPAADTAPSWKARM

序列表

<110> SmithKline Beecham Biologicals S.A.

5 <120> 新化合物

<130> BC45226

10 <160> 68

<170> FastSEQ for Windows Version 3.0

<210> 1

15 <211> 3280

<212> DNA

<213> 人

<400> 1

20 atggctgagc ctgggcacag ccaccatctc tccgccagag tcaggggaag aactgagagg 60
cgcatacccc ggctgtggcg gctgctgctc tgggctggga ccgccttcca ggtgacccag 120
ggaacggggac cggagcttca tgcctgcaaa gactctgagt accactatga gtacacggcg 180
tgtgacagca cgggttccag gtggaggggc gccgtgccgc ataccceggg cctgtgcacc 240
agcctgcctg accccgtcaa gggcaccgag tgctccttct cctgcaacgc cggggagttt 300
25 ctggatatga aggaccagtc atgtaagcca tgcgctgagg gccgctactc cctcggcaca 360
ggcattcggg ttgatgagtg ggatgagctg ccccatggct ttgccagcct ctccagccaac 420
atggagctgg atgacagtgc tgctgagtc accgggaact gtacttctc caagtgggtt 480
ccccggggcg actacatcgc ctccaacacg gacgaatgca cagccacact gatgtacgcc 540
gtcaacctga agcaatctgg caccgttaac ttcgaatact actatccaga ctccagcatc 600
30 atctttgagt ttttcgttca gaatgaccag tgccagccca atgcagatga ctccagggtgg 660
atgaagacca cagagaaagg atgggaattc cacagtgtgg agctaaatcg aggcaataat 720
gtcctctatt ggagaaccac agccttctca gtatggacca aagtacccaa gcctgtgctg 780
gtgagaaaca ttgccataac aggggtggcc tacacttcag aatgcttccc ctgcaaacct 840
ggcacgtatg cagacaagca gggctcctct tctgcaaac tttgccagc caactcttat 900
35 tcaaataaag gagaaacttc ttgccaccag tgtgacctg acaaatactc agagaaagga 960
tcttcttct gtaacgtgcg ccagcttgc acagacaaag attatttcta cacacacacg 1020
gcctgcgatg ccaacggaga gacacaactc atgtacaaat gggccaagcc gaaatctgt 1080
agcagggacc ttgagggggc agtgaagctg cctgcctctg gtgtgaagac cactgcccc 1140
cctgcaacc caggcttctt caaaaccaac aacagcact gccagccctg cccatattgt 1200
40 tctactcca atggctcaga ctgtaccgc tgccctgcag ggactgaacc tgctgtggga 1260
tttgaataca aatggtggaa cacgctgccc acaaactgg aaacgaccgt tctcattagg 1320
atcaacttcg agtacaagg catgacagg tgggaggtg ctggtgatca catttacaca 1380
gctgctggag cctcagacaa tgacttcatg attctcactc tggttgtgcc aggttttaga 1440
cctccgcagt cgggtgatggc agacacagag aataaagagg tggccagaat cacatttgtc 1500
45 tttgagacc tctgttctgt gaactgtgag ctctacttca tgggtgggtg gaattctagg 1560
accaacactc ctgtggagac gtggaaaggt tccaaaggca aacagtccta tacctacatc 1620
attgaggaga aactaccac gagcttcacc tgggccttcc agaggaccac ttttcatgag 1680
gcaagcagga agtacaccaa tgacgttgc aagatctact ccatcaatgt caccaatgtt 1740
atgaatggcg tggcctccta ctgccgtccc tgtgccctag aagcctctga tgtgggctcc 1800
50 tcttgcaact cttgtctgc tggttactat attgaccgag attcaggaac ctgcccactc 1860
tgcccccta acacaattct gaaagcccac cagccttatg gtgtccaggc ctgtgtgcc 1920
tgtggtccag ggaccaagaa caacaagatc cactctctgt gctacaatga ttgcacctc 1980
tcaagcaaca ctccaaccag gactttcaac tacaacttct ccgctttggc aaacaccgctc 2040
actcttctg gagggccaag ctctacttcc aaagggttga aatacttcca tcaacttacc 2100
55 ctcaactctc gtggaacca gggtaggaaa atgtctgtgt gcaccgaca tgctactgac 2160
ctccggatc ctgaggggtg gtcaggggtc tccaaatcta tcacagccta cgtctgccag 2220
gcagtcata tccccccaga ggtgacaggc tacaaggccg gggtttctc acagcctgtc 2280
agccttctg atcgacttat tggggtgaca acagatatga ctctggatgg aatcacctcc 2340
ccagctgaac ttttccact ggagtccttg ggaataccgg acgtgatctt cttttatagg 2400
60 tccaatgatg tgaccagtc ctgacgttct gggagatcaa ccaccatccg cgtcagggtc 2460
agtccacaga aaactgtccc tggaaagttg ctgctgccag gaacgtgctc agatgggacc 2520
tgtgatggct gcaacttcca ctctctgtgg gagagcgcgg ctgcttgcct gctctgctca 2580
gtggctgact accatgctat cgtcagcagc tgtgtggctg ggatccagaa gactacttac 2640
gtgtggcgag aaccgaagct atgctctggg ggcatttctc tgcctgagca gagagtcacc 2700
65 atctgcaaaa ccatagattt ctggctgaaa gtggcatct ctgcaggcac ctgtactgcc 2760

```

atcctgctca ccgtcttgac ctgctacttt tggaaaaaga atcaaaaact agagtacaag      2820
tactccaagc tggatgatgaa tgctactctc aaggactgtg acctgccagc agctgacagc      2880
tgcgccatca tggaaggcga ggatgtagag gacgacctca tctttaccag caagaagtca      2940
ctctttggga agatcaaatc atttacctcc aagaggactc ctgatggatt tgactcagtg      3000
5  ccgctgaaga catcctcagg aggcccagac atggacctgt gagaggcact gcttgacctca      3060
cctgctctct caccttgcat agcacctttg caagcctgct gcgatttggg tgccagcatc      3120
ctgcaacacc cactgctgga aatctcttca ttgtggcctt atcagatggt tgaatttcag      3180
atcttttttt atagagtacc caaacctctc tttctgcttg cctcaaacct gccaaatata      3240
cccacacttt gtttgtaaaa aaaaaaaaaa aaaaaaaaaa      3280

10      <210> 2
        <211> 1013
        <212> PRT
        <213> 人

15      <400> 2
Met Ala Glu Pro Gly His Ser His His Leu Ser Ala Arg Val Arg Gly
 1      5      10      15
Arg Thr Glu Arg Arg Ile Pro Arg Leu Trp Arg Leu Leu Leu Trp Ala
20      20      25      30
Gly Thr Ala Phe Gln Val Thr Gln Gly Thr Gly Pro Glu Leu His Ala
 35      40      45
Cys Lys Glu Ser Glu Tyr His Tyr Glu Tyr Thr Ala Cys Asp Ser Thr
50      55      60
25 Gly Ser Arg Trp Arg Val Ala Val Pro His Thr Pro Gly Leu Cys Thr
 65      70      75      80
Ser Leu Pro Asp Pro Val Lys Gly Thr Glu Cys Ser Phe Ser Cys Asn
 85      90      95
Ala Gly Glu Phe Leu Asp Met Lys Asp Gln Ser Cys Lys Pro Cys Ala
30      100      105      110
Glu Gly Arg Tyr Ser Leu Gly Thr Gly Ile Arg Phe Asp Glu Trp Asp
 115      120      125
Glu Leu Pro His Gly Phe Ala Ser Leu Ser Ala Asn Met Glu Leu Asp
130      135      140
35 Asp Ser Ala Ala Glu Ser Thr Gly Asn Cys Thr Ser Ser Lys Trp Val
 145      150      155      160
Pro Arg Gly Asp Tyr Ile Ala Ser Asn Thr Asp Glu Cys Thr Ala Thr
 165      170      175
Leu Met Tyr Ala Val Asn Leu Lys Gln Ser Gly Thr Val Asn Phe Glu
40      180      185      190
Tyr Tyr Tyr Pro Asp Ser Ser Ile Ile Phe Glu Phe Phe Val Gln Asn
 195      200      205
Asp Gln Cys Gln Pro Asn Ala Asp Asp Ser Arg Trp Met Lys Thr Thr
210      215      220
45 Glu Lys Gly Trp Glu Phe His Ser Val Glu Leu Asn Arg Gly Asn Asn
 225      230      235      240
Val Leu Tyr Trp Arg Thr Thr Ala Phe Ser Val Trp Thr Lys Val Pro
 245      250      255
Lys Pro Val Leu Val Arg Asn Ile Ala Ile Thr Gly Val Ala Tyr Thr
50      260      265      270
Ser Glu Cys Phe Pro Cys Lys Pro Gly Thr Tyr Ala Asp Lys Gln Gly
 275      280      285
Ser Ser Phe Cys Lys Leu Cys Pro Ala Asn Ser Tyr Ser Asn Lys Gly
290      295      300
55 Glu Thr Ser Cys His Gln Cys Asp Pro Asp Lys Tyr Ser Glu Lys Gly
 305      310      315      320
Ser Ser Ser Cys Asn Val Arg Pro Ala Cys Thr Asp Lys Asp Tyr Phe
 325      330      335
Tyr Thr His Thr Ala Cys Asp Ala Asn Gly Glu Thr Gln Leu Met Tyr
60      340      345      350
Lys Trp Ala Lys Pro Lys Ile Cys Ser Glu Asp Leu Glu Gly Ala Val
 355      360      365
Lys Leu Pro Ala Ser Gly Val Lys Thr His Cys Pro Pro Cys Asn Pro
370      375      380
65 Gly Phe Phe Lys Thr Asn Asn Ser Thr Cys Gln Pro Cys Pro Tyr Gly
 385      390      395      400

```

	Ser	Tyr	Ser	Asn	Gly	Ser	Asp	Cys	Thr	Arg	Cys	Pro	Ala	Gly	Thr	Glu
					405					410					415	
	Pro	Ala	Val	Gly	Phe	Glu	Tyr	Lys	Trp	Trp	Asn	Thr	Leu	Pro	Thr	Asn
				420					425					430		
5	Met	Glu	Thr	Thr	Val	Leu	Ser	Gly	Ile	Asn	Phe	Glu	Tyr	Lys	Gly	Met
				435				440					445			
	Thr	Gly	Trp	Glu	Val	Ala	Gly	Asp	His	Ile	Tyr	Thr	Ala	Ala	Gly	Ala
				450			455					460				
	Ser	Asp	Asn	Asp	Phe	Met	Ile	Leu	Thr	Leu	Val	Val	Pro	Gly	Phe	Arg
10	465					470					475					480
	Pro	Pro	Gln	Ser	Val	Met	Ala	Asp	Thr	Glu	Asn	Lys	Glu	Val	Ala	Arg
					485					490					495	
	Ile	Thr	Phe	Val	Phe	Glu	Thr	Leu	Cys	Ser	Val	Asn	Cys	Glu	Leu	Tyr
				500					505					510		
15	Phe	Met	Val	Gly	Val	Asn	Ser	Arg	Thr	Asn	Thr	Pro	Val	Glu	Thr	Trp
				515				520					525			
	Lys	Gly	Ser	Lys	Gly	Lys	Gln	Ser	Tyr	Thr	Tyr	Ile	Ile	Glu	Glu	Asn
				530			535					540				
20	Thr	Thr	Thr	Ser	Phe	Thr	Trp	Ala	Phe	Gln	Arg	Thr	Thr	Phe	His	Glu
	545					550					555					560
	Ala	Ser	Arg	Lys	Tyr	Thr	Asn	Asp	Val	Ala	Lys	Ile	Tyr	Ser	Ile	Asn
					565					570					575	
	Val	Thr	Asn	Val	Met	Asn	Gly	Val	Ala	Ser	Tyr	Cys	Arg	Pro	Cys	Ala
				580					585					590		
25	Leu	Glu	Ala	Ser	Asp	Val	Gly	Ser	Ser	Cys	Thr	Ser	Cys	Pro	Ala	Gly
				595				600					605			
	Tyr	Tyr	Ile	Asp	Arg	Asp	Ser	Gly	Thr	Cys	His	Ser	Cys	Pro	Pro	Asn
				610			615					620				
30	Thr	Ile	Leu	Lys	Ala	His	Gln	Pro	Tyr	Gly	Val	Gln	Ala	Cys	Val	Pro
	625					630					635					640
	Cys	Gly	Pro	Gly	Thr	Lys	Asn	Asn	Lys	Ile	His	Ser	Leu	Cys	Tyr	Asn
					645					650					655	
	Asp	Cys	Thr	Phe	Ser	Arg	Asn	Thr	Pro	Thr	Arg	Thr	Phe	Asn	Tyr	Asn
				660					665					670		
35	Phe	Ser	Ala	Leu	Ala	Asn	Thr	Val	Thr	Leu	Ala	Gly	Gly	Pro	Ser	Phe
				675				680						685		
	Thr	Ser	Lys	Gly	Leu	Lys	Tyr	Phe	His	His	Phe	Thr	Leu	Ser	Leu	Cys
				690			695					700				
	Gly	Asn	Gln	Gly	Arg	Lys	Met	Ser	Val	Cys	Thr	Asp	Asn	Val	Thr	Asp
40	705					710					715					720
	Leu	Arg	Ile	Pro	Glu	Gly	Glu	Ser	Gly	Phe	Ser	Lys	Ser	Ile	Thr	Ala
					725					730					735	
	Tyr	Val	Cys	Gln	Ala	Val	Ile	Ile	Pro	Pro	Glu	Val	Thr	Gly	Tyr	Lys
				740					745					750		
45	Ala	Gly	Val	Ser	Ser	Gln	Pro	Val	Ser	Leu	Ala	Asp	Arg	Leu	Ile	Gly
				755				760					765			
	Val	Thr	Thr	Asp	Met	Thr	Leu	Asp	Gly	Ile	Thr	Ser	Pro	Ala	Glu	Leu
				770			775					780				
	Phe	His	Leu	Glu	Ser	Leu	Gly	Ile	Pro	Asp	Val	Ile	Phe	Phe	Tyr	Arg
50	785					790					795					800
	Ser	Asn	Asp	Val	Thr	Gln	Ser	Cys	Ser	Ser	Gly	Arg	Ser	Thr	Thr	Ile
					805					810					815	
	Arg	Val	Arg	Cys	Ser	Pro	Gln	Lys	Thr	Val	Pro	Gly	Ser	Leu	Leu	Leu
					820				825						830	
55	Pro	Gly	Thr	Cys	Ser	Asp	Gly	Thr	Cys	Asp	Gly	Cys	Asn	Phe	His	Phe
				835				840					845			
	Leu	Trp	Glu	Ser	Ala	Ala	Ala	Cys	Pro	Leu	Cys	Ser	Val	Ala	Asp	Tyr
							855					860				
60	His	Ala	Ile	Val	Ser	Ser	Cys	Val	Ala	Gly	Ile	Gln	Lys	Thr	Thr	Tyr
	865					870					875					880
	Val	Trp	Arg	Glu	Pro	Lys	Leu	Cys	Ser	Gly	Gly	Ile	Ser	Leu	Pro	Glu
					885					890					895	
	Gln	Arg	Val	Thr	Ile	Cys	Lys	Thr	Ile	Asp	Phe	Trp	Leu	Lys	Val	Gly
				900					905					910		
65	Ile	Ser	Ala	Gly	Thr	Cys	Thr	Ala	Ile	Leu	Leu	Thr	Val	Leu	Thr	Cys
				915				920						925		

Tyr Phe Trp Lys Lys Asn Gln Lys Leu Glu Tyr Lys Tyr Ser Lys Leu
 930 935 940
 Val Met Asn Ala Thr Leu Lys Asp Cys Asp Leu Pro Ala Ala Asp Ser
 945 950 955 960
 5 Cys Ala Ile Met Glu Gly Glu Asp Val Glu Asp Asp Leu Ile Phe Thr
 965 970 975
 Ser Lys Lys Ser Leu Phe Gly Lys Ile Lys Ser Phe Thr Ser Lys Arg
 980 985 990
 10 Thr Pro Asp Gly Phe Asp Ser Val Pro Leu Lys Thr Ser Ser Gly Gly
 995 1000 1005
 Pro Asp Met Asp Leu
 1010

 <210> 3
 15 <211> 677
 <212> DNA
 <213> 人

 <400> 3
 20 ttttttaatt tacaacaaa gtgtgggtat atttggcagg tttgaggcaa gcagaaagga 60
 gggtttgggt actctataaa aaaagatctg aaattcaaac atctgataag gccacaatga 120
 agagatttcc agcagtgggt gttgcaggat gctggcacc aaatcgccgc acgttgcaaa 180
 ggtgctatgc aagtgagga ggcaggtgag gcaggcagt cctctcacag gtccatgtct 240
 25 gggcctcctg aggatgtctt cagcggcact gagtcaaadc catcaggagt cctcttgag 300
 gtaaatgatt tgatcttccc aaagagtgac ttcttgctgg taaagatgag gtcgtcctct 360
 acatcctcgc cttccatgat ggcgcagtgt cagctgctgg caggtcacag tccttgagag 420
 tagcattcat caccagcttg gactacttgt actctagttt ttgattcttt ttccaaaagt 480
 agcaggtcaa gacgtgagc aggatggcag tacaggtgcc tgcaagatg cccactttca 540
 gccagaaatc tatggttttg cagatggtga ctctctgctc aggcagagaa atgccaccag 600
 30 agcatagctt gggttctcgc cacacgtaag tagtcttctg gatcccagcc acacagctgc 660
 tgacgatagc atggtag 677

 <210> 4
 35 <211> 105
 <212> PRT
 <213> 人

 <400> 4
 40 Tyr His Ala Ile Val Ser Ser Cys Val Ala Gly Ile Gln Lys Thr Thr
 1 5 10 15
 Tyr Val Trp Arg Glu Pro Lys Leu Cys Ser Gly Gly Ile Ser Leu Pro
 20 25 30
 Glu Gln Arg Val Thr Ile Cys Lys Thr Ile Asp Phe Trp Leu Lys Val
 35 40 45
 45 Gly Ile Ser Ala Gly Thr Cys Thr Ala Ile Leu Leu Thr Val Leu Thr
 50 55 60
 Cys Tyr Phe Trp Lys Lys Asn Gln Lys Leu Glu Tyr Lys Tyr Ser Lys
 65 70 75 80
 50 Leu Val Met Asn Ala Thr Leu Lys Asp Cys Asp Leu Pro Ala Ala Asp
 85 90 95
 Thr Ala Pro Ser Trp Lys Ala Arg Met
 100 105

 <210> 5
 55 <211> 9
 <212> PRT
 <213> 人工序列

 <400> 5
 60 Arg Leu Trp Arg Leu Leu Leu Trp Ala
 1 5

 <210> 6
 65 <211> 9
 <212> PRT
 <213> 人工序列

<400> 6
 Ser Leu Ala Asp Arg Leu Ile Gly Val
 1 5
 5
 <210> 7
 <211> 9
 <212> PRT
 <213> 人工序列
 10
 <400> 7
 Ser Leu Pro Glu Gln Arg Val Thr Ile
 1 5
 15
 <210> 8
 <211> 9
 <212> PRT
 <213> 人工序列
 20
 <400> 8
 Lys Leu Cys Ser Gly Gly Ile Ser Leu
 1 5
 25
 <210> 9
 <211> 9
 <212> PRT
 <213> 人工序列
 30
 <400> 9
 Ala Ala Ala Cys Pro Leu Cys Ser Val
 1 5
 35
 <210> 10
 <211> 9
 <212> PRT
 <213> 人工序列
 40
 <400> 10
 Ser Ala Leu Ala Asn Thr Val Thr Leu
 1 5
 45
 <210> 11
 <211> 9
 <212> PRT
 <213> 人工序列
 50
 <400> 11
 Phe Val Phe Glu Thr Leu Cys Ser Val
 1 5
 55
 <210> 12
 <211> 9
 <212> PRT
 <213> 人工序列
 60
 <400> 12
 Glu Leu Pro His Gly Phe Ala Ser Leu
 1 5
 65
 <210> 13
 <211> 9
 <212> PRT
 <213> 人工序列
 <400> 13

Leu Ile Phe Thr Ser Lys Lys Ser Leu
 1 5

5 <210> 14
 <211> 9
 <212> PRT
 <213> 人工序列

<400> 14
 10 Lys Leu Glu Tyr Lys Tyr Ser Lys Leu
 1 5

15 <210> 15
 <211> 9
 <212> PRT
 <213> 人工序列

<400> 15
 20 Lys Thr Ile Asp Phe Trp Leu Lys Val
 1 5

25 <210> 16
 <211> 9
 <212> PRT
 <213> 人工序列

<400> 16
 30 Ser Val Ala Asp Tyr His Ala Ile Val
 1 5

35 <210> 17
 <211> 9
 <212> PRT
 <213> 人工序列

<400> 17
 Leu Leu Leu Pro Gly Thr Cys Ser Asp
 1 5

40 <210> 18
 <211> 9
 <212> PRT
 <213> 人工序列

45 <400> 18
 Ala Leu Ala Asn Thr Val Thr Leu Ala
 1 5

50 <210> 19
 <211> 9
 <212> PRT
 <213> 人工序列

<400> 19
 55 Thr Leu Cys Ser Val Asn Cys Glu Leu
 1 5

60 <210> 20
 <211> 9
 <212> PRT
 <213> 人工序列

<400> 20
 65 Asn Thr Asp Glu Cys Thr Ala Thr Leu
 1 5

<210> 21
 <211> 9
 <212> PRT
 5 <213> 人工序列

<400> 21
 Ser Leu Pro Asp Pro Val Lys Gly Thr
 1 5

10 <210> 22
 <211> 9
 <212> PRT
 <213> 人工序列

15 <400> 22
 Ser Leu Phe Gly Lys Ile Lys Ser Phe
 1 5

20 <210> 23
 <211> 9
 <212> PRT
 <213> 人工序列

25 <400> 23
 Cys Thr Ala Ile Leu Leu Thr Val Leu
 1 5

30 <210> 24
 <211> 9
 <212> PRT
 <213> 人工序列

35 <400> 24
 Ile Val Ser Ser Cys Val Ala Gly Ile
 1 5

40 <210> 25
 <211> 9
 <212> PRT
 <213> 人工序列

45 <400> 25
 Lys Met Ser Val Cys Thr Asp Asn Val
 1 5

50 <210> 26
 <211> 9
 <212> PRT
 <213> 人工序列

55 <400> 26
 Val Leu Val Arg Asn Ile Ala Ile Thr
 1 5

60 <210> 27
 <211> 9
 <212> PRT
 <213> 人工序列

60 <400> 27
 Glu Leu Asn Arg Gly Asn Asn Val Leu
 1 5

65 <210> 28

<211> 9
 <212> PRT
 <213> 人工序列

5 <400> 28
 Ala Thr Leu Met Tyr Ala Val Asn Leu
 1 5

10 <210> 29
 <211> 9
 <212> PRT
 <213> 人工序列

15 <400> 29
 Val Ala Val Pro His Thr Pro Gly Leu
 1 5

20 <210> 30
 <211> 9
 <212> PRT
 <213> 人工序列

25 <400> 30
 Arg Leu Trp Arg Leu Leu Leu Trp Ala
 1 5

30 <210> 31
 <211> 9
 <212> PRT
 <213> 人工序列

35 <400> 31
 Ser Ala Gly Thr Cys Thr Ala Ile Leu
 1 5

40 <210> 32
 <211> 9
 <212> PRT
 <213> 人工序列

45 <400> 32
 Gly Ile Ser Leu Pro Glu Gln Arg Val
 1 5

50 <210> 33
 <211> 9
 <212> PRT
 <213> 人工序列

55 <400> 33
 Lys Thr Val Pro Gly Ser Leu Leu Leu
 1 5

60 <210> 34
 <211> 9
 <212> PRT
 <213> 人工序列

65 <400> 34
 Arg Leu Ile Gly Val Thr Thr Asp Met
 1 5

70 <210> 35
 <211> 9
 <212> PRT

<213> 人工序列
 <400> 35
 5 Thr Leu Ala Gly Gly Pro Ser Phe Thr
 1 5
 <210> 36
 <211> 9
 <212> PRT
 10 <213> 人工序列
 <400> 36
 15 Tyr Ile Ile Glu Glu Asn Thr Thr Thr
 1 5
 <210> 37
 <211> 9
 <212> PRT
 20 <213> 人工序列
 <400> 37
 Ile Ala Ile Thr Gly Val Ala Tyr Thr
 1 5
 25 <210> 38
 <211> 9
 <212> PRT
 <213> 人工序列
 30 <400> 38
 Val Thr Gln Gly Thr Gly Pro Glu Leu
 1 5
 35 <210> 39
 <211> 10
 <212> PRT
 <213> 人工序列
 40 <400> 39
 Ser Leu Phe Gly Lys Ile Lys Ser Phe Thr
 1 5 10
 45 <210> 40
 <211> 10
 <212> PRT
 <213> 人工序列
 50 <400> 40
 Ala Ile Val Ser Ser Cys Val Ala Gly Ile
 1 5 10
 55 <210> 41
 <211> 10
 <212> PRT
 <213> 人工序列
 60 <400> 41
 Ser Ala Ala Ala Cys Pro Leu Cys Ser Val
 1 5 10
 65 <210> 42
 <211> 10
 <212> PRT
 <213> 人工序列

<400> 42
 His Leu Glu Ser Leu Gly Ile Pro Asp Val
 1 5 10

5
 <210> 43
 <211> 10
 <212> PRT
 <213> 人工序列

10
 <400> 43
 Lys Ile Tyr Ser Ile Asn Val Thr Asn Val
 1 5 10

15
 <210> 44
 <211> 10
 <212> PRT
 <213> 人工序列

20
 <400> 44
 Ser Leu Ala Asp Arg Leu Ile Gly Val Thr
 1 5 10

25
 <210> 45
 <211> 10
 <212> PRT
 <213> 人工序列

30
 <400> 45
 Ile Leu Lys Ala His Gln Pro Tyr Gly Val
 1 5 10

35
 <210> 46
 <211> 10
 <212> PRT
 <213> 人工序列

40
 <400> 46
 Val Met Ala Asp Thr Glu Asn Lys Glu Val
 1 5 10

45
 <210> 47
 <211> 10
 <212> PRT
 <213> 人工序列

50
 <400> 47
 Leu Leu Trp Ala Gly Thr Ala Phe Gln Val
 1 5 10

55
 <210> 48
 <211> 10
 <212> PRT
 <213> 人工序列

60
 <400> 48
 Gly Thr Cys Thr Ala Ile Leu Leu Thr Val
 1 5 10

65
 <210> 49
 <211> 10
 <212> PRT
 <213> 人工序列

65
 <400> 49
 Ile Thr Ser Pro Ala Glu Leu Phe His Leu

<210> 57
 <211> 10
 <212> PRT
 <213> 人工序列
 5
 <400> 57
 Leu Val Arg Asn Ile Ala Ile Thr Gly Val
 1 5 10
 10
 <210> 58
 <211> 10
 <212> PRT
 <213> 人工序列
 15
 <400> 58
 Gly Leu Cys Thr Ser Leu Pro Asp Pro Val
 1 5 10
 20
 <210> 59
 <211> 10
 <212> PRT
 <213> 人工序列
 25
 <400> 59
 Thr Leu Lys Asp Cys Asp Leu Pro Ala Ala
 1 5 10
 30
 <210> 60
 <211> 10
 <212> PRT
 <213> 人工序列
 35
 <400> 60
 Val Ile Phe Phe Tyr Arg Ser Asn Asp Val
 1 5 10
 40
 <210> 61
 <211> 10
 <212> PRT
 <213> 人工序列
 45
 <400> 61
 Ser Ile Thr Ala Tyr Val Cys Gln Ala Val
 1 5 10
 50
 <210> 62
 <211> 10
 <212> PRT
 <213> 人工序列
 50
 <400> 62
 Ser Val Cys Thr Asp Asn Val Thr Asp Leu
 1 5 10
 55
 <210> 63
 <211> 10
 <212> PRT
 <213> 人工序列
 60
 <400> 63
 Ser Leu Cys Gly Asn Gln Gly Arg Lys Met
 1 5 10
 65
 <210> 64
 <211> 10

```

    <212> PRT
    <213> 人工序列

    <400> 64
5   Asn Met Glu Thr Thr Val Leu Ser Gly Ile
    1                               5           10

    <210> 65
    <211> 10
10  <212> PRT
    <213> 人工序列

    <400> 65
15  Asn Ile Ala Ile Thr Gly Val Ala Tyr Thr
    1                               5           10

    <210> 66
    <211> 10
    <212> PRT
20  <213> 人工序列

    <400> 66
    Gly Ile Arg Phe Asp Glu Trp Asp Glu Leu
    1                               5           10
25

    <210> 67
    <211> 10
    <212> PRT
    <213> 人工序列
30

    <400> 67
    Phe His Phe Leu Trp Glu Ser Ala Ala Ala
    1                               5           10
35

    <210> 68
    <211> 9
    <212> PRT
    <213> 人工序列

    <400> 68
40  Phe Leu Trp Glu Ser Ala Ala Ala Cys
    1                               5

```

说明书附图

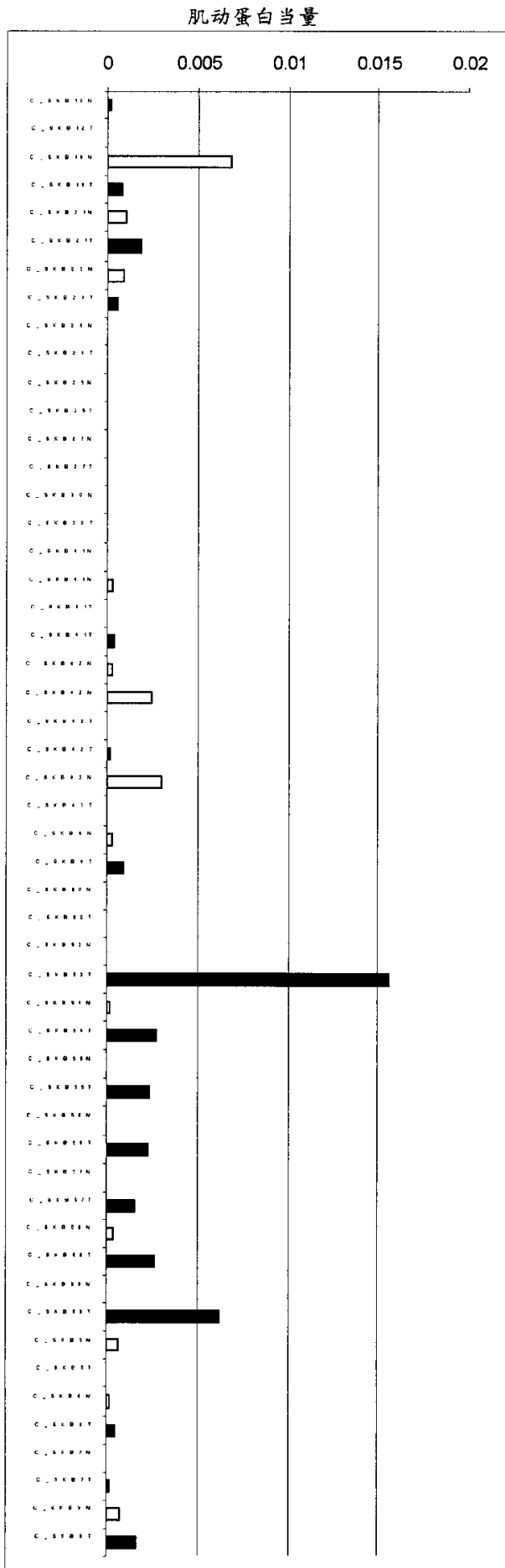


图 1: 匹配正常和癌变结肠样品的实时 PCR 数据

肌动蛋白当量

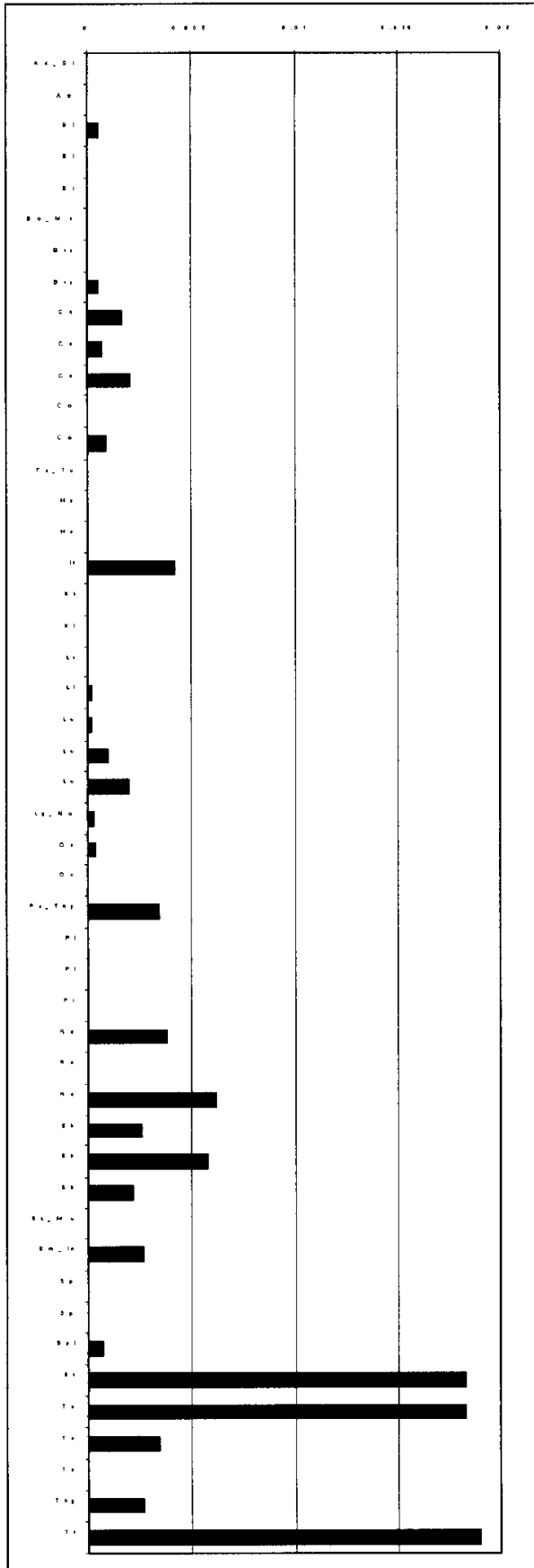


图2: 正常组织中 CASB619 的实时 PCR 数据

图例: Ad_Gl: 肾上腺; Ao: 大动脉; Bl: 膀胱, Bo_Ma: 骨髓; Bra: 脑; Ce: 子宫颈; Co: 结肠; Fa_Tu: 输卵管; He: 心脏; Il: 回肠; Ki: 肾脏; Li: 肝脏; Lu: 肺; Ly_No: 淋巴结; Oe: 食管; Ov: 卵巢; a_Thy: 甲状旁腺; Pl: 胎盘; Pr: 前列腺; Re: 直肠; Sk: 皮肤; Sk_Mu: 骨骼; Sp: 脾脏; St: 胃; Te: 睾丸; Thy: 甲状腺; Tr: 气管

专利名称(译)	新化合物		
公开(公告)号	CN1351661A	公开(公告)日	2002-05-29
申请号	CN00805597.1	申请日	2000-03-20
[标]申请(专利权)人(译)	史密斯克莱恩比彻姆生物有限公司		
申请(专利权)人(译)	史密斯克莱恩比彻姆生物有限公司		
当前申请(专利权)人(译)	史密斯克莱恩比彻姆生物有限公司		
[标]发明人	CEM布鲁克 J P卡萨特 T科切 C维纳尔斯丫德巴索尔斯		
发明人	C·E·M·布鲁克 J· - P·卡萨特 T·科切 C·维纳尔斯丫德巴索尔斯		
IPC分类号	G01N33/50 A61K31/7088 A61K38/00 A61K39/00 A61K39/39 A61K45/00 A61K48/00 A61P35/00 C07K14/47 C07K14/705 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/574 G01N37/00 A61K38/17 G01N33 /68 C07K16/30 C12N15/11		
CPC分类号	A01K2217/05 A61K38/00 C07K14/705		
优先权	1999007113 1999-03-26 GB 1999022858 1999-09-25 GB		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了CASB619多肽和多核苷酸以及通过重组技术生产这种多肽的方法。本发明还公开了在诊断中应用CASB619多肽和多核苷酸的方法, 预防和治疗癌症, 特别是卵巢癌和结肠癌, 自身免疫病和相关病症的疫苗。

