

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200780018810.8

[51] Int. Cl.

C12N 15/00 (2006.01)
A61K 31/711 (2006.01)
A61K 35/26 (2006.01)
A61K 38/00 (2006.01)
A61K 39/00 (2006.01)
A61K 45/00 (2006.01)

[43] 公开日 2009年6月3日

[11] 公开号 CN 101448938A

[51] Int. Cl. (续)

A61K 48/00 (2006.01)
A61P 35/00 (2006.01)
A61P 35/02 (2006.01)
A61P 43/00 (2006.01)
C07K 14/82 (2006.01)
C07K 16/32 (2006.01)
C12N 5/06 (2006.01)
C12N 15/09 (2006.01)
C12Q 1/68 (2006.01)
G01N 33/53 (2006.01)

[22] 申请日 2007.3.27

[21] 申请号 200780018810.8

[30] 优先权

[32] 2006. 3.28 [33] JP [31] 086820/2006

[86] 国际申请 PCT/JP2007/056279 2007.3.27

[87] 国际公布 WO2007/119515 日 2007.10.25

[85] 进入国家阶段日期 2008.11.24

[71] 申请人 佐藤升志

地址 日本北海道

共同申请人 大日本住友制药株式会社

[72] 发明人 佐藤升志 鸟越俊彦 广桥良彦
原田健司

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 梁 谋 黄可峻

权利要求书 5 页 说明书 86 页 序列表 108 页
附图 12 页

[54] 发明名称

新肿瘤抗原肽

[57] 摘要

本发明涉及新肿瘤抗原蛋白、肿瘤抗原肽以及这些物质在癌症免疫领域中的用途。更具体地说，本发明涉及含有来源于 Lengsin、BJ - TSA - 9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分肽、能够结合 HLA 抗原并因此被 CTL 识别的肽，以及包含上述肽和在药学上可接受的载体的药物组合物等。

1. 一种含有来源于 Lengsin (含谷氨酸-氨连接酶(谷氨酰胺合酶)结构域 1, 也称为 GLULD1)、BJ-TSA-9 (假定蛋白 MGC14128)、C20orf42 (URP1, 也称为 Kindlerin)、BUB1、C10orf3 或 HIFPH3 (egl9 同系物 3, 也称为 EGLN3)的部分肽、能够结合 HLA 抗原并被 CTL 识别的肽。

2. 权利要求 1 的肽, 其中所述 HLA 抗原为 HLA-A24 或 HLA-A2 抗原。

3. 权利要求 2 的肽, 所述肽含有 SEQ ID NO: 13-201 中任一项的氨基酸序列。

4. 一种肽, 所述肽含有一种氨基酸序列, 该氨基酸序列与 SEQ ID NO: 13-31、42-49、59-78、89-117、158-165、176-183 和 195-201 中任一项的氨基酸序列相同, 只是 2 位的氨基酸被酪氨酸、苯丙氨酸、甲硫氨酸或色氨酸取代, 和/或 C 末端氨基酸被苯丙氨酸、亮氨酸、异亮氨酸、色氨酸或甲硫氨酸取代, 所述肽能够结合 HLA-A24 抗原, 并被 CTL 识别。

5. 一种肽, 所述肽含有一种氨基酸序列, 该氨基酸序列与 SEQ ID NO: 32-41、50-58、79-88、118-157、166-175 和 184-194 中任一项的氨基酸序列相同, 只是 2 位的氨基酸被亮氨酸、甲硫氨酸、缬氨酸、异亮氨酸或谷氨酰胺取代, 和/或 C 末端氨基酸被缬氨酸或亮氨酸取代, 所述肽能够结合 HLA-A2 抗原, 并被 CTL 识别。

6. 一种表位肽, 所述表位肽含有权利要求 1-5 中任一项的肽。

7. 一种含有权利要求 1-6 中任一项的肽和在药学上可接受的载体的药物组合物。

8. 一种含有编码权利要求 1-6 中任一项的肽的多核苷酸的核酸。

9. 一种含有权利要求 8 的核酸和在药学上可接受的载体的药物组合物。

10. 一种含有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 和在药学上可接受的载体的药物组合物。

11. 权利要求 10 的药物组合物, 其中 Lengsin 含有 SEQ ID NO: 2 的氨基酸序列。

12. 权利要求 10 的药物组合物, 其中 BJ-TSA-9 含有 SEQ ID NO: 4 的氨基酸序列。

13. 权利要求 10 的药物组合物, 其中 C20orf42 含有 SEQ ID NO: 6 的氨基酸序列。

14. 权利要求 10 的药物组合物, 其中 BUB1 含有 SEQ ID NO: 8 的氨基酸序列。

15. 权利要求 10 的药物组合物, 其中 C10orf3 含有 SEQ ID NO: 10 的氨基酸序列。

16. 权利要求 10 的药物组合物, 其中 HIFPH3 含有 SEQ ID NO: 12 的氨基酸序列。

17. 一种含有核酸和在药学上可接受的载体的药物组合物, 所述核酸包含编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸。

18. 权利要求 17 的药物组合物, 其中编码 Lengsin 的多核苷酸包含 SEQ ID NO: 1 的碱基序列, 或编码 SEQ ID NO: 2 的氨基酸序列。

19. 权利要求 17 的药物组合物, 其中编码 BJ-TSA-9 的多核苷酸包含 SEQ ID NO: 3 的碱基序列, 或编码 SEQ ID NO: 4 的氨基酸序列。

20. 权利要求 17 的药物组合物, 其中编码 C20orf42 的多核苷酸包含 SEQ ID NO: 5 的碱基序列, 或编码 SEQ ID NO: 6 的氨基酸序列。

21. 权利要求 17 的药物组合物, 其中编码 BUB1 的多核苷酸包含 SEQ ID NO: 7 的碱基序列, 或编码 SEQ ID NO: 8 的氨基酸序列。

22. 权利要求 17 的药物组合物, 其中编码 C10orf3 的多核苷酸包含 SEQ ID NO: 9 的碱基序列, 或编码 SEQ ID NO: 10 的氨基酸序列。

23. 权利要求 17 的药物组合物, 其中编码 HIFPH3 的多核苷酸包

含 SEQ ID NO: 11 的碱基序列, 或编码 SEQ ID NO: 12 的氨基酸序列。

24. 一种制备抗原提呈细胞的方法, 其中使具有抗原提呈能力的细胞在体外与以下的任一种接触:

(a) 权利要求 1-6 中任一项的肽,

(b) 权利要求 8 的核酸,

(c) Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3,
和

(d) 含有编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸的核酸。

25. 一种通过权利要求 24 的方法制备的抗原提呈细胞。

26. 一种含有权利要求 25 的抗原提呈细胞和在药学上可接受的载体的药物组合物。

27. 一种诱导 CTL 的方法, 其中使外周血淋巴细胞在体外与以下的任一种接触:

(a) 权利要求 1-6 中任一项的肽,

(b) 权利要求 8 的核酸,

(c) Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3,
和

(d) 含有编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸的核酸。

28. 通过权利要求 27 的方法诱导的 CTL。

29. 一种含有权利要求 28 的 CTL 和在药学上可接受的载体的药物组合物。

30. 权利要求 7、9-23、26 或 29 中任一项的药物组合物, 所述药物组合物用作 CTL 的诱导物。

31. 权利要求 7、9-23、26 或 29 中任一项的药物组合物, 所述药物组合物用作癌症疫苗。

32. 一种特异性结合权利要求 1-5 中任一项的肽的抗体。

33. 一种含有权利要求 1-5 中任一项的肽和 HLA 抗原的 HLA 单体、HLA 二聚体、HLA 四聚体或 HLA 五聚体。

34. 一种用于检测对肿瘤抗原肽特异性的 CTL 的试剂, 所述肿瘤抗原肽来源于 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3, 所述试剂含有作为组分的权利要求 33 的 HLA 单体、HLA 二聚体、HLA 四聚体或 HLA 五聚体。

35. 一种由多核苷酸和/或其互补多核苷酸组成的疾病标志物, 其中所述多核苷酸含有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的碱基序列的至少 15 个连续核苷酸。

36. 权利要求 35 的疾病标志物, 所述疾病标志物用作检测癌症的探针或引物。

37. 权利要求 35 或 36 的疾病标志物, 其中来源于 Lengsin 基因的疾病标志物用于肺腺癌、肺鳞状细胞癌或胃癌。

38. 权利要求 35 或 36 的疾病标志物, 其中来源于 BJ-TSA-9 基因的疾病标志物用于白血病、肺腺癌、肺鳞状细胞癌、小细胞肺癌、口腔癌、胃癌、胰腺癌或淋巴瘤。

39. 权利要求 35 或 36 的疾病标志物, 其中来源于 C20orf42 基因的疾病标志物用于肺鳞状细胞癌、肺腺癌、肝癌、胃癌、白血病、恶性淋巴瘤组织、直肠癌、结肠癌或胰腺癌。

40. 权利要求 35 或 36 的疾病标志物, 其中来源于 BUB1 基因的疾病标志物用于乳腺癌、肺腺癌、肺鳞状细胞癌、卵巢癌、口腔鳞状细胞癌、肾癌、大肠癌(结肠癌、直肠癌)、胃癌、胰腺癌、肝癌、白血病、淋巴瘤或黑素瘤。

41. 权利要求 35 或 36 的疾病标志物, 其中来源于 C10orf3 基因的疾病标志物用于乳腺癌、结肠癌、直肠癌、肾癌、胃癌、卵巢癌、肝癌、胰腺癌、肺鳞状细胞癌、肺腺癌、小细胞肺癌或黑素瘤。

42. 权利要求 35 或 36 的疾病标志物, 其中来源于 HIFPH3 基因的疾病标志物用于乳腺癌、结肠癌、胃癌、肾癌、胰腺癌、肝癌、肺

腺癌或肺鳞状细胞癌。

43. 一种用于检测癌症的方法，所述方法包括以下步骤(a)、(b)和(c):

(a)使由受试者的生物样品制备的 RNA 或由其转录的互补多核苷酸与权利要求 35-42 中任一项的疾病标志物杂交;

(b)通过使用所述疾病标志物作为指示剂检测与疾病标志物杂交的、由生物样品制备的 RNA 或由该 RNA 转录的互补多核苷酸; 和

(c)基于(b)中的检测结果确定受试者是否患有癌症。

44. 权利要求 43 的方法，其中步骤(C)中，对受试者的检测结果与健康个体的结果进行比较且在受试者中观察到的与疾病标志物的杂交水平高于在健康个体中观察到的水平时，则确定受试者患有癌症。

45. 一种用于癌症的疾病标志物，所述疾病标志物包含特异性识别 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的抗体。

46. 权利要求 45 的疾病标志物，所述疾病标志物用作检测癌症的探针。

47. 一种用于检测癌症的方法，所述方法包括以下步骤(a)、(b)和(c):

(a)使由受试者的生物样品制备的蛋白与权利要求 45 或 46 中的疾病标志物结合;

(b)通过使用所述疾病标志物作为指示剂检测与疾病标志物结合的、由生物样品制备的蛋白或得自该蛋白的部分肽; 和

(c)基于(b)中的检测结果确定受试者是否患有癌症。

48. 权利要求 47 的方法，其中步骤(C)中，对受试者的检测结果与健康个体的结果进行比较且在受试者中观察到的与所述疾病标志物的结合水平高于在健康个体中观察到的水平时，则确定受试者患有癌症。

新肿瘤抗原肽

技术领域

[0001] 本发明涉及新肿瘤抗原蛋白和由其获得的肽以及它们在癌症免疫领域中的用途。

技术背景

[0002] 已知免疫系统,尤其是 T 细胞,在清除活体癌(肿瘤)方面发挥重要作用。实际上,已在人癌症病灶中观察到对癌细胞表现出细胞毒活性的淋巴细胞浸润(非专利文献 1),并在没有多大困难的情况下从黑色素瘤中分离到识别自体肿瘤细胞的细胞毒性 T 淋巴细胞(CTL)(非专利文献 2、3、4)。此外,通过转移 CTL 对黑色素瘤的临床治疗结果也表明了 T 细胞在癌症清除方面的重要性(非专利文献 5)。

[0003] 尽管在很长时间内并不清楚 CTL 攻击自体肿瘤细胞的靶分子,但近年来随着免疫学与分子生物学的发展,这类分子也逐渐清晰起来。具体而言,业已揭示,CTL 通过 T 细胞受体(TCR)识别一种被称作癌抗原肽的肽与主要组织相容性复合物 I 类抗原(MHC-I 类抗原,也称作 HLA 抗原)之间的复合物,并由此攻击自体肿瘤细胞。

[0004] 癌抗原肽是通过癌特异性抗原蛋白在细胞内合成后的细胞内蛋白酶体降解而生成的。这样生成的癌抗原肽结合至内质网中的 MHC-I 类抗原(HLA 抗原),形成复合物,该复合物被运输到细胞表面,作为抗原被提呈。抗原特异性 CTL 识别作为抗原被提呈的复合物,并通过细胞毒活性或者产生淋巴因子而表现出抗癌作用。由于这一系列作用的阐明,有可能通过使用癌抗原蛋白或者癌抗原肽作为通常所说的癌症疫苗来增强癌症患者体内的癌特异性 CTL,从而治疗癌症。

[0005] 就肿瘤抗原蛋白而言,T. Boon 等人于 1991 年首次从人

黑色素瘤细胞中鉴定出一种名为 MACE 的蛋白(非专利文献 6)。随后,主要由黑色素瘤细胞中鉴定出几种其它的肿瘤抗原蛋白。

[0006] 为了将肿瘤抗原蛋白和肽应用于肿瘤治疗或诊断,必须鉴定可广泛应用于腺癌如肺癌的新肿瘤抗原蛋白和肽,所述腺癌远比黑素瘤更频繁发生。

[0007] Lengsin (含谷氨酸-氨连接酶(谷氨酰胺合酶)结构域 1 (GLULD1) UniGene Hs.149585)被鉴定为在人晶状体中高度表达的新基因,根据氨基酸序列的相似性已报告其为谷氨酰胺合酶家族的可能成员。然而,该基因的生理功能仍是未知的(非专利文献 7)。

[0008] BJ-TSA-9 (假定蛋白 MGC14128)(UniGene Hs.379821)是一种在国立卫生研究院哺乳动物基因采集(MGC)计划中克隆的新基因,其生理功能仍是未知的(非专利文献 8)。

[0009] C20orf42 (或 URP1, Kindlerin)作为在肺癌和结肠癌中高度表达的基因被克隆(非专利文献 9),与秀丽线虫 (*C. elegans*) UNC-112 具有高度相似性。业已报道, C20orf42 和 UNC-112 一样与整联蛋白相互作用, C20orf42 的表达由 TGF- β 诱导(非专利文献 10)。C20orf42 还被鉴定为在 Kindler 综合征中突变的 Kindlerin 基因, Kindler 综合征是一种罕见的常染色体隐性皮肤病(非专利文献 11、12)。

[0010] BUB1 是涉及细胞周期 M 期的纺锤体关卡的丝氨酸/苏氨酸激酶,没有功能性 BUB1 导致异常染色体分离。另外, BUB1 磷酸化 Cdc2, 并抑制 APC/C^{cdc20} 的泛素连接酶活性,对异常染色体分离产生预防作用(非专利文献 13)。已报告 BUB1 在癌症组织中的表达和突变下降(非专利文献 14、15)。

[0011] C10orf3 是一种在国立卫生研究院哺乳动物基因采集(MGC)计划中克隆的新基因,其生理功能仍是未知的(非专利文献 16)。

[0012] HIFPH3 (egl9 同源物 3(egl nine homolog 3), EGLN3)作为具有 egl9 类似活性的 3 种分子中的 1 种被克隆,其被鉴定为在秀丽线虫中调节缺氧诱导因子 1 (HIF1)活性的加双氧酶(非专利文献 17)。

其被认为在缺氧条件下通过 HIF1 调节活性, 但 3 种 egl9 同源物之间的差异是未知的。

[0013] 非专利文献 1: Arch. Surg., 126:200 (1990)

非专利文献 2: Immunol. Today, 8:385 (1987)

非专利文献 3: J. Immunol., 138:989 (1987)

非专利文献 4: Int. J. Cancer, 52:52 (1992)

非专利文献 5: J. Natl. Cancer. Inst., 86:1159 (1994)

非专利文献 6: Science, 254, 1643-1647 (1991)

非专利文献 7: Mol. Vis. Jun. 15; 8:185-95 (2002)

非专利文献 8: Proc. Natl. Acad. Sci. U S A., 99(26):16899-903 (2002)

非专利文献 9: Biochim. Biophys. Acta 1637: 207-216 (2003)

非专利文献 10: J. Biol. Chem., Feb 20; 279(8):6824-33 (2004)

非专利文献 11: Hum. Molec. Genet., 12: 925-935 (2003)

非专利文献 12: Am. J. Hum. Genet., 73: 174-187 (2003)

非专利文献 13: Mol. Cell., Nov 5; 16(3):387-97 (2004)

非专利文献 14: Oncogene Jul 11;21(30):4673-9 (2002)

非专利文献 15: Leuk. Lymphoma. Feb; 43(2):385-91 (2002)

非专利文献 16: Proc. Natl. Acad. Sci. USA., Dec 24; 99(26):16899-903 (2002)

非专利文献 17: Cell 107: 43-54 (2001)

发明公开

本发明要解决的问题

[0014] 本发明的目标是提供新肿瘤抗原蛋白和肽以及它们在癌症免疫领域的用途。

解决问题的方法

[0015] 本发明人深入地尝试发现与正常组织相比在癌症组织

中显示出表达水平和频率增加的基因，并选定 6 个基因：Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因，由此发现这些基因和由其表达的产物(即蛋白)可以用作癌症的疾病标志物。

[0016] 然后，本发明人由 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因编码蛋白(Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3)的氨基酸序列选择可能结合 HLA 分子的部分肽，并检查它们与 HLA 分子的结合亲和力和 CTL 诱导活性。结果，发现这 6 种蛋白(尤其是 BJ-TSA-9、C20orf42 和 C10orf3)为肿瘤抗原蛋白，得自它们的部分肽为肿瘤抗原肽。因此，Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 以及由它们获得的肽被认为可以用作癌症疫苗，用于表现出这些蛋白表达(尤其是表达增加)的多种癌症。

由如上所述的发现实现了本发明。

[0017] 因此，本发明涉及以下内容：

(1)含有来源于 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分肽、能够结合 HLA 抗原并被 CTL 识别的肽；

(2)以上(1)的肽，其中 HLA 抗原为 HLA-A24 或 HLA-A2 抗原；

(3)以上(2)的肽，其含有 SEQ ID NO: 13-201 中任一项的氨基酸序列；

(4)含有与 SEQ ID NO: 13-31、42-49、59-78、89-117、158-165、176-183 和 195-201 中任一项的氨基酸序列相同的氨基酸序列的肽，只是 2 位的氨基酸被酪氨酸、苯丙氨酸、甲硫氨酸或色氨酸取代，和/或 C 末端氨基酸被苯丙氨酸、亮氨酸、异亮氨酸、色氨酸或甲硫氨酸取代，所述肽能够结合 HLA-A24 抗原，并被 CTL 识别；

(5)含有与 SEQ ID NO: 32-41、50-58、79-88、118-157、166-175 和 184-194 中任一项的氨基酸序列相同的氨基酸序列的肽，只是 2 位的氨基酸被亮氨酸、甲硫氨酸、缬氨酸、异亮氨酸或谷氨酰胺取代，和/或 C 末端氨基酸被缬氨酸或亮氨酸取代，所述肽能够结合 HLA-A2

抗原，并被 CTL 识别；

[0018] (6)含有以上(1)-(5)中任一项的肽的表位肽；

(7)含有以上(1)-(6)中任一项的肽和在药学上可接受的载体的药物组合物；

(8)含有编码以上(1)-(6)中任一项的肽的多核苷酸的核酸；

(9)含有以上(8)的核酸和在药学上可接受的载体的药物组合物；

(10)含有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 和在药学上可接受的载体的药物组合物；

[0019] (11)以上(10)的药物组合物，其中 Lengsin 含有 SEQ ID NO: 2 的氨基酸序列；

(12)以上(10)的药物组合物，其中 BJ-TSA-9 含有 SEQ ID NO: 4 的氨基酸序列；

(13)以上(10)的药物组合物，其中 C20orf42 含有 SEQ ID NO: 6 的氨基酸序列；

(14)以上(10)的药物组合物，其中 BUB1 含有 SEQ ID NO: 8 的氨基酸序列；

(15)以上(10)的药物组合物，其中 C10orf3 含有 SEQ ID NO: 10 的氨基酸序列；

[0020] (16)以上(10)的药物组合物，其中 HIFPH3 含有 SEQ ID NO: 12 的氨基酸序列；

(17)一种含有核酸和在药学上可接受的载体的药物组合物，所述核酸包含编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸；

(18)以上(17)的药物组合物，其中编码 Lengsin 的多核苷酸包含 SEQ ID NO: 1 的碱基序列，或编码 SEQ ID NO: 2 的氨基酸序列；

(19)以上(17)的药物组合物，其中编码 BJ-TSA-9 的多核苷酸包含 SEQ ID NO: 3 的碱基序列，或编码 SEQ ID NO: 4 的氨基酸序列；

(20)以上(17)的药物组合物，其中编码 C20orf42 的多核苷酸包含

SEQ ID NO: 5 的碱基序列, 或编码 SEQ ID NO: 6 的氨基酸序列;

[0021] (21)以上(17)的药物组合物, 其中编码 BUB1 的多核苷酸包含 SEQ ID NO: 7 的碱基序列, 或编码 SEQ ID NO: 8 的氨基酸序列;

(22)以上(17)的药物组合物, 其中编码 C10orf3 的多核苷酸包含 SEQ ID NO: 9 的碱基序列, 或编码 SEQ ID NO: 10 的氨基酸序列;

(23)以上(17)的药物组合物, 其中编码 HIFPH3 的多核苷酸包含 SEQ ID NO: 11 的碱基序列, 或编码 SEQ ID NO: 12 的氨基酸序列;

(24)一种制备抗原提呈细胞的方法, 其中使具有抗原提呈能力的细胞在体外与以下的任一种接触:

(a)以上(1)-(6)中任一项的肽,

(b)以上(8)的核酸,

(c) Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3,
和

(d)含有编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸的核酸;

(25)一种通过以上(24)的方法制备的抗原提呈细胞;

[0022] (26)一种含有以上(25)的抗原提呈细胞和在药学上可接受的载体的药物组合物;

(27)一种诱导 CTL 的方法, 其中使外周血淋巴细胞在体外与以下的任一种接触:

(a)以上(1)-(6)中任一项的肽,

(b)以上(8)的核酸,

(c) Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3,
和

(d)含有编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸的核酸;

(28)通过以上(27)的方法诱导的 CTL;

(29)一种含有以上(28)的 CTL 和在药学上可接受的载体的药物组

合物;

(30)以上(7)、(9)-(23)、(26)或(29)中任一项的药物组合物,其用作 CTL 诱导物;

[0023] (31)以上(7)、(9)-(23)、(26)或(29)中任一项的药物组合物,其用作癌症疫苗;

(32)一种特异性结合以上(1)-(5)中任一项的肽的抗体;

(33)一种含有以上(1)-(5)中任一项的肽和 HLA 抗原的 HLA 单体、HLA 二聚体、HLA 四聚体或 HLA 五聚体;

(34)一种用于检测对肿瘤抗原肽特异性的 CTL 的试剂,所述肿瘤抗原肽来源于 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3,所述试剂含有作为组分的以上(33)的 HLA 单体、HLA 二聚体、HLA 四聚体或 HLA 五聚体;

(35)一种由多核苷酸和/或其互补多核苷酸组成的疾病标志物,其中所述多核苷酸含有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的碱基序列的至少 15 个连续核苷酸;

[0024] (36)以上(35)的疾病标志物,其用作检测癌症的探针或引物;

(37)以上(35)或(36)的疾病标志物,其中来源于 Lengsin 基因的疾病标志物用于肺腺癌、肺鳞状细胞癌或胃癌;

(38)以上(35)或(36)的疾病标志物,其中来源于 BJ-TSA-9 基因的疾病标志物用于白血病、肺腺癌、肺鳞状细胞癌、小细胞肺癌、口腔癌、胃癌、胰腺癌或淋巴瘤;

(39)以上(35)或(36)的疾病标志物,其中来源于 C20orf42 基因的疾病标志物用于肺鳞状细胞癌、肺腺癌、肝癌、胃癌、白血病、恶性淋巴瘤组织、直肠癌、结肠癌或胰腺癌;

(40)以上(35)或(36)的疾病标志物,其中来源于 BUB1 基因的疾病标志物用于乳腺癌、肺腺癌、肺鳞状细胞癌、卵巢癌、口腔鳞状细胞癌、肾癌、大肠癌(结肠癌、直肠癌)、胃癌、胰腺癌、肝癌、白血病、

淋巴瘤或黑素瘤;

[0025] (41)以上(35)或(36)的疾病标志物,其中来源于 C10orf3 基因的疾病标志物用于乳腺癌、结肠癌、直肠癌、肾癌、胃癌、卵巢癌、肝癌、胰腺癌、肺鳞状细胞癌、肺腺癌、小细胞肺癌或黑素瘤;

(42)以上(35)或(36)的疾病标志物,其中来源于 HIFPH3 基因的疾病标志物用于乳腺癌、结肠癌、胃癌、肾癌、胰腺癌、肝癌、肺腺癌或肺鳞状细胞癌;

(43)一种用于检测癌症的方法,所述方法包括以下步骤(a)、(b)和(c):

(a)使由受试者的生物样品制备的 RNA 或由其转录的互补多核苷酸与以上(35)-(42)中任一项的疾病标志物杂交;

(b)通过使用所述疾病标志物作为指示剂检测与疾病标志物杂交的、由生物样品制备的 RNA 或由其转录的互补多核苷酸; 和

(c)基于(b)中的检测结果确定受试者是否患有癌症;

(44)以上(43)的方法,其中在步骤(C)中,将受试者的检测结果与健康个体的结果进行比较且在受试者中观察到的与疾病标志物的杂交水平高于在健康个体中观察到的该水平时,则确定受试者患有癌症;

(45)一种用于癌症的疾病标志物,所述疾病标志物含有特异性识别 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的抗体;

[0026] (46)以上(45)的疾病标志物,所述疾病标志物用作检测癌症的探针;

(47)一种用于检测癌症的方法,所述方法包括以下步骤(a)、(b)和(c):

(a)使由受试者的生物样品制备的蛋白与以上(45)或(46)的疾病标志物结合;

(b)通过使用所述疾病标志物作为指示剂检测与疾病标志物结合的、由生物样品制备的蛋白或得自该蛋白的部分肽; 和

(c)基于(b)中的检测结果确定受试者是否患有癌症;

(48)以上(47)的方法, 其中在步骤(C)中, 对受试者的检测结果与健康个体的结果进行比较且在受试者中观察到的与疾病标志物的结合水平高于在健康个体中观察到的该水平时, 则确定受试者患有癌症。

本发明的作用

[0027] 本发明的新肿瘤抗原肽、编码其的核酸等可以用作癌症疫苗。此外, 所述肿瘤抗原肽还可以用作 HLA 四聚体等的组分, 以检测 CTL。

实施本发明的最佳方式

[0028] 本文使用的氨基酸、(多)肽、(多)核苷酸等的缩写依照 IUPAC-IUB (IUPAC-IUB Communication on Biological Nomenclature, Eur. J. Biochem., 138: 9 (1984))的规则、“撰写含有碱基序列或氨基酸序列的说明书的指南”(日本特许办公室)和本领域常用的符号。

[0029] 1)本发明的蛋白

本发明的蛋白为 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3。

本发明的蛋白 Lengsin 含有 SEQ ID NO: 2 的氨基酸序列或类似于上述氨基酸序列的氨基酸序列。本发明的蛋白 Lengsin 可以为来源于天然来源(例如肺腺癌细胞系 1-87)的蛋白或重组蛋白。在此, SEQ ID NO: 2 的氨基酸序列以登录号 NM_016571、登录号 NP_057655 向 GenBank 数据库登记, 并代表人 Lengsin (含谷氨酸-氨连接酶(谷氨酰胺合酶)结构域 1, 也称为 GLULD1)。人 Lengsin 公开于非专利文献 7 (Mol. Vis. Jun. 15; 8:185-95 (2002))。

[0030] 本发明的蛋白 BJ-TSA-9 含有 SEQ ID NO: 4 的氨基酸序列或类似于上述氨基酸序列的氨基酸序列。本发明的蛋白 BJ-TSA-9 可以为来源于天然来源(例如肺腺癌细胞系 1-87)的蛋白或重组蛋白。

在此, SEQ ID NO: 4 的氨基酸序列以登录号 NM_032899、登录号 NP_116288 向 GenBank 数据库登记, 并代表人 BJ-TSA-9 (假定蛋白 MGC14128)。人 BJ-TSA-9 公开于非专利文献 8 (Proc. Natl. Acad. Sci. USA., 99(26):16899-903 (2002))。

[0031] 本发明的蛋白 C20orf42 含有 SEQ ID NO: 6 的氨基酸序列或类似于上述氨基酸序列的氨基酸序列。本发明的蛋白 C20orf42 可以为来源于天然来源(例如结肠癌细胞系 SW480)的蛋白或重组蛋白。在此, SEQ ID NO: 6 的氨基酸序列以登录号 NM_017671、登录号 NP_060141 向 GenBank 数据库登记, 并代表人 C20orf42 (URP1, 也称为 Kindlerin)。人 C20orf42 公开于非专利文献 9 (Biochim. Biophys. Acta 1637: 207-216 (2003))。

[0032] 本发明的蛋白 BUB1 含有 SEQ ID NO: 8 的氨基酸序列或类似于上述氨基酸序列的氨基酸序列。本发明的蛋白 BUB1 可以为来源于天然来源(例如胰腺癌细胞系 PUN)的蛋白或重组蛋白。在此, SEQ ID NO: 8 的氨基酸序列以登录号 NM_004336、登录号 NP_004327 向 GenBank 数据库登记, 并代表人 BUB1。人 BUB1 公开于文献 (Genomics 46:379-388(1997))。

[0033] 本发明的蛋白 C10orf3 含有 SEQ ID NO: 10 的氨基酸序列或类似于上述氨基酸序列的氨基酸序列。本发明的蛋白 C10orf3 可以为来源于天然来源(例如肺腺癌细胞系 1-87)的蛋白或重组蛋白。在此, SEQ ID NO: 10 的氨基酸序列以登录号 NM_018131、登录号 NP_060601 向 GenBank 数据库登记, 并代表人 C10orf3。人 C10orf3 公开于非专利文献 16 (Proc. Natl. Acad. Sci. USA., Dec 24; 99(26):16899-903 (2002))。

[0034] 本发明的蛋白 HIFPH3 含有 SEQ ID NO: 12 的氨基酸序列或类似于上述氨基酸序列的氨基酸序列。本发明的蛋白 HIFPH3 可以为来源于天然来源(例如肾癌细胞系 SMKTR-1)的蛋白或重组蛋白。在此, SEQ ID NO: 12 的氨基酸序列以登录号 NM_022073、登录号

NP_071356 向 GenBank 数据库登记, 并代表人 HIFPH3 (egl9 同源物 3, 也称为 EGLN3)。人 HIFPH3 公开于非专利文献 17 (Cell 107: 43-54 (2001))。

[0035] Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 中的每一种在本文都还被称为“本发明的蛋白”。

[0036] 具体地说, “含有 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的蛋白”包括由 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列组成的蛋白, 以及由包含在 N 和/或 C 末端连接额外的氨基酸序列的 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的氨基酸序列组成的蛋白。“额外的氨基酸序列”可为来源于和本发明蛋白不同的其它结构基因的氨基酸序列。

[0037] 具体地说, “含有与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列类似的氨基酸序列的蛋白”包括以下的蛋白(a)-(c):

(a)含有除缺失、取代和/或添加一个或多个氨基酸以外与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列相同的氨基酸序列、并具有作为肿瘤抗原蛋白的活性的蛋白;

(b)含有与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列具有至少 70%序列同一性的氨基酸序列、并具有作为肿瘤抗原蛋白的活性的蛋白;

(c)由能够在严格条件下与编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的多核苷酸的互补链杂交的多核苷酸编码、并具有作为肿瘤抗原蛋白的活性的蛋白。

[0038] 优选的实例包括由与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列类似的氨基酸序列组成的蛋白。该蛋白的实例包括以下的蛋白(a')-(c'):

(a')由除缺失、取代和/或添加一个或多个氨基酸以外与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列相同的氨基酸序列组成、并具有作为肿瘤抗原蛋白的活性的蛋白;

(b')由与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列具有至少 70%序列同一性的氨基酸序列组成、并具有作为肿瘤抗原蛋白的活性的蛋白;

(c')由能够在严格条件下与编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的多核苷酸的互补链杂交的多核苷酸编码、并具有作为肿瘤抗原蛋白的活性的蛋白。

[0039] 在以上(a)中“含有除缺失、取代和/或添加一个或多个氨基酸以外与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列相同的氨基酸序列的蛋白”是指人工生产的蛋白,即例如在活体中存在的修饰过的(变体)蛋白或等位基因变体。

在这方面,关于蛋白中的修饰(突变)的数量或位置没有限制,只要保持本发明蛋白的活性。标准(人们可基于其确定要缺失、取代和/或添加的氨基酸残基的数量或位置而不降低活性)可使用本领域众所周知的计算机程序获得,例如 DNA Star 软件。例如,突变数通常应在总氨基酸残基的 10%之内,优选应在总氨基酸残基的 5%之内。而且,考虑到保持结构,通过取代引入的氨基酸优选与被取代的氨基酸具有类似特征,所述特征包括极性、电荷、溶解性、疏水性、亲水性、两亲性等。例如,Ala、Val、Leu、Ile、Pro、Met、Phe 和 Trp 被分类为非极性氨基酸;Gly、Ser、Thr、Cys、Tyr、Asn 和 Gln 被分类为不带电的氨基酸;Asp 和 Glu 被分类为酸性氨基酸;而 Lys、Arg 和 His 被分类为碱性氨基酸。本领域一般技术人员可以基于这些标准选择属于相同类别的适宜氨基酸。

[0040] 在以上(b)中“含有与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列具有至少 70%序列同一性的氨基酸序列的蛋白”包括含有的氨基酸序列与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列具有至少约 70%、优选约 80%、更优选约 90%、再更优选约 95%的序列同一性的蛋白,具体地说,为由 SEQ ID NO: 2、4、6、8、10 或 12 的部分氨基酸序列组成的蛋白。

[0041] 本文使用的术语“序列同一性”是指两种蛋白之间的同一性和同源性。“序列同一性”通过比较两个在待比较序列区域内最优比对的序列来确定。在本文中，待比较的最优比对的蛋白可具有添加或缺失(例如“空位”)。可通过使用例如 Vector NTI、ClustalW 算法(Nucleic Acid Res., 22 (22): 4673-4680(1994))进行比对计算序列同一性。序列同一性可使用用于序列分析的软件(具体地说为 Vector NTI 或 GENETYX-MAC)或由公共数据库提供的测序工具来确定。这样的公共数据库通常可在 Web 站点(<http://www.ddbj.nig.ac.jp>)获得。

[0042] 在以上(c)中“能够在严格条件下与编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的多核苷酸的互补链杂交的多核苷酸”包括含有的碱基序列与编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的多核苷酸具有至少约 40%、优选约 60%、更优选约 70%、再更优选约 80%、甚至更优选约 90%以及最优选约 95%的序列同一性的多核苷酸。具体地说，例举的多核苷酸含有的碱基序列与 SEQ ID NO: 1、3、5、7、9 或 11 的碱基序列具有至少约 40%、优选约 60%、更优选约 70%、再更优选约 80%、甚至更优选约 90%以及最优选约 95%的序列同一性。更具体地说，例举的多核苷酸由 SEQ ID NO: 1、3、5、7、9 或 11 的碱基序列的部分序列组成。

[0043] 可以按照自身已知的方法或与其等效的方法进行杂交，例如在基础教科书如“Molecular Cloning, 第 2 版, Cold Spring Harbor Laboratory Press (1989)”等中描述的方法。另外，可使用市售可得的文库按照其随附的说明书进行杂交。

本文使用的“严格条件”可如在文献(Berger 和 Kimmel, 1987, “Guide to Molecular Cloning Techniques Methods in Enzymology”, 152 卷, Academic Press, San Diego Calif.; 或“Molecular Cloning”第 2 版, Cold Spring Harbor Laboratory Press (1989), 出处同上)中所述基于形成复合物的核酸或结合探针的核酸的解链温度(T_m)来确定。

[0044] 例如，杂交可在含有 6xSSC (20xSSC 表示 333 mM 柠檬

酸钠、333 mM NaCl)、0.5% SDS 和 50%甲酰胺的溶液于 42°C 进行或在含有 6xSSC 的溶液(无 50%甲酰胺)于 65°C 进行。

杂交后的洗涤可在约“1xSSC、0.1% SDS、37°C”的条件下进行。优选地，在这样的洗涤条件下洗涤时，互补链保持与靶正向链结合。更严格的杂交条件可包括约“0.5xSSC、0.1% SDS、42°C”的洗涤条件，再更严格的杂交条件为约“0.1xSSC、0.1% SDS、65°C”，但其不限于此。

[0045] 本发明的蛋白具有作为肿瘤抗原蛋白的活性。术语“作为肿瘤抗原蛋白的活性”是指通过用于肿瘤抗原蛋白活性的常规测定检测的活性。具体地说，其是指表达 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的细胞被 CTL 识别的特征，即细胞显现出对 CTL 的反应性，换句话说，本发明的蛋白或由其获得的抗原肽活化或诱导 CTL。

[0046] 在该方面，“细胞”优选表达 HLA 抗原。因此，更具体地说，“作为肿瘤抗原蛋白的活性”是指以下特征：当本发明的蛋白在表达 HLA 抗原如 HLA-A24 或 HLA-A2 的细胞中表达时，源自本发明蛋白的肿瘤抗原肽和 HLA 抗原之间的复合物被提呈至细胞表面上，从而使该细胞被 CTL 识别，换句话说，CTL 被活化(诱导)。

[0047] 如上陈述的本发明蛋白的特征可易于通过已知方法或其等效方法确定，例如 ^{51}Cr 释放测定(J. Immunol., 159: 4753, 1997)、使用 LDH 细胞毒性检测试剂盒的 LDH 释放测定(Takara Bio, Inc.)、细胞因子检测等。详细测定方案将在后文阐述。

[0048] 首先，用含编码本发明蛋白的 DNA 的表达载体和含编码 HLA 抗原的 DNA 的表达载体共转染宿主细胞，如 293-EBNA 细胞(Invitrogen)。编码 HLA 抗原的 DNA 包括编码 HLA-A24 抗原或 HLA-A2 抗原的 DNA。编码 HLA-A24 抗原的 DNA 的实例包括 HLA-A2402 cDNA (Cancer Res., 55: 4248-4252 (1995), Genbank 登录号 M64740)。编码 HLA-A2 抗原的 DNA 的实例包括 HLA-A0201 cDNA

(GenBank 登录号 M84379)。

[0049] 如上所述的转染可通过使用脂转染胺试剂(GIBCO BRL)的脂转染法等进行。然后,加入限于 HLA 抗原所用的 CTL,并使其反应,之后例如通过诸如 ELISA 的方法检测活化的(反应性)CTL 产生的多种细胞因子(例如 IFN- γ)。本文能够使用的 CTL 可通过用本发明的蛋白刺激外周血淋巴细胞制备,或者按照 Int. J. Cancer, 39, 390-396, 1987, N. Eng. J. Med, 333, 1038-1044, 1995 的方法等建立。

[0050] 本发明蛋白的 CTL 诱导活性还可以通过其中使用对人的模型动物的测定在体内检验(WO 02/47474; Int. J. Cancer. 100, 565-570 (2002))。

本发明的蛋白可通过用于由天然产物(例如癌细胞系)纯化蛋白的自身已知的方法制备,或通过后文所述方法制备,该方法包括培养携带核酸的转化体,所述核酸包含编码本发明蛋白的多核苷酸。

[0051] 2)来源于本发明蛋白的肽

本发明的肽可被称为“本发明的肽”或“本发明的肿瘤抗原肽”,是含有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分肽的肿瘤抗原肽,能够结合 HLA 抗原,并被 CTL 识别。因此,本发明的肽可以包含对应于本发明蛋白的氨基酸序列的任何位置的肽,并为任何长度,只要该肽含有如上文定义的本发明蛋白的部分氨基酸序列,并可以与 HLA 抗原形成被 CTL 识别的复合物。

[0052] 可如下鉴定本发明的肽:合成候选肽,其为 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分片段,并对候选肽进行测定,以检验 CTL 是否识别候选肽与 HLA 抗原之间的复合物,也就是说,候选肽是否具有作为肿瘤抗原肽的活性。

可按照一般用于肽化学领域的方法进行该肽的合成。这样的方法可见于文献,包括 Peptide Synthesis, Interscience, New York, 1966; The Proteins, 第 2 卷, Academic Press Inc., New York, 1976; Peptide-Gosei, Maruzen, Inc., 1975; Peptide-Gosei no Kiso to Jikken, Maruzen, Inc.,

1985; 和 Iyaku hin no Kaihatsu (Zoku), 第 14 卷, Peptide-Gosei, Hirokawa-syoten, 1991.

[0053] 鉴定本发明的肿瘤抗原肽的方法将在后文详述。

针对某些 HLA 类型, 例如 HLA-A1、-A0201、-A0204、-A0205、-A0206、-A0207、-A11、-A24、-A31、-A6801、-B7、-B8、-B2705、-B37、-Cw0401 和 -Cw0602, 已阐述了被提呈的、结合 HLA 抗原的肿瘤抗原肽的氨基酸序列的规律性(基序)。参见 Immunogenetics, 41: 178 页, 1995 等。例如, HLA-A24 的基序已知具有 8-11 个氨基酸的氨基酸序列, 其中在 2 位的氨基酸为酪氨酸、苯丙氨酸、甲硫氨酸或色氨酸, C 末端氨基酸为苯丙氨酸、亮氨酸、异亮氨酸、色氨酸或甲硫氨酸(J. Immunol., 152, 3913 页, 1994, Immunogenetics, 41: 178 页, 1995. J. Immunol., 155 :4307 页, 1994)。对于 HLA-A2 的基序, 列于表 1 的那些是已知的(Immunogenetics, 41, 178 页, 1995, J. Immunol., 155: 4749 页, 1995)。

[0054] [表 1]

HLA-A2 型	在 N-末端的 2 位的氨基酸	在 C-末端的氨基酸
HLA-A0201	L、M	V、L
HLA-A0204	L	L
HLA-A0205	V、L、I、M	L
HLA-A0206	V、Q	V、L
HLA-A0207	L	L

*所有肽都为 8-11 个氨基酸长。

[0055] 最近, 已有可能使用 NIH 的 BIMAS 软件 (http://bimas.dcrt.nih.gov/molbio/hla_bind/)经互联网检索预期能够结合 HLA 抗原的肽序列。

关于肽的长度, 结合多种 HLA 分子的抗原肽的分析揭示, 其一般为约 8-14 个氨基酸(Immunogenetics, 41: 178, 1995)。然而, 对于 HLA-DR、-DP 和 -DQ 而言, 还已知由 14 个以上的氨基酸组成的肽。

[0056] 考虑到所述基序，易于由本发明蛋白的氨基酸序列选择对应于该肽的部分。例如，利用 BIMAS 软件，可容易地选择预期能够结合 HLA 抗原的序列。可如下鉴定本发明的肽：通过上述方法合成选定的候选肽，并检验该候选肽是否结合 HLA 抗原并被 CTL 识别，即该候选肽是否具有作为肿瘤抗原肽的活性。

具体地说，可通过在 *J. Immunol.*, 154, 2257 页, 1995 中所述的方法进行鉴定。因此，加入候选肽，以在体外刺激分离自人类受试者的外周血淋巴细胞，所述人类受试者对预期提呈所述候选肽的 HLA 抗原为阳性。在特异性识别用候选肽脉冲的 HLA 阳性细胞的 CTL 被诱发时，该候选肽可能为肿瘤抗原肽。例如，通过使用 ELISA 等检测 CTL 响应于抗原提呈细胞产生的多种细胞因子(例如 IFN- γ)的量，可以检验是否发生 CTL 诱导。或者，还可以通过 ^{51}Cr 释放测定检验 CTL 诱导，在 ^{51}Cr 释放测定中，检测 CTL 针对用 ^{51}Cr 标记的抗原提呈细胞的细胞毒性(*Int. J. Cancer*, 58: 317 页, 1994)。

[0057] 此外，可通过用候选肽脉冲诸如 293-EBNA 细胞 (Invitrogen) 的细胞检验 CTL 的诱导，其中所述细胞已导入编码预期存在候选肽的 HLA 抗原类型的 cDNA 的表达质粒，使细胞与限于预期提呈候选肽的前述类型的 HLA 抗原的 CTL 反应，并检测由 CTL 产生的多种细胞因子(例如 IFN- γ)(*J. Exp. Med.*, 187: 277, 1998)。

HLA 抗原的实例包括 HLA-A24 抗原和 HLA-A2 抗原。为选择限于 HLA-A24 的肿瘤抗原肽，可使用 HLA-A2402 cDNA (*Cancer Res.*, 55: 4248-4252 (1995), Genbank 登录号 M64740) 作为编码 HLA 抗原的 cDNA。为选择限于 HLA-A2 的肿瘤抗原肽，可使用 HLA-A0201 cDNA (GenBank 登录号 M84379) 作为编码 HLA 抗原的 cDNA。

[0058] 对于 CTL，除了通过用肽刺激人外周血淋巴细胞获得的那些以外，还可以使用通过在文献(*Int. J. Cancer*, 39, 390-396, 1987; *N. Eng. J. Med.*, 333, 1038-1044, 1995) 中描述的方法建立的 CTL。

本发明肽的体内活性可通过使用针对人的动物模型的测定(WO

02/47474, Int J. Cancer 100, 565-570 (2002))来确定。

[0059] 在以上情况中, 肿瘤抗原肽序列的规律性(基序)是已知的; 然而, 当肽的基序为未知时, HLA-B55 或 HLA-A26 就是这种情况, 本发明的肿瘤抗原肽可按照例如描述于 WO97/46676 的方法鉴定, 只有能够识别 HLA 抗原和肿瘤抗原肽之间的复合物的 CTL 细胞系是可用的。

[0060] 本发明肽的具体实例包括来源于以下的部分肽: 由 SEQ ID NO: 2 的氨基酸序列组成的 Lengsin、由 SEQ ID NO: 4 的氨基酸序列组成的 BJ-TSA-9、由 SEQ ID NO: 6 的氨基酸序列组成的 C20orf42、由 SEQ ID NO: 8 的氨基酸序列组成的 BUB1、由 SEQ ID NO: 10 的氨基酸序列组成的 C10orf3 或由 SEQ ID NO: 12 的氨基酸序列组成的 HIFPH3, 其能够结合 HLA 抗原, 并被 CTL 识别。考虑到本发明肽结合的 HLA 抗原, 优选实例包括能够结合 HLA-A24 或 HLA-A2 抗原的肽。肽的长度可优选为 8-14 个氨基酸, 更优选为 8-11 个氨基酸。

[0061] 具体地说, 本发明的肽包括含 SEQ ID NO: 13-201 中任一个的氨基酸序列、能够结合 HLA 抗原并被 CTL 识别的肽。该肽的长度可优选为 9-14 个氨基酸, 更优选为 9-11 个氨基酸。更具体地说, 作为结合 HLA-A24 的肿瘤抗原肽, 例举的肽由 SEQ ID NO: 13-31、42-49、59-78、89-117、158-165、176-183 和 195-201 中的任一个氨基酸序列组成, 能够结合 HLA-A24 抗原, 并被 CTL 识别(参见以下的表 5、7、9、11、13 和 15)。优选地, 例举的肽由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 或 183 的氨基酸序列组成。

更优选地, 例举的肽由 SEQ ID NO: 158、22、23、26、27、198、200、201、177、178、179 或 183 的氨基酸序列组成。

[0062] 此外, 作为结合 HLA-A2 的肿瘤抗原肽, 例举的肽由

SEQ ID NO: 32-41、50-58、79-88、118-157、166-175 和 184-194 中的任一个氨基酸序列组成，能够结合 HLA-A2 抗原，并被 CTL 识别(参见以下的表 6、8、10、12、14 和 16)。

[0063] 在本发明范围内，本发明的肽不仅包括由 SEQ ID NO: 2、4、6、8、10 或 12 的部分氨基酸序列组成的肽，而且包括通过部分修饰前述肽产生的变体(修饰)肽，只要该变体肽具有能够结合 HLA 抗原并被 CTL 识别的特征。具体地说，所包含的氨基酸序列除了已引入的至少一个氨基酸修饰以外与本发明肽的氨基酸序列(由 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分氨基酸序列组成，具体地说是 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列)相同，且具有作为肿瘤抗原肽的活性(即能够结合 HLA 抗原并被 CTL 识别)的变体肽，属于本发明的范围。

[0064] 氨基酸残基的“修饰”是指氨基酸残基的取代、缺失和/或添加，包括向肽的 N-和/或 C-末端添加，优选为氨基酸残基的取代。当该修饰包括氨基酸残基取代时，可随意选择要取代的氨基酸残基的数量或位置，只要保持作为肿瘤抗原肽的活性；然而，优选所述取代包括 1 至数个氨基酸残基，因为肿瘤抗原肽一般为如上提及的约 8-14 个氨基酸长。

[0065] 本发明的变体肽优选为 8-14 个氨基酸长(然而，就 HLA-DR、-DP 或 -DQ 而言，可接受由 14 个以上的氨基酸组成的肽)。

如上所述，对于某些 HLA 类型，例如 HLA-A1、-A0201、-A0204、-A0205、-A0206、-A0207、-A11、-A24、-A31、-A6801、-B7、-B8、-B2705、-B37、-Cw0401 和 -Cw0602，已知结合 HLA 抗原并被提呈的抗原肽中的基序。此外，有可能经由互联网(http://bimas.dcrn.nih.gov/molbio/hla_bind/)检索预期能够结合 HLA 抗原的肽序列。因此，人们可以基于该基序等制备以上的变体肽。

[0066] 例如，如上文所述，能够结合 HLA-A24 并被提呈的抗原肽的基序已知为特征如下的序列：在 8-11 个氨基酸的肽中，2 位的

氨基酸为酪氨酸、苯丙氨酸、甲硫氨酸或色氨酸，C末端氨基酸为苯丙氨酸、亮氨酸、异亮氨酸、色氨酸或甲硫氨酸(J. Immunol., 152: 3913 页, 1994; Immunogenetics, 41: 178 页, 1995; J. Immunol., 155: 4307 页, 1994)。对于 HLA-A2，基序已知为特征如下的序列：在 8-11 个氨基酸的肽中，2 位的氨基酸为亮氨酸、甲硫氨酸、缬氨酸、异亮氨酸或谷氨酰胺，C 末端氨基酸为缬氨酸或亮氨酸(Immunogenetics, 41: 178 页, 1995; J. Immunol., 155: 4749 页, 1995)。此外，预期能够结合 HLA 抗原的某些肽序列经由互联网([http://bimas.dcrt.nih.gov/molbio/hla bind/](http://bimas.dcrt.nih.gov/molbio/hla_bind/))公开。与以上基序的可用氨基酸具有类似特征的氨基酸也是可接受的。因此，本发明包括这样的变体肽：其含有的氨基酸序列与本发明肽的氨基酸序列相同，只是根据基序在可用于取代的位置(就 HLA-A24 和 HLA-A2 而言，为 2 位和 C-末端)的氨基酸被另一个氨基酸取代，优选被来自例如上述互联网站的预期提供结合活性的氨基酸取代，具有结合 HLA 抗原的活性，并被 CTL 识别。

[0067] 更优选地，本发明包括这样的变体肽：其在上述位置的氨基酸残基被根据基序已知可用的另一个氨基酸取代，并具有作为肿瘤抗原肽的活性。因此，对于如示于 SEQ ID NO: 13-31、42-49、59-78、89-117、158-165、176-183 和 195-201 的 HLA-A24 结合肽，变体肽的实例包括这样的肽：其含有的氨基酸序列与 SEQ ID NO: 13-31、42-49、59-78、89-117、158-165、176-183 和 195-201 中的任一个氨基酸序列相同，只是在 2 位的氨基酸被酪氨酸、苯丙氨酸、甲硫氨酸或色氨酸取代，和/或 C 末端氨基酸被苯丙氨酸、亮氨酸、异亮氨酸、色氨酸或甲硫氨酸取代，能够结合 HLA-A24 抗原，并被 CTL 识别。尤其是，更优选其 2 位的氨基酸被酪氨酸取代的肽。

[0068] 更优选地，变体肽由这样的氨基酸序列组成：其与 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、

198、199、200、201、177、178、179 和 183 中的任一个氨基酸序列相同，只是 2 位的氨基酸被酪氨酸、苯丙氨酸、甲硫氨酸或色氨酸取代，和/或 C 末端氨基酸被苯丙氨酸、亮氨酸、异亮氨酸、色氨酸或甲硫氨酸取代，该变体肽能够结合 HLA-A24 抗原，并被 CTL 识别。更优选地，变体肽由这样的氨基酸序列组成：其与 SEQ ID NO: 158、22、23、26、27、198、200、201、177、178、179 和 183 中的任一个氨基酸序列相同，只是在 2 位的氨基酸和/或 C 末端的氨基酸如上所述被取代，该变体肽能够结合 HLA-A24 抗原，并被 CTL 识别。

[0069] 就示于 SEQ ID NO: 32-41、50-58、79-88、118-157、166-175 和 184-194 的 HLA-A2 结合肽而言，优选的变体肽含有的氨基酸序列与 SEQ ID NO: 32-41、50-58、79-88、118-157、166-175 和 184-194 中的任一个氨基酸序列相同，只是 2 位的氨基酸被亮氨酸、甲硫氨酸、缬氨酸、异亮氨酸或谷氨酰胺取代，和/或 C 末端氨基酸被缬氨酸或亮氨酸取代，能够结合 HLA-A2 抗原，并被 CTL 识别。

[0070] 本发明的肽还包括含有如上所述的本发明肿瘤抗原肽的表位肽。

最近，业已表明，由连接在一起的多个(多于 1 个) CTL 表位(抗原肽)组成的肽(“表位肽”)有效诱导 CTL。例如，业已报道，由源自肿瘤抗原蛋白 PSA (各自限于 HLA-A2-、-A3、-A11 或 B53)的 CTL 表位连接在一起组成的肽(约 30 聚体)在体内诱导对各自的 CTL 表位特异性的 CTL (Journal of Immunology 1998, 161: 3186-3194)。

[0071] 另外，业已表明，由 CTL 表位和辅助表位连接在一起组成的肽(表位肽)有效诱导 CTL。在本文中，“辅助表位”是指能够激活 CD4 阳性 T 细胞的肽(Immunity., 1:751, 1994)，其实例包括乙型肝炎病毒来源的 HBVc128-140、破伤风毒素来源的 TT947-967 等。由辅助表位激活的 CD4⁺ T 细胞表现出一些活性，包括诱导并维持 CTL，以及活化诸如巨噬细胞的效应子，并因此被认为在免疫学抗肿瘤应答中有重要作用。作为由辅助表位与 CTL 表位连接在一起组成的肽的具

体实例,有报道指出,由6种HBV来源的HLA-A2限制性抗原肽、3种HLA-A11限制性抗原肽及辅助表位组成的DNA(微基因)在体内有效诱导针对各自表位的CTL(*Journal of Immunology* 1999, 162: 3915-3925)。实际上,由CTL表位(对应于黑色素瘤抗原gp100的280-288位的肿瘤抗原肽)和辅助表位(来源于破伤风毒素的T辅助表位)连接组成的肽已进行临床试验(*Clinical Cancer Res.*, 2001, 7:3012-3024)。

因此,本发明的肿瘤抗原肽还包括由含有连接在一起的本发明肽并具有CTL诱导活性的多个表位组成的肽(表位肽)。

[0072] 在这方面,“表位肽”被定义为(1)由2个或多个CTL表位(肿瘤抗原肽)连接在一起组成的肽,或(2)由CTL表位和辅助表位连接在一起组成的肽,其在抗原提呈细胞中被加工,产生肿瘤抗原肽,然后被细胞提呈,并诱导CTL。

[0073] 当CTL表位与本发明的肽连接时,CTL表位可来源于示于SEQ ID NO: 2的Lengsin、示于SEQ ID NO: 4的BJ-TSA-9、示于SEQ ID NO: 6的C20orf42、示于SEQ ID NO: 8的BUB1、示于SEQ ID NO: 10的C10orf3或示于SEQ ID NO: 12的HIFPH3的氨基酸序列,并限于HLA-A1、-A0201、-A0204、-A0205、-A0206、-A0207、-A11、-A24、-A31、-A6801、-B7、-B8、-B2705、-B37、-B55、-Cw0401、-Cw0602等。来源于其它肿瘤抗原蛋白的CTL表位也是可用的。多个CTL表位可以连接在一起,基于结合多种HLA分子的抗原肽的分析,CTL表位的长度可为约8-14个氨基酸(*Immunogenetics*, 41: 178, 1995)。

当辅助表位连接本发明的肽时,辅助表位可为前述乙型肝炎病毒来源的HBVc128-140、破伤风毒素来源的TT947-967等。辅助表位长度可为约13-30个氨基酸,优选为约13-17个氨基酸。

[0074] 由多个表位连接在一起组成的肽(表位肽)可通过前述的常规肽合成方法来制备。也可以通过常规DNA合成方法以及基于多核苷酸(其编码由多个表位连接在一起组成的表位肽)的序列信息进行

基因工程来制备。也就是说,可通过将编码表位肽的多核苷酸插入已知的表达载体,使用获得的重组表达载体转化宿主细胞,培养转化子并从培养物中回收目的表位肽,制备由多个表位连接在一起组成的表位肽。这些方法可以根据例如文献中描述的方法(Molecular Cloning, T. Maniatis 等, CSH Laboratory (1983), DNA Cloning, DM. Glover, IRL PRESS (1985))进行。

如上所述产生的由多个表位连接在一起组成的表位肽,可通过如上所述的测定检验体外 CTL 诱导活性,或者利用 WO02/47474 或 Int J. Cancer. 100, 565-570 (2002)中描述的测定使用针对人的模型动物检验体内 CTL 诱导活性。

[0075] 另外,可以修饰本发明的肿瘤抗原肽的 N-末端氨基酸的氨基或 C-末端氨基酸的羧基。经历此修饰的肽也属于本发明的范围。

N-末端氨基酸氨基的修饰包括 1-3 个选自例如 C₁₋₆ 烷基、苯基、环烷基和酰基的基团。具体地说,酰基包括 C₁₋₆ 烷酰基、被苯基取代的 C₁₋₆ 烷酰基、被 C₅₋₇ 环烷基取代的羧基、C₁₋₆ 烷基磺酰基、苯基磺酰基、C₂₋₆ 烷氧基羰基、被苯基取代的烷氧基羰基、被 C₅₋₇ 环烷氧基取代的羧基、苯氧基羰基等。

在 C-末端氨基酸的羧基被修饰的肽可以为酯或酰胺形式。具体地说,所述酯包括 C₁₋₆ 烷基酯、被苯基取代的 C₀₋₆ 烷基酯、C₅₋₇ 环烷基酯等。具体地说,所述酰胺包括酰胺、被 1 个或 2 个 C₁₋₆ 烷基取代的酰胺、被 1 个或 2 个 C₀₋₆ 烷基(其中烷基被苯基取代)取代的酰胺、形成 5-元至 7-元氮杂环烷(包含酰胺基团的氮原子)的酰胺等。

[0076] 3)本发明的核酸

具体地说,本发明的核酸是指

(1)包含编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸的核酸,和

(2)包含编码本发明肽的多核苷酸的核酸。

[0077] (1)编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3

或 HIFPH3 的多核苷酸以及含有该多核苷酸的核酸

编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸可以为多种细胞或组织的 cDNA 或 mRNA、cRNA 或基因组 DNA(例如来源于前列腺癌的那些), 或合成 DNA。其可为单链的或双链的。具体地说, 所述多核苷酸包括以下的:

(a)含有 SEQ ID NO: 1、3、5、7、9 或 11 的碱基序列的多核苷酸;

(b)含有编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的碱基序列的多核苷酸; 和

含有的碱基序列类似于多核苷酸(a)或(b)的碱基序列的多核苷酸。

[0078] 在这方面, SEQ ID NO: 1 的碱基序列对应于人 Lengsin 基因的碱基序列, 其以登录号 NM_016571 向 GenBank 数据库登记。SEQ ID NO: 2 的氨基酸序列对应于人 Lengsin 的氨基酸序列, 其以登录号 NM_016571、登录号 NP_057655 向 GenBank 数据库登记。人 Lengsin 基因公开于非专利文献 7 (Mol. Vis. Jun. 15;8:185-95 (2002))。

[0079] SEQ ID NO: 3 的碱基序列对应于人 BJ-TSA-9 基因的碱基序列, 其以登录号 NM_032899 向 GenBank 数据库登记。SEQ ID NO: 4 的氨基酸序列对应于人 BJ-TSA-9 的氨基酸序列, 其以登录号 NM_032899、登录号 NP_116288 向 GenBank 数据库登记。人 BJ-TSA-9 基因公开于非专利文献 8 (Proc. Natl. Acad. Sci. USA., 99(26): 16899-903 (2002))。

[0080] SEQ ID NO: 5 的碱基序列对应于人 C20orf42 基因的碱基序列, 其以登录号 NM_017671 向 GenBank 数据库登记。SEQ ID NO: 6 的氨基酸序列对应于人 C20orf42 的氨基酸序列, 其以登录号 NM_017671、登录号 NP_060141 向 GenBank 数据库登记。人 C20orf42 基因公开于非专利文献 9 (Biochim. Biophys. Acta 1637: 207-216 (2003))。

[0081] SEQ ID NO: 7 的碱基序列对应于人 BUB1 基因的碱基序列, 其以登录号 NM_004336 向 GenBank 数据库登记。SEQ ID NO: 8

的氨基酸序列对应于人 BUB1 的氨基酸序列，其以登录号 NM_004336、登录号 NP_004327 向 GenBank 数据库登记。人 BUB1 基因公开于文献(Genomics 46:379-388(1997))。

[0082] SEQ ID NO: 9 的碱基序列对应于人 C10orf3 基因的碱基序列，其以登录号 NM_018131 向 GenBank 数据库登记。SEQ ID NO: 10 的氨基酸序列对应于人 C10orf3 的氨基酸序列，其以登录号 NM_018131、登录号 NP_060601 向 GenBank 数据库登记。人 C10orf3 基因公开于非专利文献 16 (Proc. Natl. Acad. Sci. USA., Dec 24; 99(26):16899-903 (2002))。

[0083] SEQ ID NO: 11 的碱基序列对应于人 HIFPH3 基因的碱基序列，其以登录号 NM_022073 向 GenBank 数据库登记。SEQ ID NO: 12 的氨基酸序列对应于人 HIFPH3 的氨基酸序列，其以登录号 NM_022073、登录号 NP_071356 向 GenBank 数据库登记。人 HIFPH3 基因公开于非专利文献 17 (Cell 107: 43-54 (2001))。

[0084] 具体地说，前述的(a)含有 SEQ ID NO: 1、3、5、7、9 或 11 的碱基序列的多核苷酸和(b)含有编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的碱基序列的多核苷酸包括由 SEQ ID NO: 1、3、5、7、9 或 11 的碱基序列组成的多核苷酸和由编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的碱基序列组成的多核苷酸。其它实例包括由含有 SEQ ID NO: 1、3、5、7、9 或 11 的碱基序列或编码 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的碱基序列(在其 5'-和/或 3'-末端添加额外的碱基序列)组成的多核苷酸。“额外的碱基序列”可以为编码非 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 的结构基因的碱基序列。

此编码 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的多核苷酸的特征在于由该多核苷酸编码的蛋白具有作为肿瘤抗原蛋白的活性。其活性和测定方法描述于“1)本发明的蛋白”。

[0085] 可使用在 GenBank 登录号 NM_016571 中公开的或本文

在 SEQ ID NO: 1 中公开的碱基序列的适宜部分作为杂交探针或 PCR 引物, 通过筛选由例如肺腺癌细胞系(如 1-87)获得的 cDNA 文库, 克隆含有 SEQ ID NO: 1 的碱基序列的多核苷酸。

[0086] 可使用在 GenBank 登录号 NM_032899 中公开的或本文在 SEQ ID NO: 3 中公开的碱基序列的适宜部分作为杂交探针或 PCR 引物, 通过筛选由例如肺腺癌细胞系(如 1-87)获得的 cDNA 文库, 克隆含有 SEQ ID NO: 3 的碱基序列的多核苷酸。

[0087] 可使用在 GenBank 登录号 NM_017671 中公开的或本文在 SEQ ID NO: 5 中公开的碱基序列的适宜部分作为杂交探针或 PCR 引物, 通过筛选由例如结肠癌细胞系(例如 SW480 (ATCC 编号: CCL-228))获得的 cDNA 文库, 克隆含有 SEQ ID NO: 5 的碱基序列的多核苷酸。

[0088] 可使用在 GenBank 登录号 NM_004336 中公开的或本文在 SEQ ID NO: 7 中公开的碱基序列的适宜部分作为杂交探针或 PCR 引物, 通过筛选由例如胰腺癌细胞系(如 PUN)获得的 cDNA 文库, 克隆含有 SEQ ID NO: 7 的碱基序列的多核苷酸。

[0089] 可使用在 GenBank 登录号 NM_018131 中公开的或本文在 SEQ ID NO: 9 中公开的碱基序列的适宜部分作为杂交探针或 PCR 引物, 通过筛选由例如肺腺癌细胞系(如 1-87)获得的 cDNA 文库, 克隆含有 SEQ ID NO: 9 的碱基序列的多核苷酸。

[0090] 可使用在 GenBank 登录号 NM_022073 中公开的或本文在 SEQ ID NO: 11 中公开的碱基序列的适宜部分作为杂交探针或 PCR 引物, 通过筛选由例如肾癌细胞系(如 SMKTR-1)获得的 cDNA 文库, 克隆含有 SEQ ID NO: 11 的碱基序列的多核苷酸。

[0091] 本领域一般技术人员可按照 Molecular Cloning, 第 2 版, Cold Spring Harbor Laboratory Press (1989) 等中所述方法容易地进行克隆。

[0092] 具体地说, 所含碱基序列与以上(a)或(b)的多核苷酸的碱

基序列类似的多核苷酸包括以下的:

(c)在严格条件下能够与多核苷酸(a)或(b)的互补链杂交的多核苷酸,其编码的蛋白具有作为肿瘤抗原蛋白的活性;

(d)含有的碱基序列与多核苷酸(a)或(b)具有至少 70%序列同一性的多核苷酸,其编码的蛋白具有作为肿瘤抗原蛋白的活性;和

(e)一种多核苷酸,其编码蛋白含有的氨基酸序列与多核苷酸(a)或(b)编码的氨基酸序列相同,只是缺失、取代和/或添加一个或多个氨基酸,其中所述蛋白具有作为肿瘤抗原蛋白的活性。

[0093] 优选实例包括由与以上的多核苷酸(a)或(b)的碱基序列类似的碱基序列组成的多核苷酸。由与以上的多核苷酸(a)或(b)的碱基序列类似的碱基序列组成的多核苷酸包括以下的多核苷酸(c')-(e'):

(c')在严格条件下能够与多核苷酸(a)或(b)的互补链杂交的多核苷酸,其编码的蛋白具有作为肿瘤抗原蛋白的活性;

(d')由与多核苷酸(a)或(b)具有至少 70%序列同一性的碱基序列组成的多核苷酸,其编码的蛋白具有作为肿瘤抗原蛋白的活性;和

(e')一种多核苷酸,其编码蛋白由与多核苷酸(a)或(b)编码的氨基酸序列相同的氨基酸序列组成,不同的是其中缺失、取代和/或添加一个或多个氨基酸,其中所述蛋白具有作为肿瘤抗原蛋白的活性。

[0094] “在严格条件下能够与以上多核苷酸(a)或(b)的互补链杂交的多核苷酸”的实例包括含有以下碱基序列的多核苷酸:其与以上多核苷酸(a)或(b)的碱基序列具有至少约 40%、优选约 60%、更优选约 70%、再更优选约 80%、还更优选约 90%、最优选约 95%的序列同一性,具体地说,包括由以上多核苷酸(a)或(b)的部分序列组成的多核苷酸。

[0095] 可以按照自身已知的方法或与其等效的方法进行杂交,例如在基础教科书“Molecular Cloning, 第2版, Cold Spring Harbor Laboratory Press (1989)”中描述的方法等。另外,可使用市售可得文库按照其随附的说明书进行杂交。

本文使用的“严格条件”可如在文献(Berger 和 Kimmel, 1987, “Guide to Molecular Cloning Techniques Methods in Enzymology”, 152 卷, Academic Press, San Diego CA; 或 “Molecular Cloning” 第 2 版, Cold Spring Harbor Laboratory Press (1989))中所述基于形成复合物的核酸或结合探针的核酸的解链温度(T_m)确定。

[0096] 例如, 杂交可在含有 6xSSC (20xSSC 对应于 333 mM 柠檬酸钠、333 mM NaCl)、0.5% SDS 和 50% 甲酰胺的溶液于 42°C 进行, 或在含有 6xSSC 的溶液(无 50% 甲酰胺)于 65°C 进行。

杂交后的洗涤可在约 “1xSSC、0.1% SDS、37°C” 的条件下进行。优选地, 在这样的洗涤条件下洗涤时, 互补链保持与靶正向链结合。更严格的杂交条件可包括在约 “0.5xSSC、0.1% SDS、42°C” 的条件下洗涤, 再更严格的杂交条件为约 “0.1xSSC、0.1% SDS、65°C”, 但其不限于此。

[0097] “含有的碱基序列与以上(a)或(b)的多核苷酸具有至少 70% 序列同一性的多核苷酸” 包括含有的碱基序列与以上(a)或(b)的多核苷酸的碱基序列具有至少约 70%、优选约 80%、更优选约 90% 和最优选约 95% 的序列同一性的多核苷酸, 具体地说, 包括由以上(a)或(b)的多核苷酸的部分序列组成的多核苷酸。

[0098] 本文使用的术语“序列同一性”是指两个多核苷酸之间的同一性或同源性。通过比较两个在对应于待比较序列的区域内最优比对的序列确定“序列同一性”。在本文中, 待比较的两个多核苷酸的最佳比对可具有添加或缺失(例如“空位”)。可通过使用例如 Vector NTI、ClustalW 算法(Nucleic Acid Res., 22 (22): 4673-4680(1994))进行比对计算此序列同一性。序列同一性可使用序列分析软件(具体地说为 Vector NTI 或 GENETYX-MAC)或由公共数据库提供的测序工具来测定。这样的公共数据库通常可在 Web 站点(<http://www.ddbj.nig.ac.jp>)获得。

[0099] 具有此序列同一性的多核苷酸可按照前述杂交方法或

常规 PCR 反应或后文描述的用于修饰多核苷酸的反应(缺失、添加或取代)制备。

[0100] “多核苷酸, 其编码蛋白含有的氨基酸序列与以上的多核苷酸(a)或(b)编码的蛋白的氨基酸序列相同, 只是缺失、取代和/或添加一个或多个氨基酸”包括编码人工产生的变体蛋白或活体中存在的等位基因变体的核酸。

在这方面, 对于氨基酸修饰(突变)的数目或位置没有限制, 只要保持本发明蛋白的活性。标准(人们可基于其确定要缺失、取代和/或添加的氨基酸残基的数量或位置而不降低活性)可使用本领域众所周知的计算机程序获得, 例如 DNA Star 软件。例如, 突变数通常应在总氨基酸残基的 10%之内, 优选应在总氨基酸残基的 5%之内。而且, 考虑到结构保持, 通过取代引入的氨基酸优选与被除去的氨基酸具有类似特征, 所述特征例如为极性、电荷、溶解性、疏水性、亲水性、两亲性等。例如, Ala、Val、Leu、Ile、Pro、Met、Phe 和 Trp 被分类为非极性氨基酸; Gly、Ser、Thr、Cys、Tyr、Asn 和 Gln 被分类为不带电的氨基酸; Asp 和 Glu 被分类为酸性氨基酸; 而 Lys、Arg 和 His 被分类为碱性氨基酸。本领域一般技术人员可以基于这些标准选择相同类别内的适宜氨基酸。

[0101] 编码这些变体蛋白的多核苷酸可通过多种方法制备, 例如在 Molecular Cloning, 第 2 版, Cold Spring Harbor Laboratory Press (1989)中描述的定点诱变和 PCR 技术。其还可以使用市售可得的试剂盒通过已知方法诸如 Gapped duplex 或 Kunkel 法制备。

[0102] 编码如上所述的本发明蛋白的多核苷酸编码具有作为肿瘤抗原蛋白活性的蛋白。“具有作为肿瘤抗原蛋白的活性”是指蛋白在用于肿瘤抗原蛋白活性的常规测定中为阳性。具体地说, 其是指以下特征: 表达编码本发明蛋白的多核苷酸的细胞被 CTL 识别, 即细胞表现出对 CTL 的反应性, 换句话说, 本发明的蛋白或由其获得的肿瘤抗原肽活化或诱导 CTL。其活性和测定方法描述于以上的“1)本发

明的蛋白”。

[0103] 含有本发明多核苷酸的核酸可以为单链的或双链的，并可以为 DNA 或 RNA。当本发明的多核苷酸为双链时，可以通过将上述多核苷酸掺入表达载体中构建用于表达本发明蛋白的表达载体。因此，本发明的核酸包括通过将本发明的双链多核苷酸插入表达载体构建的重组表达载体。

[0104] 适宜的表达载体可以根据要使用的宿主、用途等选择，并包括质粒、噬菌体载体、病毒载体等。

当宿主为大肠杆菌时，载体可以为质粒载体，如 pUC118、pUC119、pBR322 和 pCR3；或噬菌体载体，如 λ ZAPII 和 λ gt11。当宿主为酵母时，载体可以为 pYES2、pYEUra3 等。当宿主为昆虫细胞时，载体可以为 pAcSGHisNT-A 等。当宿主为动物细胞时，载体可以为质粒载体，例如 pCEP4、pKCR、pCDM8、pGL2、pcDNA3.1、pRc/RSV 和 pRc/CMV；或者为病毒载体，例如逆转录病毒载体、腺病毒载体和腺相关病毒载体。

表达载体可以任选地包含诸如能够诱导表达的启动子、编码信号序列的基因、用于选择的标记基因和终止子的因子。

[0105] 此外，表达载体可以含有额外的序列，用于将蛋白表达为具有硫氧还蛋白、His-标签、GST (谷胱甘肽 S-转移酶)等的融合蛋白，以便有利于蛋白的分离和纯化。在此情况下能够使用的载体包括含有在宿主细胞如 pGEX4T 中有功能的适宜启动子(lac、tac、trc、trp、CMV、SV40 早期启动子等)的 GST 融合蛋白载体；含有 Tag 序列(Myc、His 等)的载体，如 pcDNA3.1/Myc-His；以及能够表达具有硫氧还蛋白和 His 的融合蛋白的载体，如 pET32a。

通过用在以上获得的表达载体转化宿主细胞，可以制备含有本发明载体的转化体。

[0106] 本文可以使用的宿主细胞包括大肠杆菌、酵母、昆虫细胞和动物细胞。大肠杆菌的实例包括大肠杆菌 K-12 的菌株，例如

HB101、C600、JM109、DH5 α 和 AD494 (DE3)。酵母的实例包括酿酒酵母 (*Saccharomyces cerevisiae*)。动物细胞的实例包括 L929、BALB/c3T3、C127、CHO、COS、Vero、Hela 和 293-EBNA 细胞。昆虫细胞的实例包括 sf9。

可以使用适于以上各自对应的宿主细胞的常规方法将表达载体导入宿主细胞中。具体地说，其可以用磷酸钙法、DEAE-葡聚糖法、电穿孔法或使用用于基因转移的脂质 (Lipofectamine, Lipofectin; Gibco-BRL) 的方法进行。在导入后，在含有选择标记的常规培养基中培养细胞，由此可以选择含有表达载体的转化体。

[0107] 可以通过在合适的条件下培养转化体生产本发明的蛋白。产生的蛋白可以按照标准生物化学方法进一步分离并纯化。纯化方法包括盐析、离子交换层析、吸附层析、亲和层析、凝胶过滤层析等。当本发明的蛋白表达为如上所述的具有硫氧还蛋白、His 标签、GST 等的融合蛋白时，其可以通过利用这些融合蛋白或标签的特征的适宜纯化方法分离和纯化。

[0108] (2)编码本发明肽的多核苷酸和含有该多核苷酸的核酸如上所述，含有编码本发明肽的多核苷酸的核酸属于本发明核酸的范围。

编码本发明肽的多核苷酸可以为 DNA 或 RNA，并可以为单链的或双链的。编码本发明肽的多核苷酸可以基于关于肽的氨基酸序列或编码其的 DNA 的信息容易地制备。具体地说，其可以通过常规方法如 DNA 合成或 PCR 扩增来制备。

[0109] 具体地说，编码本发明肽的多核苷酸包括编码如上所述的表位肽的多核苷酸。

含有编码本发明肽的多核苷酸的核酸可以为单链的或双链的，并还可以为 DNA 或 RNA。当本发明的多核苷酸为双链时，用于表达本发明肽(表位肽)的重组表达载体可以通过将上述多核苷酸导入表达载体中构建。

本文使用的表达载体、宿主细胞、转化宿主细胞的方法等类似于以上(1)中所述的那些。

[0110] 4)本发明的抗原提呈细胞

可通过使具有抗原提呈能力的细胞在体外与上述的任一种本发明的蛋白、肽和核酸接触来制备抗原提呈细胞。具体地说，本发明提供制备抗原提呈细胞的方法，其特征在于使分离自肿瘤患者的具有抗原提呈能力的细胞在体外与任一种本发明的蛋白、肽和核酸接触，并由此制备抗原提呈细胞。

[0111] 在本文中，“具有抗原提呈能力的细胞”不限于具体的细胞，可为任何在表面上表达能够提呈本发明肽的 HLA 抗原的细胞；然而，优选已知具有特别高的抗原提呈能力的树突细胞。

此外，本发明的任何蛋白、肽和核酸均可用于由具有抗原提呈能力的细胞制备本发明的抗原提呈细胞。

[0112] 本发明的抗原提呈细胞可以如下制备：由肿瘤患者分离具有抗原提呈能力的细胞，用本发明的蛋白或肽在体外脉冲所述细胞，并使所述细胞提呈 HLA 抗原和本发明肽之间的复合物(Cancer Immunol. Immunother., 46: 82, 1998; J. Immunol. 158: 1796 页, 1997; Cancer Res., 59:1184, 1999)。在使用树突细胞时，本发明的抗原提呈细胞例如可如下制备：使用 Ficoll 法由肿瘤患者的外周血分离淋巴细胞，除去非粘附细胞，在 GM-CSF 和 IL-4 存在下温育粘附细胞，以诱导树突细胞，并用本发明的蛋白或肽温育和脉冲所述树突细胞。

当本发明的抗原提呈细胞是通过将本发明的核酸导入具有抗原提呈能力的细胞中制备时，核酸可为 DNA 或 RNA 形式。具体地说，可以按照在 Cancer Res., 56:5672, 1996 或 J. Immunol., 161: 5607 页, 1998 中的教导使用 DNA，并按照例如 J. Exp. Med., 184:465 页, 1996 中的教导使用 RNA。

[0113] 本发明的抗原提呈细胞的特征在于，该细胞提呈本发明的肽和 HLA 抗原之间的复合物，具体地说，包括其为树突细胞并且

在细胞表面上提呈由以下氨基酸序列组成的肽和 HLA-A24 抗原之间的复合物的抗原提呈细胞,所述氨基酸序列如下: SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 或 183 的氨基酸序列(优选为 158、22、23、26、27、198、200、201、177、178、179 或 183 的氨基酸序列)。这样的抗原提呈细胞可以如下制备: 由 HLA-A24⁺前列腺癌患者分离具有抗原提呈能力的细胞, 用由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 或 183 的氨基酸序列(优选为 158、22、23、26、27、198、200、201、177、178、179 或 183 的氨基酸序列)组成的肽体外脉冲所述细胞, 并使所述细胞提呈所述肽和 HLA-A24 抗原之间的复合物。

[0114] 5)本发明的 CTL

任一种本发明的蛋白、肽和核酸在与外周血淋巴细胞接触时均可以用于在体外诱导 CTL。具体地说, 本发明提供诱导 CTL 的方法, 其中使得自肿瘤患者的外周血淋巴细胞与任一种本发明的蛋白、肽和核酸在体外接触, 并由此诱导 CTL。

[0115] 对于黑素瘤, 过继免疫治疗已显示出疗效, 其中大量离体培养得自患者的肿瘤浸润 T 细胞, 并返回到同一患者中(J. Natl. Cancer. Inst., 86: 1159, 1994)。此外, 在小鼠黑素瘤中, 通过用肿瘤抗原肽 TRP-2 在体外刺激脾细胞, 以诱导对肿瘤抗原肽特异性的 CTL 增殖, 并将 CTL 给予黑素瘤移植小鼠, 已观察到转移抑制(J. Exp. Med., 185:453, 1997)。这由特异性识别抗原提呈细胞上的 HLA 抗原和肿瘤抗原肽之间的复合物的 CTL 的体外增殖产生。因此, 一般认为这样的治疗方法是有效的: 该方法包括用本发明的蛋白、肽或核酸在体外刺

激患者的外周血淋巴细胞，以增殖肿瘤特异性 CTL，并将 CTL 返回患者中。

[0116] 用于过继免疫治疗的 CTL 可以如下制备：由患者分离外周血淋巴细胞，并用本发明的蛋白、肽或核酸在体外刺激淋巴细胞 (Journal of Experimental Medicine 1999, 190:1669)。

[0117] 本发明的 CTL 的特征在于其通过使外周血淋巴细胞与任一种本发明的蛋白、肽和核酸体外接触而被诱导，并可以为单个 CTL 克隆或由多个 CTL 克隆组成的 CTL 混合物(群)。此 CTL 的具体实例为特异性识别由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 或 183 的氨基酸序列组成的肽和 HLA-A24 抗原之间的复合物的 CTL。此 CTL 的更优选的实例为特异性识别由 SEQ ID NO: 158、22、23、26、27、198、200、201、177、178、179 或 183 的氨基酸序列组成的肽和 HLA-A24 抗原之间的复合物的 CTL。

[0118] 6) 本发明的药物组合物

如上所述的本发明的蛋白、肽、核酸、抗原提呈细胞和 CTL 可以为对每种物质适宜形式的药物组合物的活性成分。本发明的药物组合物可以为 CTL 诱导物的活性成分，即癌症疫苗，详见下文。

[0119] (1) 含有本发明的蛋白作为活性成分的 CTL 诱导物

本发明的蛋白(Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3、HIFPH3)具有诱导 CTL 的活性，因此，本发明的蛋白可以用作肿瘤治疗或预防药物的活性成分(癌症疫苗)。因此，含有本发明的蛋白作为活性成分的 CTL 诱导物针对肿瘤发挥治疗或预防作用。所述蛋白在给予肿瘤患者时，被抗原提呈细胞掺入，并在胞内降解；通过胞内降解产生的肿瘤抗原肽结合 HLA 抗原，形成复合物；然后该复合物被提呈至抗原提呈细胞表面；对该复合物特异性的 CTL 在体内有效增殖，

并破坏肿瘤细胞。以此方式实现肿瘤的治疗或预防。

[0120] 可以将含有本发明蛋白作为活性成分的 CTL 诱导物给予 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 蛋白为阳性的任何肿瘤患者。

具体地说,含有 Lengsin 作为活性成分的 CTL 诱导物可以用于预防或治疗癌症(肿瘤),例如肺腺癌、肺鳞状细胞癌或胃癌。

含有 BJ-TSA-9 作为活性成分的 CTL 诱导物可以用于预防或治疗癌症(肿瘤),例如白血病、肺腺癌、肺鳞状细胞癌、小细胞肺癌、口腔癌、胃癌、胰腺癌或淋巴瘤。

[0121] 含有 C20orf42 作为活性成分的 CTL 诱导物可以用于预防或治疗癌症(肿瘤),例如肺鳞状细胞癌、肺腺癌、肝癌、胃癌、白血病、恶性淋巴瘤组织、直肠癌、结肠癌或胰腺癌。

含有 BUB1 作为活性成分的 CTL 诱导物可以用于预防或治疗癌症(肿瘤),例如乳腺癌、肺腺癌、肺鳞状细胞癌、卵巢癌、口腔鳞状细胞癌、肾癌、大肠癌(结肠癌、直肠癌)、胃癌、胰腺癌、肝癌、白血病、淋巴瘤或黑素瘤。

[0122] 含有 C10orf3 作为活性成分的 CTL 诱导物可以用于预防或治疗癌症(肿瘤),例如乳腺癌、结肠癌、直肠癌、肾癌、胃癌、卵巢癌、肝癌、胰腺癌、肺鳞状细胞癌、肺腺癌、小细胞肺癌或黑素瘤。

含有 HIFPH3 作为活性成分的 CTL 诱导物可以用于预防或治疗癌症(肿瘤),例如乳腺癌、结肠癌、胃癌、肾癌、胰腺癌、肝癌、肺腺癌或肺鳞状细胞癌。

[0123] 含有本发明蛋白作为活性成分的 CTL 诱导物可以作为与药学可接受的载体(例如适宜的佐剂)的混合物给予,或与该载体一起给予,以便能有效建立细胞免疫。

[0124] 可用的佐剂的实例包括描述于文献(Clin. Microbiol. Rev., 7:277-289, 1994)中的那些。具体地说,包括以下的:微生物来源的组分或其衍生物、细胞因子、植物来源的组分或其衍生物、海洋生

物来源的组分或其衍生物、诸如氢氧化铝的矿物凝胶、表面活性剂(如溶血卵磷脂及 Pluronic®多元醇、聚阴离子、肽、油乳化剂(乳化剂))等。另外,还包括脂质体制剂、其中成分结合在几 μm 直径的珠子上的颗粒制剂、其中成分附着至油脂的制剂、微球制剂和微囊。

[0125] 在本文中,“微生物来源的组分或其衍生物”可以被具体地分类为(1)被杀死的细菌,(2)来源于细菌的细胞壁骨架(后文为“CWS”),和(3)微生物来源的特定组分及其衍生物。

(1) 被杀死的细菌的实例包括粉状溶血性链球菌(例如 Picibanil®, Chugai Co., Ltd.)、被杀死的细菌混悬液的混合物(例如 Broncasma Berna®, Sanwa Kagaku Kenkyusho Co., Ltd)或被杀死的结核分枝杆菌细菌。

(2) 细菌来源的 CWS 的实例包括得自分枝杆菌的 CWS (例如牛分枝杆菌 CWS)、得自诺卡氏菌的 CWS (例如红色诺卡氏菌 CWS)、棒状杆菌 CWS。

[0126] (3) 微生物来源的具体组分及其衍生物的实例包括微生物来源的多糖,例如来自结核分枝杆菌的多糖(例如 Ancer®, Zeria Pharmaceutical Co., Ltd.); 来自担子菌类的多糖(Lentinan®, Ajinomoto, Co., Ltd.; Krestin®, Sankyo, Co., Ltd.; Basidiomycetes, Coriolus versicolor (Fr) Quel); 胞壁酰二肽(MDP)相关化合物; 脂多糖(LPS); 脂质 A (MPL)相关化合物; 糖脂海藻糖二霉菌酸酯(TDM); 细菌 DNA (例如 CpG 寡核苷酸); 及其衍生物。

这些微生物来源的组分及其衍生物可得自商业来源,或者可以按照在已知文献(例如 Cancer Res., 33, 2187-2195 (1973); J. Natl. Cancer Inst., 48, 831-835(1972), J. Bacteriol., 94, 1736-1745 (1967); Gann, 69, 619-626 (1978), J. Bacteriol., 92, 869-879 (1966)或 J. Natl. Cancer Inst., 52, 95-101 (1974))中描述的方法生产和分离。

[0127] 术语“细胞因子”指例如 IFN-α、IL-12、GM-CSF、IL-2、IFN-γ、IL-18 或 IL-15。细胞因子可以为天然产物或基因工程产物。当

细胞因子市售可得时，人们可以购买和使用该细胞因子。或者，可如下重组制备细胞因子：基于向数据库如 GenBank、EMBL 或 DDBJ 登记的碱基序列以常规方式克隆所需基因，将该基因连接入适宜的表达载体中，用所获重组表达载体转化宿主细胞，并使所述细胞表达和生产预期的细胞因子。

[0128] “植物来源的组分或其衍生物”的实例包括皂苷来源的组分 Quil A (Accurate Chemical & Scientific Corp)、QS-21 (Aquila Biopharmaceuticals Inc.)或甘草甜素(SIGMA-ALDRICH 等)。

“海洋生物来源的组分或其衍生物”的实例包括海绵来源的糖脂 α -半乳糖苷神经酰胺。

油乳化剂(乳化制剂)的实例包括油包水型(w/o)、水包油型(o/w)和水包油包水型(w/o/w)的乳化制剂。

[0129] 在油包水型(w/o)乳化制剂中，活性成分分散在用作分散相的水中。在水包油型(o/w)乳化制剂中，活性成分分散在用作分散介质的水中。此外，在水包油包水型(w/o/w)乳化制剂中，活性成分分散在为最内相的水中。这些乳化制剂可以按照在例如 JP-A-8-985、JP-A-9-122476 等中的教导生产。

[0130] “脂质体制剂”是指微粒，其中活性成分囊化在具有脂质双层结构的脂质体的水相中或脂质双层内。用于制备脂质体的代表性脂质包括磷脂酰胆碱和神经鞘磷脂。还可以加入赋予电荷的磷酸二鲸蜡酯、磷脂酸、磷脂酰丝氨酸等，用于稳定脂质体。生产脂质体的方法包括超声法、乙醇注入法、乙醚注入法、反相蒸发法、高压破碎提取法(French press extraction method)等。

[0131] “微球制剂”是指由均一聚合物基质组成的微粒，其中活性成分分散在基质中。该基质可以由生物可降解的聚合物组成，所述生物可降解的聚合物例如为白蛋白、明胶、几丁质、壳聚糖、淀粉、聚乳酸、聚氰基丙烯酸烷基酯等。微球制剂可以通过任何已知的方法制备，所述方法包括但不限于在文献(Eur. J. Pharm. Biopharm.

50:129-146, 2000; Dev. Biol. Stand. 92:63-78, 1998; Pharm. Biotechnol. 10:1-43, 1997 等)中描述的那些。

[0132] “微囊制剂”是指含有活性成分作为用膜包封的核心物质的微粒。用于膜的包衣物质包括成膜聚合物, 例如羧甲基纤维素、醋酸纤维素酞酸酯、乙基纤维素、明胶、明胶/阿拉伯胶、硝酸纤维素、聚乙烯醇、羟丙基纤维素等。微囊制剂可以通过凝聚法、表面聚合等制备。

例如可以皮内、皮下、肌内或静脉内实现给药。尽管在要给予的制剂中的本发明蛋白的剂量可以根据例如要治疗的疾病、患者的年龄和体重进行适当调整, 但其通常在 0.0001 mg-1000 mg、优选 0.001 mg-100 mg、更优选 0.01 mg-10 mg 的范围内, 其可以每几天至每几个月给药一次。

[0133] (2)含有本发明肽作为活性成分的 CTL 诱导物

本发明的肽具有诱导 CTL 的活性。诱导的 CTL 可以通过淋巴因子的细胞毒性作用或生产发挥抗肿瘤作用。因此, 本发明的肽可以用作治疗或预防肿瘤的药物(癌症疫苗)的活性成分。当将含有本发明肽作为活性成分的 CTL 诱导物给予肿瘤患者时, 本发明的肽被提呈至抗原提呈细胞中的 HLA 抗原。然后, 所提呈的 HLA 抗原和本发明肽之间的结合复合物的特异性 CTL 增殖, 其后破坏肿瘤细胞。以此方式可在患者中实现肿瘤的治疗或预防。

含有本发明肽作为活性成分的 CTL 诱导物可以给予任何对本发明蛋白为阳性的肿瘤患者。具体地说, 其可用于预防或治疗如在以上 6)-(1)中所述的癌症(肿瘤)。

[0134] 包含本发明肽作为活性成分的 CTL 诱导物可以包含单个 CTL 表位(本发明的肽)或由本发明肽和其它肽(CTL 表位或辅助表位)连接在一起组成的表位肽作为活性成分。最近, 业已表明, 由多个(多于 1 个) CTL 表位(抗原肽) 连接在一起组成的表位肽具有有效诱导 CTL 的活性。例如, 已有报道指出, 由起源于肿瘤抗原蛋白 PSA、各

自限于 HLA-A2-、-A3-、-A11 或 B53 的 CTL 表位相连接组成的约 30 聚体的表位肽诱导对各自的 CTL 表位特异性 CTL (Journal of Immunology 1998, 161: 3186-3194)。另外, 业已报道, 由相连的 CTL 表位和辅助表位组成的表位肽可以有效诱导 CTL。当本发明的肽以表位肽的形式给予时, 该肽被掺入抗原提呈细胞中; 然后通过胞内降解产生的抗原肽结合 HLA 抗原, 形成复合物; 复合物被以高密度提呈至抗原提呈细胞表面; 对该复合物特异性的 CTL 在体内有效增殖, 并破坏肿瘤细胞。以此方式实现肿瘤的治疗或预防。

[0135] 包含本发明肽作为活性成分的 CTL 诱导物的具体实例包括含有由 SEQ ID NOS: 13-201 的任一个氨基酸序列组成的肿瘤抗原肽作为活性成分的 CTL 诱导物。CTL 诱导物的优选实例为含有由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 和 183 的任一个氨基酸序列组成的肽作为活性成分的 CTL 诱导物。CTL 诱导物的更优选实例为含有由 SEQ ID NO: 158、22、23、26、27、198、200、201、177、178、179 和 183 的任一个氨基酸序列组成的肽作为活性成分的 CTL 诱导物。

[0136] 含有本发明肽作为活性成分的 CTL 诱导物可以在与药学可接受的载体(例如适宜的佐剂)的混合物中给予或与该载体一起给予, 使得可有效建立细胞免疫。

可用的佐剂的实例包括描述于文献(Clin. Microbiol. Rev., 7:277-289, 1994)中的那些。具体地说, 包括以下的: 微生物来源的组分或其衍生物、细胞因子、植物来源的组分或其衍生物、海洋生物来源的组分或其衍生物、诸如氢氧化铝的矿物凝胶、表面活性剂(如溶血卵磷脂及 Pluronic®多元醇、聚阴离子、肽、油乳化剂(乳化制剂))等。另外, 还包括脂质体制剂、其中成分结合在几 μm 直径的珠子上的颗

粒制剂、其中成分附着至油脂的制剂、微球制剂和微囊。这些佐剂的具体实例与在以上“6)-1)含有本发明蛋白作为活性成分的 CTL 诱导物”中所述的那些相同。

[0137] 例如可以皮内、皮下、肌肉或静脉内实现给药。尽管本发明肽在要给予的制剂中的剂量可以根据例如要治疗的疾病、患者的年龄和体重进行适当调整,但其通常在 0.0001 mg-1000 mg、优选 0.001 mg-100 mg、更优选 0.1 mg-10 mg 的范围内,其可以每几天至每几个月给药一次。

[0138] (3)含有本发明的核酸作为活性成分的 CTL 诱导物

本发明的核酸具有诱导 CTL 的活性。由此可以为用作治疗或预防肿瘤的药物(癌症疫苗)的活性成分。含有本发明核酸作为活性成分的 CTL 诱导物在给予肿瘤患者时,可以通过表达所述核酸对肿瘤发挥治疗性或预防性作用。

[0139] 例如,当以下述方式将掺入到表达载体中的本发明核酸给予肿瘤患者时,肿瘤抗原蛋白在抗原提呈细胞中高度表达。此后,通过胞内降解产生的肿瘤抗原肽与 HLA 抗原形成复合物;然后复合物以高密度被提呈至抗原提呈细胞的表面;肿瘤特异性 CTL 在体内有效增殖,并破坏肿瘤细胞。以此方式实现肿瘤的治疗或预防。

含有本发明的核酸作为活性成分的 CTL 诱导物可以给予任何对本发明蛋白为阳性的肿瘤患者。具体地说,其可用于预防或治疗如在以上 6)-(1)中所述的癌症(肿瘤)。

[0140] 可以通过使用病毒载体或通过其它程序(Nikkei-Science, April, 1994, 20-45 页; Gekkan-Yakuji, 36(1), 23-48 (1994); Jikken-Igaku-Zokan, 12(15), 1994, 以及其中提及的参考文献)实现将本发明的核酸给予和导入细胞中。

导入病毒载体的方法可以包括将本发明的 DNA 掺入 DNA 或 RNA 病毒中,例如逆转录病毒、腺病毒、腺相关病毒、疱疹病毒、痘苗病毒、痘病毒、脊髓灰质炎病毒或新培斯病毒,并将所述病毒导入

细胞中。其中尤其优选涉及逆转录病毒、腺病毒、腺相关病毒或痘苗病毒的方法。

[0141] 除了以上方法，还例举了其中直接肌肉注射表达质粒的方法(DNA 接种)、脂质体法、脂转染法、微注射、磷酸钙法和电穿孔，在它们之中特别优选 DNA 接种和脂质体法。

在实践中，例如在其中将核酸直接导入体内的体内方法中，或在其中将核酸体外导入得自人类受试者的某些细胞中并将细胞再导入受试者体内的离体方法中，本发明的核酸可以用作药物(Nikkei-Science, April, 1994, 20-45 页; Gekkan-Yakuji, 36(1), 23-48 (1994); Jikkenn-Igaku-Zokan, 12(15), 1994; 以及其中提及的参考文献)。更优选体内方法。

[0142] 就体内方法而言，根据要治疗的疾病和症状以及其它因素，可通过任何适宜的途径实现给药。例如，其可以经由静脉内、动脉内、皮下、皮内、肌肉内途径等给予。当以体内方法给予时，本发明的核酸可以配制为液体制剂，通常配制为含有本发明核酸作为活性成分的可注射形式，如有需要，还可以向其加入在药学上可接受的载体。对于含有本发明核酸的脂质体或膜融合脂质体(例如仙台病毒(HVJ)-脂质体)，可以接受混悬剂、冷冻制剂、离心浓缩的冷冻制剂等形式的脂质体制剂。

尽管在要给予的制剂中的本发明核酸的剂量可以根据例如要治疗的疾病、患者的年龄和体重进行适当调整，但核酸中的多核苷酸的量通常在 0.0001 mg-100 mg、优选 0.001 mg-10mg 的范围内，其可以每几天至每几个月给药一次。

[0143] 最近，已表明编码多个(1 个以上) CTL 表位(肿瘤抗原肽)相连接或者 CTL 表位与辅助表位相连接组成的表位肽的多核苷酸在体内有效诱导 CTL。例如，已报道一种编码由 6 种 HBV 来源的 HLA-A2 限制性抗原肽、3 种 HLA-A11 限制性抗原肽及辅助表位相连接组成的表位肽的 DNA (微基因)在体内有效诱导针对各自表位的 CTL (Journal

of Immunology 1999, 162: 3915-3925)。

因此，通过连接一个或多个各自编码本发明肽的多核苷酸和任选的编码不同肽的其它多核苷酸制备的多核苷酸，在导入到适宜的表达载体中时，可以为 CTL 诱导物的活性成分。这样的 CTL 诱导物可以以上述的相同给药方式或形式施用。

[0144] (4)含有本发明的抗原提呈细胞作为活性成分的 CTL 诱导物

本发明的抗原提呈细胞具有诱导 CTL 的活性，并因此可以为用于治疗或预防肿瘤的药物(癌症疫苗)的活性成分。含有本发明的抗原提呈细胞作为活性成分的 CTL 诱导物可以通过给予肿瘤患者抗原提呈细胞对肿瘤发挥治疗性或预防性作用。

含有本发明的抗原提呈细胞作为活性成分的 CTL 诱导物可以给予对本发明蛋白为阳性的任何肿瘤患者。具体地说，其可用于预防或治疗如以上 6)-(1)所述的癌症(肿瘤)。

[0145] 含有抗原提呈细胞作为活性成分的 CTL 诱导物优选含有生理盐水、磷酸缓冲盐水(PBS)、媒介物等，以稳定保持抗原提呈细胞。其可以例如静脉内、皮下或皮内给予。在先前的文献中例举了诱导物的剂量。将含有抗原提呈细胞作为活性成分的 CTL 诱导物再导入对本发明蛋白为阳性的患者中，可以在患者体内产生特异性 CTL 的有效诱导，并产生肿瘤治疗作用。

[0146] (5)含有本发明的 CTL 作为活性成分的癌症疫苗

本发明的 CTL 具有抗肿瘤细胞的细胞毒活性，并因此可以为用于治疗或预防肿瘤的药物(癌症疫苗)的活性成分。

含有本发明的 CTL 作为活性成分、用于治疗或预防肿瘤的药物组合物可以给予对本发明蛋白为阳性的任何肿瘤患者。具体地说，其可用于预防或治疗如以上 6)-(1)中所述的癌症(肿瘤)。

含有本发明的 CTL 作为活性成分、用于治疗或预防肿瘤的药物组合物优选含有生理盐水、磷酸缓冲盐水(PBS)、媒介物等，以稳定保

持 CTL。其可以例如静脉内、皮下或皮内给予。将含有本发明的 CTL 作为活性成分的药物组合物再导入对本发明蛋白为阳性的患者中，可以对患者体内抗肿瘤细胞的 CTL 的细胞毒活性产生促进作用，之后破坏肿瘤细胞，产生肿瘤治疗作用。

[0147] 7)针对本发明肽的抗体

本发明提供特异性结合本发明肽的抗体。就形式而言，本发明的抗体不受限制，可以为针对本发明肽产生的多克隆抗体或单克隆抗体。

本发明的抗体在任何意义上均可不受限制，只要其特异性结合如上所述的本发明肽，具体实例是特异性结合肿瘤抗原肽的抗体，所述肿瘤抗原肽由 SEQ ID NO: 13-201 中的任一个氨基酸序列组成，优选由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 或 183 的氨基酸序列组成，更优选由 158、22、23、26、27、198、200、201、177、178、179 或 183 的氨基酸序列组成。

[0148] 制备抗体的方法在本领域众所周知，本发明的抗体可以按照这些常规方法(Current protocols in Molecular Biology, Ausubel 等人编辑, (1987) Publish. John Wiley and Sons. 11.12-11.13 章, Antibodies; A Laboratory Manual, Lane, H, D.等人编辑, Cold Spring Harber Laboratory Press, New York 1989)中的任一种制备。

[0149] 具体地说,可以通过使用本发明的肽(例如由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 和 183 中的任一个氨基酸序列组成的肽)作为免疫原免疫非人动物如兔,并以常规方式由免疫动物血清回收

抗体，获得本发明的抗体。当抗体为单克隆抗体时，其可如下获得：用本发明的肽(例如由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 和 183 中的任一个氨基酸序列组成的肽)免疫非人动物如小鼠，使所获脾细胞与骨髓瘤细胞进行细胞融合，以制备杂交瘤细胞，并由杂交瘤细胞回收抗体(Current protocols in Molecular Biology, Ausubel 等人编辑, (1987) Publish. John Wiley and Sons. 11.4-11.11 章)。

[0150] 针对本发明肽的抗体还可以在根据宿主使用不同佐剂增强免疫应答的同时生产。佐剂的实例包括弗氏佐剂；诸如氢氧化铝的矿物凝胶；表面活性剂，如溶血卵磷脂及 Pluronic®多元醇、聚阴离子、肽、油乳化剂、匙孔蛾血蓝蛋白和二硝基苯酚；人用佐剂，如 BCG (卡介苗)和棒状杆菌。

[0151] 如上所述，识别本发明肽的抗体和中和该肽活性的抗体可易于通过以常规方式免疫动物制备。所述抗体可用于亲和层析、免疫诊断方法等。免疫诊断方法可适当地选自免疫印迹、放射性免疫测定(RIA)、酶联免疫吸附测定(ELISA)、荧光或发光测定等。免疫诊断方法应有效诊断表达编码本发明蛋白的基因的癌症。

[0152] 8)本发明的 HLA 单体、HLA 二聚体、HLA 四聚体和 HLA 五聚体

本发明还提供含有本发明的肿瘤抗原肽和 HLA 抗原的 HLA 单体、HLA 二聚体、HLA 四聚体或 HLA 五聚体。

在癌症免疫疗法中，可通过在治疗前检验患者中针对肿瘤抗原(肿瘤抗原肽)的 CTL 前体细胞的频率和量，或在用肿瘤抗原(肿瘤抗原肽)进行治疗的患者中检验 CTL 的频率或量，获得用于选择对肿瘤抗原(肿瘤抗原肽)高度响应的患者、监测疗效或评价治疗适用性的主要指示剂。各自均含有肿瘤抗原肽和 HLA 抗原的 HLA 单体、HLA 二聚

体、HLA 四聚体和 HLA 五聚体可在对抗原(抗原肽)特异性的 CTL 检测中用作试剂,具体地说,可用于检测 CTL 的频率和量。

[0153] 本文使用的 HLA 四聚体是指如下制备的四聚体:生物素化由缔合肽(抗原肽)的 HLA 抗原 α -链和 β 2-微球蛋白组成的复合物(HLA 单体),并使其结合抗生物素蛋白,用于四聚化(Science 279: 2103-2106 (1998); 和 Science 274: 94-96 (1996))。

HLA 单体是用于制备上述 HLA 四聚体的单体,通过生物素化缔合的 HLA 抗原 α -链、 β 2-微球蛋白和抗原肽而形成。

[0154] HLA 二聚体是如下制备的二聚体:融合 HLA 抗原 α -链和 Ig (免疫球蛋白,例如 IgG1),并使所获融合体与 β 2-微球蛋白和抗原肽结合(Proc. Natl. Acad. Sci. USA 90: 6671-6675 (1993))。对结合 HLA 二聚体的抗原肽特异性的 CTL 可通过例如使标记的抗 IgG1 抗体结合 IgG1 来检测。

HLA 五聚体是最近开发的技术,是指含有 5 个分子复合物的五聚体,该复合物为通过卷曲螺旋结构域聚合的 HLA 抗原和抗原肽的复合物。因为 HLA 抗原和抗原肽的复合物可以用荧光等标记,所以可以类似于 HLA 四聚体通过流式细胞术等进行分析(参见 <http://www.proimmune.co.uk/>)。

[0155] 如上所述的 HLA 单体、二聚体、四聚体和五聚体全部都可以通过向诸如 ProImmune 或 BD Biosciences 的生产商定制生产而获得。目前,含有不同抗原肽的 HLA 四聚体等市售可得(Medical & Biological Laboratories Co., Ltd.等)。

[0156] 具体地说,本发明的 HLA 单体、二聚体、四聚体和五聚体的实例包括 HLA 单体、二聚体、四聚体和五聚体,其各自均含有由 SEQ ID NO: 42、43、44、45、46、47、49、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、158、159、160、161、15、16、21、22、23、24、25、26、27、30、195、197、198、199、200、201、177、178、179 和 183 的氨基

酸序列、优选 158、22、23、26、27、198、200、201、177、178、179 和 183 的氨基酸序列和 HLA-A24 抗原组成的肽。尤其优选 HLA 四聚体或 HLA 五聚体来检测 CTL。

[0157] 优选用荧光标记 HLA 单体、HLA 四聚体和 HLA 五聚体,使得结合的 CTL 可易于被分拣出来,或易于通过已知检测方法(如流式细胞术、荧光显微镜等)来检测。实例包括用藻红蛋白(PE)、异硫氰酸荧光素(FITC)、多甲藻叶绿素蛋白(PerCP)、别藻蓝蛋白(APC)等标记的 HLA 单体、四聚体和五聚体。

[0158] 在可以为本发明的 HLA 单体、二聚体、四聚体和五聚体的组分的 HLA 抗原中,HLA-A24 抗原(HLA-A24 抗原的 α -链)可基于在 Genbank 登录号 M64740 中公开的 HLA-A2402 的已知碱基序列通过常规方法如 PCR 容易地克隆。HLA-A2 抗原(HLA-A2 抗原的 α -链)可基于在 Genbank 登录号 M84379 中公开的 HLA-A0201 基因的已知碱基序列通过常规方法如 PCR 容易地克隆。

[0159] 为本发明的 HLA 单体、二聚体、四聚体和五聚体的组分的 β 2-微球蛋白优选源于人。人 β 2-微球蛋白可基于在 Genbank 登录号 AB021288 中公开的人 β 2-微球蛋白的已知碱基序列通过常规方法如 PCR 容易地克隆。

[0160] 由上述的对应文献公知用于制备 HLA 单体、二聚体、四聚体和五聚体的方法。有关 HLA-A24 的 HLA 四聚体的制备简述如下。

首先,用 HLA-A24 α -链表达载体和 β 2-微球蛋白表达载体转化适宜的宿主细胞,如能够表达蛋白的大肠杆菌或哺乳动物细胞,并使其表达它们。在本文优选使用大肠杆菌(例如 BL21)。然后将所获的单体形式的 HLA-A24 复合物和本发明的肽混合,形成可溶性 HLA-肽复合物。用 BirA 酶生物素化 HLA-肽复合物的 HLA-A24 α -链的 C-末端序列。当生物素化的 HLA-肽复合物和荧光标记的抗生物素蛋白以 4:1 的摩尔比率混合时,形成 HLA 四聚体。优选通过凝胶过滤等纯化以

上每个步骤所获的蛋白。

[0161] 上述 HLA 单体、二聚体、四聚体和五聚体可有效用作对来源于本发明蛋白的肿瘤抗原肽特异性的 CTL 的检测试剂。

本发明的 CTL 检测试剂可以用于例如以下用途。

1)在启动采用本发明的蛋白、肽或核酸的治疗之前检验本发明的肿瘤抗原肽的 CTL 前体的频率和量。这样,可以评价患者对肿瘤抗原肽的反应性。

2)检验患者在本发明的蛋白、肽或核酸治疗下 CTL 的频率或量。这样,可进行疗效监测、评价治疗的合适性、确证治疗进展良好等。

[0162] 具体地说,可如下进行 CTL 的检测:由受试患者分离含有 CTL 的生物样品(例如 PBMC),使本发明的 HLA 四聚体等与该生物样品接触,并通过例如流式细胞术检测对结合 HLA 四聚体的本发明肽特异性的现有 CTL 的频率和量。

[0163] 9)本发明的疾病标志物

本发明还提供利用 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因或其表达产物的癌症(肿瘤)的疾病标志物。

[0164] 如上所述,本发明人发现,相比于正常组织,Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因在癌组织中的表达水平和/或频率显著增加。因此,识别这些基因或其表达产物(即蛋白)的抗体可用作表达这些基因的癌症的疾病标志物。

后文阐述了识别 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因的多核苷酸或其表达产物(即蛋白)的抗体作为疾病标志物的用途。

[0165] (1)癌症疾病标志物及其应用

(1-1)多核苷酸

如上所述,人 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因是本领域已知的。

已经提到,相比于正常组织,Lengsin、BJ-TSA-9、C20orf42、

BUB1、C10orf3 和 HIFPH3 基因在癌症患者组织中的表达水平和/或频率明确增加，本发明基于这一发现，以下想法也基于此：检测这些基因是否表达或这些基因的表达水平可以对癌症的有或无、严重性或康复提供明确测定，并因此提供对癌症的精确诊断。

因此，本发明提供可用作工具(即疾病标志物)的多核苷酸，其可通过在受试者中检测这些基因是否表达或这些基因的表达水平，诊断受试者中癌症的有或无或严重性。

[0166] 本文使用的“多核苷酸”包括 RNA 和 DNA，并可以为单链或双链的。“DNA”包括 cDNA、基因组 DNA 和合成 DNA。

“RNA”包括总 RNA、mRNA 和合成 RNA。本发明疾病标志物的特征在于其由多核苷酸和/或其互补多核苷酸组成，其中多核苷酸包含 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的碱基序列的至少 15 个连续核苷酸。

[0167] 具体地说，本发明的疾病标志物可以由多核苷酸和/或其互补多核苷酸组成，其中多核苷酸含有 Lengsin 基因(SEQ ID NO: 1)、BJ-TSA-9 基因(SEQ ID NO: 3)、C20orf42 基因(SEQ ID NO: 5)、BUB1 基因(SEQ ID NO: 7)、C10orf3 基因(SEQ ID NO: 9)或 HIFPH3 基因(SEQ ID NO: 11)的碱基序列的至少 15 个连续核苷酸。

[0168] 在本文中，术语“互补多核苷酸(互补链、反向链)”是指基于碱基配对原则(例如 A:T、G:C)与以下互补的多核苷酸：由以上 SEQ ID NO 的任一个碱基序列组成的多核苷酸的完整序列，或其含有任一个这些碱基序列的至少 15 个连续核苷酸的部分序列，为方便起见，其在本文还被称为“正向链”。“互补链”可以为与目标正向链的碱基序列完全互补的序列，或者为具有的互补性足以在严格条件下与目标正向链的碱基序列杂交的序列。本文使用的严格条件可以基于如在文献中所述的形成复合物或结合探针的核酸的解链温度(T_m)来确定(Berger 和 Kimmel, 1987, “Guide to Molecular Cloning Techniques Methods in Enzymology”, 152 卷, Academic Press, San Diego CA)。例

如，杂交后的洗涤通常可以在约“1 x SSC、0.1% SDS、37°C”的条件下进行。互补链在这些洗涤条件下洗涤时优选保持结合目标正向链。更严格的杂交条件可以涉及在约“0.5 x SSC、0.1% SDS、42°C”的条件下洗涤，再更严格的杂交条件包括约“0.1 x SSC、0.1% SDS、65°C”的洗涤条件，但其不限于此。具体地说，这样的互补链可以由与目标正向链的碱基序列相比具有完全互补性或至少 90%、优选 95%的同源性的碱基序列组成。

[0169] 正链的多核苷酸可以具有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH 基因的碱基序列的完整或部分序列，或者可以与上述互补链的碱基序列互补。

如上所述的正向链或互补(反向)链的多核苷酸可以以单链或双链形式用作疾病标志物。

[0170] 具体地说，本发明的疾病标志物可以为由 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的碱基序列(完整序列)或其互补序列组成的多核苷酸。或者，疾病标志物可以为由如上所述的完整序列的部分序列或其互补序列组成的多核苷酸，只要该多核苷酸选择性(特异性)识别 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因或源自其的多核苷酸。部分序列可以为含有适宜地选自上述完整序列或其互补序列的碱基序列的至少 15 个连续核苷酸的多核苷酸。

[0171] 本文使用的术语“选择性(特异性)识别”意指就 RNA 印迹而言，可以特异性检测 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因或由其获得的多核苷酸，就 RT-PCR 方法而言，生产 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因或由其获得的的多核苷酸。然而，以上定义不是限制性的，可以使用任何标准，只要本领域一般技术人员可以通过使用源自这些基因中的一个的多核苷酸确定要检测的物质或产生的产物。

[0172] 例如，本发明的疾病标志物可以利用引物 3

(<http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi>) 或载体 NTI (Infomax)由 Lengsin 基因(SEQ ID NO: 1)、BJ-TSA-9 基因(SEQ ID NO: 3)、C20orf42 基因(SEQ ID NO: 5)、BUB1 基因(SEQ ID NO: 7)、C10orf3 基因(SEQ ID NO: 9)或 HIFPH3 基因(SEQ ID NO: 11)的碱基序列设计。具体地说, 通过使用引物 3 或载体 NTI 软件由 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的碱基序列获得的候选序列, 或者至少含有候选序列一部分的序列, 可以用作引物或探针。

本发明的疾病标志物含有如上所述的至少 15 个连续核苷酸长度, 根据标志物的应用, 可以适宜地选择和设计长度。

[0173] (1-2)作为探针或引物的多核苷酸

通过在受试者的生物样品中检验 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因中的至少一种是否表达或其表达水平进行本发明的癌症检测(诊断)。在此, 本发明的疾病标志物可以用作特异性识别和扩增由任一个这些基因表达的 RNA 或由其获得的的多核苷酸的引物, 或者用作特异性检测此 RNA 或由其获得的的多肽的探针。

[0174] 当本发明的疾病标志物用作检测癌症(即用于遗传诊断)的引物时, 该引物一般可以含有 15-100 bp、优选 15-50 bp、更优选 15-35 bp 的长度。在用作检测癌症的探针时, 该探针一般可以含有 15 bp-1 kp、优选 100 bp-1 kb 的长度。

本发明的疾病标志物可以在特异性检测特定基因的任何已知方法中以常规方式用作引物或探针, 所述方法例如为 RNA 印迹、RT-PCR、原位杂交等。然后, 可以检验涉及癌症的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因是否表达或其表达水平。

[0175] 待检验的样品可以为用常规方法由受试者的靶组织的生物样品(例如活检样品)制备的总 RNA, 或由该总 RNA 制备的多核苷酸。

在生物样品中的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3

或 HIFPH3 基因的表达水平可以通过使用 DNA 芯片检测或定量。在此,本发明的疾病标志物可以用作 DNA 芯片的探针(就基因芯片人类基因组 U95 A、B、C、D、E (Affymetrix)而言,用作多核苷酸探针,其为 25 bp 长度)。当含有本发明的疾病标志物作为其探针的 DNA 芯片与由生物样品的 RNA 制备的标记-DNA 或 RNA 杂交时,本发明的疾病标志物(即探针)和标记的 DNA 或 RNA 形成复合物。通过利用标记的 DNA 或 RNA 的标记检测复合物,可以检验 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因是否表达或其表达水平。

[0176] DNA 芯片可以含有一个或多个本发明的疾病标志物,其可以结合 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因中的任一个。当 DNA 芯片含有 1 个以上的标志物时,可以在 1 个生物样品中同时检验 1 个以上的基因是否表达或其表达水平。

[0177] 本发明的疾病标志物可用于诊断或检测癌症,也就是说,诊断癌症是否存在或癌症的严重性。具体地说,可以通过评价受试者和健康个体之间的组织生物样品中的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的表达水平差异,使用疾病标志物进行癌症诊断。

在此,两种受试者的表达水平之间的差异可以为有或无表达,或者在两种受试者中均观察到表达时,该差异可以为至少 1.5 倍,优选 2 倍,更优选 3 倍。

[0178] 具体地说,来源于 Lengsin 基因的多核苷酸可用作肺腺癌、肺鳞状细胞癌或胃癌的疾病标志物。

具体地说,来源于 BJ-TSA-9 基因的多核苷酸可用作白血病、肺腺癌、肺鳞状细胞癌、小细胞肺癌、口腔癌、胃癌、胰腺癌或淋巴瘤的疾病标志物。

具体地说,来源于 C20orf42 基因的多核苷酸可用作肺鳞状细胞癌、肺腺癌、肝癌、胃癌、白血病、恶性淋巴瘤组织、直肠癌、结肠癌或胰腺癌的疾病标志物。

[0179] 具体地说,来源于BUB1基因的多核苷酸可用作乳腺癌、肺腺癌、肺鳞状细胞癌、卵巢癌、口腔鳞状细胞癌、肾癌、大肠癌(结肠癌、直肠癌)、胃癌、胰腺癌、肝癌、白血病、淋巴瘤或黑素瘤的疾病标志物

具体地说,来源于C10orf3基因的多核苷酸可用作乳腺癌、结肠癌、直肠癌、肾癌、胃癌、卵巢癌、肝癌、胰腺癌、肺鳞状细胞癌、肺腺癌、小细胞肺癌或黑素瘤的疾病标志物。

具体地说,来源于HIFPH3基因的多核苷酸可用作乳腺癌、结肠癌、胃癌、肾癌、胰腺癌、肝癌、肺腺癌或肺鳞状细胞癌的疾病标志物。

[0180] (1-3)抗体

本发明提供特异性识别 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的表达产物(即蛋白 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3)的抗体,其可以用作癌症的疾病标志物。

具体地说,本发明提供特异性识别分别具有 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 蛋白的抗体。

[0181] 已经提到,相比于正常细胞或组织, Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 基因在癌症患者组织中的表达水平和/或频率明确增加,本发明基于这一发现,以下想法也基于此:检测这些基因是否表达或其表达水平可以对癌症的有或无、严重性或康复提供明确测定,并因此提供对癌症的精确诊断。

因此,所述抗体可用作工具(即疾病标志物),其可通过在受试者中检测该蛋白是否表达或该蛋白的表达水平,诊断受试者是否存在癌症或其严重性。

[0182] 如上所述,人 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 蛋白在本领域是已知的。

就形式而言，本发明的抗体不受限制，可为针对 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 产生的多克隆抗体或单克隆抗体。此外，本发明的抗体包括对由任一个这些蛋白的氨基酸序列的一般至少 8 个、优选 15 个、更优选 20 个连续氨基酸组成的多肽具有亲和力的抗体。

[0183] 制备此抗体的方法在本领域众所周知，本发明的抗体可以按照任一个这些常规方法制备(Current protocols in Molecular Biology, Ausubel 等人编辑, (1987) Publish. John Wiley and Sons. 11.12-11.13 章)。具体地说，当抗体为多克隆抗体时，其可如下获得：用以常规方式在大肠杆菌中表达并由大肠杆菌纯化的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 蛋白作为免疫原，或者用具有以常规方式合成的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分氨基酸序列的寡肽作为免疫原，免疫非人动物如兔，并以常规方式由免疫动物的血清回收抗体。当抗体为单克隆抗体时，其可如下获得：用以常规方式在大肠杆菌中表达并由大肠杆菌纯化的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 蛋白或者用具有以常规方式合成的任一种这些蛋白的部分氨基酸序列的寡肽免疫非人动物如小鼠，并使所获脾细胞与骨髓瘤细胞进行细胞融合，以制备杂交瘤细胞，并由杂交瘤细胞回收抗体(Current protocols in Molecular Biology, Ausubel 等人编辑, (1987) Publish. John Wiley and Sons. 11.4-11.11 章)。

[0184] 可以通过诸如 DNA 克隆、构建对应质粒、将质粒转染入宿主细胞中、培养所获转化体并由转化体培养物回收蛋白的程序，由本文提供的 SEQ ID NO 1、3、5、7、9 和 11 中任一个的序列获得在抗体制备中用作免疫原的蛋白 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3。这些程序可以通过本领域技术人员已知的或在文献(Molecular Cloning, T. Maniatis 等, CSH Laboratory (1983), DNA Cloning, DM. Glover, IRL PRESS (1985))中描述的任一种方法进行。

[0185] 具体地说, 用作制备本发明抗体的免疫原的蛋白可以如下获得: 构建在选定的宿主细胞中提供 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因表达的重组 DNA (即表达载体), 将该 DNA 转染入宿主细胞中, 以提供转化体, 培养转化体, 由转化体培养物回收目标蛋白。另外, Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分肽可以基于本文提供的氨基酸序列 (SEQ ID NO: 2、4、6、8、10 或 12) 通过常用的化学(肽)合成方法制备。

[0186] 在此, 本发明的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 不仅包括 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列的蛋白, 而且包括其同源物。同源物包括由以下氨基酸序列组成的蛋白: 该氨基酸序列与 SEQ ID NO: 2、4、6、8、10 或 12 的氨基酸序列相同, 只是缺失、取代和/或添加一个或多个氨基酸, 并具有与该蛋白的已知免疫学活性等效的免疫学活性。

具有与该蛋白的已知免疫学活性等效的免疫学活性的同源物为例如能够在适宜的动物或其细胞中诱导特异性免疫反应并特异性结合抗 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的抗体的蛋白。

[0187] 关于蛋白中的氨基酸突变的数量或位置没有限制, 只要保持该蛋白的免疫学活性。人们可基于其确定要缺失、取代和/或添加的氨基酸残基的数量或位置而不降低免疫学活性的标准可使用本领域众所周知的计算机程序获得, 例如 DNA Star 软件。例如, 突变数通常应在总氨基酸残基的 10% 之内, 优选应在总氨基酸残基的 5% 之内, 更优选在总氨基酸残基的 1% 之内。要取代的氨基酸不受限制, 只要所获蛋白保持与 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 等效的免疫学活性。考虑到保持结构, 通过取代引入的氨基酸优选与被取代的氨基酸具有类似特征, 其中所述特征包括极性、电荷、溶解性、疏水性、亲水性、两亲性等。例如, Ala、Val、Leu、Ile、Pro、Met、Phe 和 Trp 被分类为非极性氨基酸; Gly、Ser、Thr、

Cys、Tyr、Asn 和 Gln 被分类为不带电的氨基酸；Asp 和 Glu 被分类为酸性氨基酸；而 Lys、Arg 和 His 被分类为碱性氨基酸。本领域一般技术人员可以基于这些标准选择属于相同类别的适宜氨基酸。

[0188] 可通过使用具有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的部分氨基酸序列的寡肽获得本发明的抗体。用于制备抗体的寡肽不一定具有功能性生物活性，但其需要具有与对应蛋白类似的免疫原性。优选地，寡肽具有这样的免疫原性，并由 BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的氨基酸序列的至少 8 个、优选 15 个、更优选 20 个连续氨基酸组成。

[0189] 为制备针对该寡肽的抗体，可将多种佐剂与免疫原一起同时给予宿主，以便增强免疫反应性。佐剂的实例为(但不限于)弗氏佐剂；诸如氢氧化铝的矿物凝胶；表面活性剂，如溶血卵磷脂、Pluronic®多元醇、聚阴离子、肽、油乳化剂、匙孔蛾血蓝蛋白和二硝基苯酚；人用佐剂，如 BCG (卡介苗)或棒状杆菌。

另外，本发明可以使用商品化抗体，包括抗 Bub1 抗体(Upstate)、抗-C10orf3 抗体(Abnova)、抗-C20orf42 抗体(Imgenex)和抗-HIFPH3 抗体(Bethyl)。

[0190] 本发明的抗体具有特异性结合 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的特征，因此可特异性检测在受试者的组织中表达的蛋白(及其同源物)。因此，本发明的抗体可以用作检测 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 在受试者的组织中是否表达或其表达水平的探针。

[0191] 具体地说，Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 可如下检测：通过活检等获得患者靶组织的生物样品，以常规方式制备组织提取物或由其获得的蛋白，并通过已知方法如蛋白质印迹、ELISA 等使用所述抗体作为探针以常规方式进行检测。

通过检验 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 和 HIFPH3 中的至少一种在受试者的组织中的水平，并与健康个体中的该水平相

比较, 可提供癌症诊断。蛋白水平的差异可为有或无蛋白, 或者可以为至少 2 倍, 优选 3 倍。

[0192] 特异性识别 Lengsin 的抗体尤其可用作肺腺癌、肺鳞状细胞癌或胃癌的疾病标志物。

特异性识别 BJ-TSA-9 的抗体尤其可以用作白血病、肺腺癌、肺鳞状细胞癌、小细胞肺癌、口腔癌、胃癌、胰腺癌或淋巴瘤的疾病标志物。

特异性识别 C20orf42 的抗体尤其可以用作肺鳞状细胞癌、肺腺癌、肝癌、胃癌、白血病、恶性淋巴瘤组织、直肠癌、结肠癌或胰腺癌的疾病标志物。

[0193] 特异性识别 BUB1 的抗体尤其可以用作乳腺癌、肺腺癌、肺鳞状细胞癌、卵巢癌、口腔鳞状细胞癌、肾癌、大肠癌(结肠癌、直肠癌)、胃癌、胰腺癌、肝癌、白血病、淋巴瘤或黑素瘤的疾病标志物。

特异性识别 C10orf3 的抗体尤其可用作乳腺癌、结肠癌、直肠癌、肾癌、胃癌、卵巢癌、肝癌、胰腺癌、肺鳞状细胞癌、肺腺癌、小细胞肺癌或黑素瘤的疾病标志物。

特异性识别 HIFPH3 的抗原尤其可以用作乳腺癌、结肠癌、胃癌、肾癌、胰腺癌、肝癌、肺腺癌或肺鳞状细胞癌的疾病标志物。

[0194] (2)检测(或诊断)癌症的方法

本发明提供一种利用如上所述的本发明疾病标志物检测(或诊断)癌症的方法。

具体地说, 本发明的检测方法包括通过活检等由受试者收集靶组织的生物样品, 检测和测量 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的表达水平(量)或由该基因编码的蛋白的表达水平(量), 并检测(或诊断)癌症的有或无或严重性。本发明的检测(诊断)方法可用于检测(或诊断)为改善癌症而给予癌症患者的治疗剂实际上是否可以改善疾病, 和/或所述治疗剂可以提供多大的改善。

[0195] 本发明的癌症检测方法包括以下的步骤(a)、(b)和(c):

(a)使由受试者制备的生物样品与本发明的疾病标志物接触;

(b)通过使用疾病标志物作为指示剂,检测样品中 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的表达水平,和

(c)基于(b)的结果确定受试者是否患有癌症。

[0196] 本文使用的生物样品可以为由受试者的靶组织制备的样品。具体地说,生物样品可以为含有由(靶)组织制备的 RNA 的样品、含有由该 RNA 制备的多核苷酸的样品或含有由靶组织制备的蛋白的样品。该含有 RNA、多核苷酸或蛋白的样品可在通过活检等收集受试者的一部分组织后通过常规方法制备。

本发明的诊断方法可根据要检验的生物样品如下所述进行。

[0197] (2-1)当要检验的生物样品为 RNA 时

当要检验的生物样品为 RNA 时,癌症的检测可通过含有以下步骤(a)、(b)和(c)的方法进行:

(a)使本发明的疾病标志物(即具有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的碱基序列的至少 15 个连续核苷酸的多核苷酸和/或其互补多核苷酸)与由受试者的生物样品制备的 RNA 或由其转录的互补多核苷酸杂交;

b)使用疾病标志物作为指示剂检测与疾病标志物杂交的、由生物样品制备的 RNA 或由其转录的互补多核苷酸; 和

(c)基于(b)中的检测结果确定受试者是否患有癌症。

[0198] 当要检验 RNA 时,本发明的检测(诊断)方法可通过检测和测量 RNA 中 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的表达水平实现。具体地说,该方法可通过使用本发明的疾病标志物(即含有 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的碱基序列的至少 15 个连续核苷酸的多核苷酸和/或其互补多核苷酸)作为引物或探针,通过已知方法如 RNA 印迹、RT-PCR、DNA 芯片分析、原位杂交等进行。

[0199] 就 RNA 印迹而言, 本发明的疾病标志物用作探针, 并可以检测和测量 RNA 中 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因是否表达或其表达水平。例如, 本发明的疾病标志物(其为互补链)用放射性同位素(^{32}P 、 ^{33}P 等: RI)或荧光物质标记。然后, 标记的疾病标志物以常规方式与被转移到尼龙膜等上的、由受试者的组织制备的 RNA 杂交。此后, 可通过用放射性检测器(BAS-1800II, Fuji Photo Film., Inc.)或荧光检测器检测和测量标记(RI 或荧光物质)的信号检测在标记的疾病标志物(DNA)和 RNA 之间形成的双链体。或者, 可以用 AlkPhos 直接标记和检测系统(Amersham Pharmacia Biotech)按照生产商的方案标记疾病标志物(即探针 DNA), 并使其与由受试者的组织制备的 RNA 杂交, 然后可以使用 Multi-biomeasure STORM860 (Amersham Pharmacia Biotech)检测和测量疾病标志物上的标记的信号。

[0200] 就 RT-PCR 而言, 本发明的疾病标志物用作引物, 可以检测和测量 RNA 中 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因是否表达或其表达水平。例如, 以常规方式由来源于受试者的组织的 RNA 制备 cDNA, 该 cDNA 与由本发明的疾病标志物制备的引物对(正向引物与为反向链的上述 cDNA 杂交, 反向引物结合正向链)杂交, 以便使用该 cDNA 作为模板扩增对应于 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的靶区域。然后, 以常规方式进行 PCR, 检测所产生的已扩增 DNA 双链体。已扩增 DNA 双链体的检测例如可通过以下方法实现: 其中采用先前用 RI 或荧光物质标记的引物进行 PCR, 检测产生的标记 DNA 双链体; 其中产生的 DNA 双链体被转移到尼龙膜等之上, 然后由于与用作探针的标记的疾病标志物杂交而被检测。所产生的标记 DNA 双链体可以使用 Agilent 2100 Bioanalyser (Yokogawa Analytical Systems Inc.)等检测。检测还可以如下实现: 使用 SYBR Green RT-PCR 试剂(Applied Biosystems)按照生产商的方案制备 RT-PCR 反应溶液, 使用 ABI

PRISM 7700 序列检测系统(Applied Biosystems)进行反应, 并检测反应产物。

[0201] 就 DNA 芯片分析而言, 例如, 在其上连接本发明的疾病标志物的 DNA 芯片被制备为 DNA 探针(其为单链的或双链的 DNA 探针), DNA 芯片以常规方式与由来源于受试者组织的 RNA 制备的 cRNA 杂交。然后, 由 DNA 和 cRNA 形成的双链体通过使由本发明的疾病标志物制备的标记探针与该双链体结合而被检测。能够检测和测量 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的表达水平的 DNA 芯片可用于本发明。这样的 DNA 芯片的实例为基因芯片人类基因组 U95 A、B、C、D、E (Affymetrix)。在以下实施例中描述了使用这样的 DNA 芯片检测和测量 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因在受试者的 RNA 中的表达水平。

[0202] (2-2)当要检验的生物样品为蛋白时

当要检验的生物样品为蛋白时, 本发明的检测(诊断)方法通过检测和测量生物样品中的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的水平(量)进行。具体地说, 该方法可包含以下步骤(a)、(b)和(c):

(a)使由受试者的生物样品制备的蛋白与为抗体的本发明疾病标志物(即识别 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的抗体)结合;

b)通过使用疾病标志物作为指示剂, 检测与疾病标志物结合的、由生物样品制备的蛋白或由其获得的的部分肽; 和

(c)基于(b)中的检测结果确定受试者是否患有癌症。

[0203] 具体地说, 该方法可为已知方法, 例如蛋白质印迹, 用于通过使用为抗体的本发明疾病标志物(即特异性识别 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 的抗体)检测和测量 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3。

蛋白质印迹可以通过使用抗体进行,即本发明的疾病标志物为一抗,用放射性同位素如 ^{125}I 、荧光物质等标记的抗体作为结合一抗的二抗,并使用放射性检测器(BAS-1800II, Fuji Photo Film., Inc.等)、荧光检测器等检测和测量所产生的标记复合物中的放射性同位素或荧光物质的信号。或者,在使用本发明的疾病标志物作为一抗后,可以用 ECL Plus 蛋白质印迹检测系统(Amersham Pharmacia Biotech)按照生产商的方案和用 Multi-biomeasure STORM860 (Amersham Pharmacia Biotech)进行检测和测量。

[0204] (2-3)癌症的诊断

癌症的诊断可如下进行:比较受试者组织中 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因的表达水平或 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 蛋白的表达水平(或量)和正常组织中的该水平,并评价二者之间的差异。

[0205] 由所需的正常组织收集或制备的生物样品(RNA 或蛋白)可以通过活检或在其同意的情况下通过收集死后组织得自无癌症的受试者的组织或癌症患者的非癌性的正常组织。“无癌症的受试者”是指至少无症状的受试者,优选为任何其它诊断方法如 X-射线检查结果未被诊断为癌症的受试者。“无癌症的受试者”在本文可被简称为“健康个体”。

[0206] 可以通过在受试者和健康个体二者的平行生物样品中测量,对受试者组织中的基因或蛋白的表达水平与正常组织(即没有癌症的受试者的组织或癌症患者的非癌性正常组织)中的该表达水平进行比较。健康个体中的基因或蛋白的表达水平不一定是平行检测的水平,在恒定条件下可以使用通过检测多个(至少 2 个,优选等于或多于 3 个,更优选等于或多于 5 个)正常组织确定的 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 基因或蛋白表达水平的平均值或统计学中值。

[0207] 当 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或

HIFPH3 基因或其表达产物(即 Lengsin、BJ-TSA-9、C20orf42、BUB1、C10orf3 或 HIFPH3 蛋白)在受试者组织中的表达水平为健康个体水平的 2 倍、优选 3 倍高时,受试者被确定为患有癌症。当受试者中的所述基因或蛋白的表达水平高于健康个体中的该水平时,受试者被确定为患有癌症或疑似患有癌症。

[0208] 具体地说,在以上的基因和蛋白中,当 Lengsin 基因或蛋白的表达水平在受试者的肺和胃中比在对应的正常组织中高时,受试者疑似患有癌症。

此外,当 BJ-TSA-9 基因或蛋白的表达水平在受试者的白细胞、肺、口腔、胃、胰腺或淋巴结中比在对应的正常组织中高时,受试者疑似患有癌症。

此外,当 C20orf42 基因或蛋白的表达水平在受试者的肺、肝脏、胃、白细胞、淋巴结、直肠、结肠或胰腺中比在对应的正常组织中高时,受试者疑似患有癌症。

[0209] 此外,当 BUB1 基因或蛋白的表达水平在受试者的乳腺、肺、卵巢、口腔、肾脏、大肠(结肠、直肠)、胃、胰腺、肝脏、白细胞、淋巴结或皮肤中比在对应的正常组织中高时,受试者疑似患有癌症。

此外,当 C10orf3 基因或蛋白的表达水平在受试者的乳腺、结肠、直肠、肾脏、胃、卵巢、肝脏、胰腺、肺或皮肤中比在对应的正常组织中高时,受试者疑似患有癌症。

此外,当 HIFPH3 基因或蛋白的表达水平在受试者的乳腺、结肠、胃、肾脏、胰腺、肝脏、肺中比在对应的正常组织中高时,受试者疑似患有癌症。

[0210] 本发明由以下实施例具体解释,但不能在任何意义上认为所述实施例限制本发明。

实施例 1

[0211] 由人组织样品和人癌细胞系制备总 RNA

通过常规方法由如下的人主要器官的癌症组织和正常组织的样品制备总 RNA: 肺腺癌(57 个样品)、正常肺(得自肺腺癌患者)(46 个样品)、肺鳞状细胞癌(48 个样品)、正常肺(得自肺鳞状细胞癌患者)(18 个样品)、结肠癌(108 个样品)、正常结肠(得自结肠癌患者)(114 个样品)、直肠癌(43 个样品)、正常直肠(得自直肠癌患者)(31 个样品)、胃癌(38 个样品)、正常胃(得自胃癌患者)(13 个样品)、乳腺癌(237 个样品)、正常乳腺(得自乳腺癌患者)(39 个样品)、卵巢癌(37 个样品)、正常卵巢(得自卵巢癌患者)(5 个样品)、肝癌(19 个样品)、正常肝脏(得自肝癌患者)(8 个样品)、肾癌(89 个样品)、正常肾脏(得自肾癌患者)(64 个样品)、胰腺癌(55 个样品)、正常胰腺(得自胰腺癌患者)(16 个样品)、白血病(6 个样品)、淋巴瘤(90 个样品)、正常骨髓(2 个样品)、黑素瘤(10 个样品)和正常皮肤(72 个样品)。

实施例 2

[0212] DNA 芯片分析

使用由实施例 1 中所述的样品制备的总 RNA 进行 DNA 芯片分析。对于 DNA 芯片分析, 使用基因芯片人类基因组 U133 套件(Affymetrix)。具体地说, 分析包含以下步骤: (1)由总 RNA 制备 cDNA, (2)由 cDNA 制备标记的 cRNA, (3)片段化标记的 cRNA, (4)使片段化的 cRNA 与探针阵列杂交, (5)染色探针阵列, (6)筛选探针阵列, 和(7)分析基因表达。

[0213] (1)由总 RNA 制备 cDNA

将含有在实施例 1 中制备的每种总 RNA (10 μg)和 T7-(dT)24 引物(Amersham)(100 pmol)(11 μl)的混合物加热至 70°C 达 10 分钟, 并在冰上冷却。在冷却后, 加入 5x 首链 cDNA 缓冲液(4 μL)、0.1 M DTT (二硫苏糖醇)(2 μl)和 10 mM dNTP Mix (1 μl)(用于 cDNA 合成的 SuperScript 选择系统(Gibco-BRL)), 将混合物加热至 42°C 达 2 分钟。

然后, 加入 Super ScriptII RT (2 μ l, 400U)(包含在上述试剂盒中), 将混合物加热至 42°C 达 1 小时, 然后在冰上冷却。在冷却后, 加入用 DEPC (nacalai tesque)(91 μ l)处理的无菌蒸馏水和 5x 第二链反应缓冲液(30 μ L)、10 mM dNTP Mix (3 μ l)、大肠杆菌 DNA 连接酶(1 μ l, 10U)、大肠杆菌 DNA 聚合酶 I (4 μ l, 40U)和大肠杆菌 RNaseH (1 μ l, 2U)(包括在如上所述的试剂盒中), 将混合物于 16°C 温育 2 小时。在加入 T4 DNA 聚合酶(2 μ l, 10U)(包括在如上所述的试剂盒中)之后, 将混合物再于 16°C 温育 5 分钟, 然后添加 0.5 M EDTA (10 μ l)。加入苯酚/氯仿/异戊醇溶液(NIPPON GENE CO., LTD.)(162 μ l)并混合。将混合物移至预先于室温以 14,000 rpm 离心 30 秒的 Phase Lock Gel Light (Eppendorf), 并于室温以 14,000 rpm 离心 2 分钟, 将其水相(145 μ l)移至 Eppendorf 管。向获得的溶液添加 7.5 M 乙酸铵溶液(72.5 μ l)和乙醇(362.5 μ l), 并混合, 然后于 4°C 以 14,000 rpm 离心 20 分钟。在离心后, 通过废弃上清液获得含有所制备的 cDNA 的沉淀。然后, 向沉淀添加 80%乙醇(0.5mL), 并于 4°C 以 14,000 rpm 离心 5 分钟, 废弃上清液。在重复相同的步骤后, 干燥沉淀, 用 DEPC-处理的水(12 μ l)裂解。结果, 由在实施例 1 中制备的每种总 RNA 获得 cDNA。

[0214] (2)由 cDNA 制备标记的 cRNA

将每种 cDNA 溶液(5 μ l)与 DEPC-处理的水(17 μ l)和 10x HY 反应缓冲液(4 μ l)、10x 生物素标记的核糖核苷酸(4 μ l)、10x DTT (4 μ l)、10x RNA 酶抑制剂混合物(4 μ l)和 20x T7 RNA 聚合酶(2 μ l)(BioArray 高产 RNA 转录物标记试剂盒(ENZO))混合, 并于 37°C 温育 5 小时, 以制备标记 cRNA。在温育后, 向溶液加入 DEPC-处理的水(60 μ l), 利用 RNeasy 小型试剂盒(GIAGEN)按照生产商的方案纯化所制备的标记 cRNA。

[0215] (3)片段化标记的 cRNA

向 5x 片段化缓冲液(200 mM Tris-乙酸 pH 8.1 (Sigma)、500 mM 乙酸钾(Sigma)、150 mM 乙酸镁(Sigma))(8 μ l)加入含有每种标记的

cRNA (20 μg)的溶液, 加热溶液(40 μl)至 94°C 达 35 分钟, 并置于冰上。结果, 实现了标记 cRNA 的片段化。

[0216] (4)片段化的 cRNA 与探针阵列的杂交

将每种片段化的 cRNA (40 μl)加入 5 nM 对照寡聚物 B2 (Amersham)(4 μl)、100x 对照 cRNA 混合物(4 μl)、鲑鱼精 DNA (Promega) (40 μg)、乙酰化 BSA (Gibco-BRL)(200 μg)、2x MES 杂交缓冲液(200 mM MES, 2 M $[\text{Na}^+]$, 40 mM EDTA, 0.02%吐温 20 (Pierce), pH 6.5-6.7)(200 μl)和 DEPC 处理的水(144 μl), 产生杂交混合物(400 μl)。将获得的每种杂交混合物加热至 99°C 达 5 分钟, 然后于 45°C 达 5 分钟。在加热后, 于室温以 14,000 rpm 离心杂交混合物 5 分钟, 以获得上清液。

[0217] 同时, 如下制备检验阵列: 用 1x MES 杂交缓冲液装载人类基因组 U133 探针阵列(Affymetrix), 在杂交烘箱中于 45°C 以 60 rpm 旋转该阵列 10 分钟, 然后除去 1xMES 杂交缓冲液。将如上获得的每种杂交混合物的上清液(200 μl)加入探针阵列, 在杂交烘箱中于 45°C 以 60 rpm 旋转该探针阵列 16 小时, 以使片段化的 cRNA 与探针阵列杂交。

[0218] (5)探针阵列的染色

在由每个与片段化的 cRNA 杂交的探针阵列除去杂交混合物之后, 用非严格洗涤缓冲液(6x SSPE (由 20x SSPE (nacalai tesque)稀释)、0.01%吐温 20 和 0.005%防沫剂 0-30 (Sigma)装载探针阵列。然后, 将与片段化的 cRNA 杂交的探针阵列置于装载非严格洗涤缓冲液和严格洗涤缓冲液(100 mM MES、0.1 M NaCl、0.01%吐温 20)的 GeneChip Fluidics Station 400 (Affymetrix)中的给定位置。此后, 按照染色方案 EuKGE-WS2, 用第一染色溶液(10 $\mu\text{g}/\text{mL}$ 链霉抗生物素蛋白藻红蛋白 (SAPE)(MolecuLar Probe)、2 mg/mL 乙酰化 BSA、100 mM MES、1 M NaCl (Ambion)、0.05%吐温 20、0.005%防沫剂 0-30)和第二染色溶液(100 $\mu\text{g}/\text{mL}$ 山羊 IgG (Sigma)、3 $\mu\text{g}/\text{mL}$ 生物素化抗-链霉抗生物素蛋白

抗体(Vector Laboratories)、2 mg/mL 乙酰化 BSA、100 mM MES、1 M NaCl、0.05%吐温 20、0.005%防沫剂 0-30)染色探针阵列。

[0219] (6)探针阵列的扫描和(7)基因表达水平分析

将染色的探针阵列运用于 HP 基因阵列扫描仪(Affymetrix), 并读取染色模式。基于该染色模式, 通过基因芯片工作站系统(Affymetrix)分析探针阵列上的基因表达。然后, 在按照分析方案标准化后, 计算每个样品中每种探针(即每种基因)的表达水平(平均差异)以及是否存在表达。对于每种探针, 计算不同种类样品的每一种的基因表达水平的平均值, 并获得不同种类样品之间的表达水平和频率的差异。

实施例 3

[0220] 表达变异分析

如下所述, 选择显示出在源自多种主要器官(肺、结肠、直肠、胃、乳腺、肝脏、肾脏、卵巢、胰腺)中癌组织的表达水平和/或频率比正常组织增加的基因。

[0221] 按照基因表达的 DNA 芯片分析结果, 从一组其在癌组织中的表达水平和/或频率相比于对应的正常组织增加的探针中, 选择在许多癌组织中的表达水平和/或频率显示明确增加的探针。然后, 检查并选择对应于这些选定探针的基因。

结果, 选定 6 个基因: BUB1、C10orf3、C20orf42、Lengsin、HIFPH3 和 BJ-TSA-9 基因。这 6 种基因的表达变异比率描述于表 2, 表达频率描述于表 3。

[0222][表 2]

基因	乳腺癌	结肠癌	肾癌	肝癌	肺腺癌	鳞状细胞癌	胰腺癌	胃癌	卵巢癌	直肠癌	白血病	恶性淋巴瘤	黑色素瘤
BUB1	2.1	2.0	2.6	3.1	2.0	4.0	2.0	2.4	4.2	2.3	1.4	1.3	3.1
C10orf3	2.5	2.3	2.9	6.7	3.0	6.2	2.4	2.4	9.0	2.9	0.9	0.9	3.9
C20orf42	0.5	2.8	0.4	2.0	6.0	15.4	4.2	2.4	0.3	2.4	2.0	1.3	0.4
Lengsin	1.0	1.0	0.4	0.7	8.5	2.9	2.1	2.3	1.9	1.2	1.3	1.8	0.9
HIFPH3	2.3	1.3	15.5	2.6	2.3	15.9	4.4	1.2	0.9	0.9	0.6	0.6	0.2
BJ-TSA-9	1.0	0.8	1.0	1.5	9.4	8.1	2.6	3.6	4.3	0.9	14.9	2.3	0.5

[0223][表 3]

基因	乳腺癌	正常乳腺	结肠癌	正常结肠	肾癌	正常肾脏	肝癌	正常肝脏	肺腺癌	正常肺
BUB1	39	8	69	21	8	0	26	0	39	11
C10orf3	62	15	89	54	15	0	21	0	60	13
C20orf42	2	3	97	66	0	0	0	0	18	0
Lengsin	0	0	0	1	0	17	11	13	53	0
HIFPH3	14	5	35	25	78	2	11	0	25	7
BJ-TSA-9	2	3	0	1	0	0	0	0	39	0

基因	肺鳞状 细胞癌	正常肺	胰腺癌	正常胰 腺	胃癌	正常胃	卵巢癌	正常卵 巢	直肠癌	正常直 肠
BUB1	63	11	22	6	61	23	70	0	74	13
C10orf3	79	6	56	6	89	46	84	20	84	48
C20orf42	48	0	53	0	68	38	3	20	95	81
Lengsin	8	0	0	0	13	0	5	0	5	0
HIFPH3	52	0	25	0	16	8	27	20	26	32
BJ-TSA-9	23	0	4	0	3	0	5	0	0	0

基因	白血病	恶性淋 巴瘤	正常骨 髓	黑色素 瘤	正常皮 肤
BUB1	100	76	100	70	3
C10orf3	83	81	100	50	7
C20orf42	0	1	0	10	24
Lengsin	0	0	0	0	0
HIFPH3	17	0	0	0	63
BJ-TSA-9	17	0	0	0	1

[0224] 表2中的值指示与对应的正常组织相比在多种癌症组织中的表达变化比率,其中在每种癌症组织中的表达水平相对于用人基因组 U133 芯片分析的在每种正常组织中的表达水平(被视为 1)确定。表3中的值指示所述基因在多种癌症组织和正常组织中的表达频率(%)。

[0225] BUB1 基因显示出在检验的所有癌症组织中的表达水平相比于正常组织增加,该增加在肺鳞状细胞癌和卵巢癌中特别显著。表达频率也增加,该增加在肺鳞状细胞癌、卵巢癌和黑素瘤组织中特别显著。

C10orf3 基因显示出在乳腺癌、结肠癌、肾癌、肝癌、肺腺癌、肺鳞状细胞癌、胰腺癌、胃癌、卵巢癌、直肠癌和黑素瘤组织中的表达水平相比于正常组织增加,该增加在卵巢癌、肝癌和肺鳞状细胞癌中特别显著。如对表达水平所观察到的一样,表达频率在除血癌之外的所有癌症中也增加,该增加在肺鳞状细胞癌、胰腺癌和卵巢癌中特别显著。

[0226] C20orf42 基因显示出在结肠癌、肝癌、肺腺癌、肺鳞状细胞癌、胰腺癌、胃癌、直肠癌、白血病和恶性淋巴瘤组织中的表达水平相比于正常组织增加,该增加在肺鳞状细胞癌和肺腺癌中特别显著。表达频率也增加,该增加在肺腺癌、肺鳞状细胞癌和胰腺癌中特别显著。

Lengsin 基因显示出在肺腺癌、肺鳞状细胞癌和胃癌中的表达水平相比于正常组织增加,该增加在肺腺癌中特别显著。如对表达水平所观察到的一样,表达频率在肺腺癌、肺鳞状细胞癌和胃癌中也增加,该增加在肺腺癌中特别显著。

[0227] HIFPH3 基因显示出在乳腺癌、结肠癌、肾癌、肝癌、肺腺癌、肺鳞状细胞癌、胰腺癌和胃癌组织中的表达水平相比于正常组织增加,该增加在肾癌和肺鳞状细胞癌中特别显著。如对表达水平所观察到的一样,表达频率在乳腺癌、结肠癌、肾癌、肝癌、肺腺癌、

肺鳞状细胞癌、胰腺癌和胃癌中也增加,该增加在肾癌、肺鳞状细胞癌和胰腺癌中特别显著。

BJ-TSA-9 基因显示出在肺腺癌、肺鳞状细胞癌和白血病中的表达水平相比于正常组织增加,该增加在白血病、肺腺癌和肺鳞状细胞癌中特别显著。如对表达水平所观察到的一样,表达频率在肺腺癌、肺鳞状细胞癌和白血病中也增加,该增加在肺腺癌和肺鳞状细胞癌中特别显著。

[0228] 如上所示,选定的6种基因显示出在癌症组织中的表达水平和频率相比于对应的正常组织明确增加。因此,发现这6种基因及其表达产物(即蛋白)可用作癌症的疾病标志物(尤其是对于以上提及的癌症)。

实施例 4

[0229] 通过 RT-PCR 法对不同细胞和组织中的肿瘤抗原基因进行表达分析

通过 RT-PCR 法分析以上6种基因在不同的癌细胞系、癌症组织和正常组织中的表达。使用 Isogen 试剂(Nippon Gene)或市售可得(clontech)的来源于正常组织的 RNA,用寡 dT 引物由提取自不同癌细胞系和癌症组织的 RNA 制备 cDNA,并使用描述于表 4 的引物组合(BUB1 基因,引物 1: SEQ ID NO: 208,引物 2: SEQ ID NO: 209; C10orf3 基因,引物 1: SEQ ID NO: 210,引物 2: SEQ ID NO: 211; C20orf42 基因,引物 1: SEQ ID NO: 206,引物 2: SEQ ID NO: 207; Lengsin 基因,引物 1: SEQ ID NO: 202,引物 2: SEQ ID NO: 203; HIFPH3 基因,引物 1: SEQ ID NO: 212,引物 2: SEQ ID NO: 213; BJ-TSA-9 基因,引物 1: SEQ ID NO: 204,引物 2: SEQ ID NO: 205)通过 PCR 反应(40 个循环)扩增。此后,通过电泳分离扩增的 cDNA,以进行分析。作为阳性对照,以和上述相同的方式通过 RT-PCR 法证实 G3PDH 基因的表达。所用的一部分癌症和正常组织在得到知情同

意后主要由手术样品制备。结果示于图 1-6。

[0230][表 4]

基因	引物 1	引物 2
lengsin	tggcactggaagaagatcaa	tcaccacaactttgttgtttgtt
BJ-TSA-9	cagatccaggtgcctctgac	tcaggtaatccaaccaccttg
C20orf42	tgaatttctgaggccttgct	tgttctcagcagcaaacagg
BUB1	ttatctgctggcttggcact	gcttttgccttaacaaatcca
C10orf3	tgtccattgtaagaggtggtg	tgagagggctacatgggttt
HIFPH3	catccctgtcttgtgtgtgg	ccaacagccctggattaaga

[0231] BUB1 基因在大部分正常组织中不表达或以低水平表达，但在口腔鳞状细胞癌、肾癌、大肠癌、胃癌、胰腺癌、肝癌、淋巴瘤和黑素瘤细胞系中高度表达。

C10orf3 基因在除睾丸以外的大部分正常组织中不表达或以低水平表达，但在肺腺癌、肺鳞状细胞癌和小细胞肺癌细胞系中高度表达。

C20orf42 基因在大部分正常组织中不表达，但在结肠癌和胰腺癌细胞系中高度表达。

Lengsin 基因在大部分正常组织中不表达，但在肺腺癌和肺鳞状细胞癌细胞系中高度表达。

HIFPH3 基因在大部分正常组织中表达，但在肾癌细胞系中以较高水平表达。另外，相比于正常组织，HIFPH3 基因在源自肾癌患者的许多癌症组织中以较高水平表达。

BJ-TSA-9 在大部分正常组织中不表达，但在肺腺癌、肺鳞状细胞癌、小细胞肺癌、口腔癌、胃癌、胰腺癌和淋巴瘤细胞系中高度表达。

[0232] 总之，这 6 种基因在多种癌细胞系和组织中高度表达，尽管癌症的类型彼此不同，但不在正常组织中表达或以比在癌细胞或组织中低的水平在正常组织中表达，除了 DNA 芯片分析结果以外，这些结果强有力地提示，这 6 种基因及其表达产物(即蛋白)可以用作癌症的疾病标志物(尤其是用于如上所述的癌症)。

实施例 5

[0233] 候选肽的合成和选择

(1)候选肽的选择

根据 Lengsin 的氨基酸序列(SEQ ID NO: 2), 选择由 SEQ ID NO: 13-31 和 195-201 的氨基酸序列以及 SEQ ID NO: 32-41 的氨基酸序列组成的肽, 分别作为潜在结合 HLA-A24 和 HLA-A2 分子的候选肽。

根据 BJ-TSA-9 的氨基酸序列(SEQ ID NO: 4), 选择由 SEQ ID NO: 42-49 的氨基酸序列和 SEQ ID NO: 50-58 的氨基酸序列组成的肽分别作为潜在结合 HLA-A24 和 HLA-A2 分子的候选肽。

根据 C20orf42 的氨基酸序列(SEQ ID NO: 6), 选择由 SEQ ID NO: 59-78 的氨基酸序列和 SEQ ID NO: 79-88 的氨基酸序列组成的肽, 作为分别潜在结合 HLA-A24 和 HLA-A2 分子的候选肽。

[0234] 根据 BUB1 的氨基酸序列(SEQ ID NO: 8), 选择由 SEQ ID NO: 89-117 的氨基酸序列和 SEQ ID NO: 118-157 的氨基酸序列组成的肽, 作为分别潜在结合 HLA-A24 和 HLA-A2 分子的候选肽。

根据 C10orf3 的氨基酸序列(SEQ ID NO: 10), 选择由 SEQ ID NO: 158-165 的氨基酸序列和 SEQ ID NO: 166-175 的氨基酸序列组成的肽, 作为分别潜在结合 HLA-A24 和 HLA-A2 分子的候选肽。

根据 HIFPH3 的氨基酸序列(SEQ ID NO: 12), 选择由 SEQ ID NO: 176-183 的氨基酸序列和 SEQ ID NO: 184-194 的氨基酸序列组成的肽, 作为分别潜在结合 HLA-A24 和 HLA-A2 分子的候选肽。得到所述肽的氨基酸序列上的位置以及所述肽的长度和氨基酸序列列于以下的表 5-16。

[0235][表 5]

Lengsin A24 肽

肽	氨基酸序列	SEQ ID NO
Lengsin_295(10)	KYNYIASFFI	13
Lengsin_368(9)	RYSKDRKDL	14
Lengsin_447(9)	FYQVEPSEI	15
Lengsin_489(9)	KYELENEEI	16
Lengsin_49(9)	DMSNSNDCM	17
Lengsin_56(9)	CMRDSSQIL	18
Lengsin_72(9)	RMKHIRQAM	19
Lengsin_86(10)	QFVRFEATDL	20
Lengsin_118(10)	CMPRGYLEVI	21
Lengsin_142(9)	CFNSDIVLM	22
Lengsin_160(10)	PWADRTARVI	23
Lengsin_172(10)	TFTVTGEPLL	24
Lengsin_199(9)	GFSLLSAFI	25
Lengsin_238(10)	PFMQELVDGL	26
Lengsin_265(10)	QMEISFLPEF	27
Lengsin_318(10)	LWDVDRKKNM	28
Lengsin_383(10)	TWGYNDNSCI	29
Lengsin_385(9)	GYNDNSCIF	30
Lengsin_479(9)	TFIRYFVAM	31
Lengsin_149(11)	LMPELSTFRVL	195
Lengsin_326(12)	NMFCSTSGTEQL	196
Lengsin_486(12)	AMKKYELENEEI	197
Lengsin_185(11)	RYIAKRQLSHL	198
Lengsin_247(11)	LYHTGANVESF	199
Lengsin_482(11)	RYFVAMKKYEL	200
Lengsin_207(12)	IYDFCIFGVPEI	201

[0236][表 6]

Lengsin A2 肽

肽	氨基酸序列	SEQ ID NO
Lengsin_313(9)	ILSHSLWDV	32
Lengsin_239(9)	FMQELVDGL	33
Lengsin_246(9)	GLYHTGANV	34
Lengsin_206(10)	FIYDFCIFGV	35
Lengsin_149(10)	LMPELSTFRV	36
Lengsin_79(10)	AMAKNRLQFV	37
Lengsin_312(10)	GILSHSLWDV	38
Lengsin_270(10)	FLPEFGISSA	39
Lengsin_231(10)	FLNNHDQPFM	40
Lengsin_336(10)	QLTITGKKWL	41

[0237][表 7]

BJ-TSA-9 A24 肽

肽	氨基酸序列	SEQ ID NO
BJ-TSA-9_20(10)	QWVRPARADF	42
BJ-TSA-9_120(10)	SWASAEKPYL	43
BJ-TSA-9_138(10)	YFQTVKHNNI	44
BJ-TSA-9_165(10)	LMDVFTDVEI	45
BJ-TSA-9_227(10)	IYCAKSGRKF	46
BJ-TSA-9_273(9)	KFTGQAVEL	47
BJ-TSA-9_281(9)	LFDEEFRHL	48
BJ-TSA-9_289(9)	LYASSKPVM	49

[0238][表 8]

BJ-TSA-9 A2 肽

肽	氨基酸序列	SEQ ID NO
BJ-TSA-9_164(9)	ILMDVFTDV	50
BJ-TSA-9_252(9)	VLSGSYSFT	51
BJ-TSA-9_100(9)	LQSGTYFPV	52
BJ-TSA-9_99(10)	SLQSGTYFPV	53
BJ-TSA-9_191(10)	VLLDQGGVKL	54
BJ-TSA-9_199(10)	KLFQEMCDKV	55
BJ-TSA-9_62(10)	FLSSVEAQYI	56
BJ-TSA-9_270(10)	ILSKFTGQAV	57
BJ-TSA-9_163(10)	VILMDVFTDV	58

[0239][表 9]

C20orf42 A24 肽

肽	氨基酸序列	SEQ ID NO
C20orf42_79(10)	KYGVQADAKL	59
C20orf42_182(9)	LYSKTMTPI	60
C20orf42_283(10)	KYDAVRINQL	61
C20orf42_292(10)	LYEQARWAIL	62
C20orf42_319(9)	QYHISKLSL	63
C20orf42_381(10)	SYQPEVLNIL	64
C20orf42_655(9)	EYIGGYIFL	65
C20orf42_52(9)	VMLKLVEQI	66
C20orf42_68(9)	CWLLKTHWTL	67
C20orf42_89(9)	LFTPQHKML	68
C20orf42_202(10)	TMTWFSDSPL	69
C20orf42_249(9)	GWLDSSRSL	70
C20orf42_271(10)	RFKYYSFFDL	71
C20orf42_392(10)	YWFIFKDTSI	72
C20orf42_460(10)	QYAQWMAACM	73
C20orf42_510(9)	NMDMNPECF	74
C20orf42_545(9)	QMPLVEAKL	75
C20orf42_554(9)	RFIQAWQSL	76
C20orf42_558(10)	AWQSLPEFGL	77
C20orf42_603(10)	TWRFTNIKQW	78

[0240][表 10]

C20orf42 A2 肽

肽	氨基酸序列	SEQ ID NO
C20orf42_69(9)	WLLKTHWTL	79
C20orf42_361(9)	SLLEDITDI	80
C20orf42_377(9)	KLFRPKKLL	81
C20orf42_417(9)	KLNLRGCEV	82
C20orf42_300(9)	ILLEEIDCT	83
C20orf42_610(9)	KQWNVNWET	84
C20orf42_621(9)	VVIEFDQNV	85
C20orf42_291(10)	QLYEQARWAI	86
C20orf42_95(10)	KMLRLRLPNL	87
C20orf42_88(10)	LLFTPQHKML	88

[0241][表 11]

Bub1 A24 肽

肽	氨基酸序列	SEQ ID NO
Bub1_43(10)	EYLITLLEHL	89
Bub1_61(9)	KYHNDPRFI	90
Bub1_89(9)	LYNHGIGTL	91
Bub1_101(9)	LYIAWAGHL	92
Bub1_139(9)	QYRLFQTRL	93
Bub1_670(9)	MYSASLLRL	94
Bub1_801(9)	VYEATQGDL	95
Bub1_870(9)	SYGTLLNAI	96
Bub1_1061(9)	HYTNKIRAL	97
Bub1_29(10)	RYIQWVEENF	98
Bub1_77(10)	EYNSDLHQFF	99
Bub1_206(9)	NMERRVITI	100
Bub1_235(9)	VMYCKEKLI	101
Bub1_292(9)	KMDELHKKL	102
Bub1_353(9)	TYQQTPVNM	103
Bub1_414(10)	EFKPQSGAEI	104
Bub1_474(10)	MFQAPTLPDI	105
Bub1_552(10)	TFGERSVSRL	106
Bub1_574(9)	EFLDDSTVW	107
Bub1_581(10)	VWGIRCNKTL	108
Bub1_741(10)	PWDDKLIFKL	109
Bub1_837(9)	LMERLKPSM	110
Bub1_880(10)	LYKNTPEKVM	111
Bub1_922(9)	NFILGNGFL	112
Bub1_952(10)	DMKLFPKGTI	113
Bub1_995(10)	VYCMLFGTYM	114
Bub1_1018(10)	LFRRLPHLDM	115
Bub1_1037(10)	MWNEFFHVML	116
Bub1_1057(10)	VFQQHYTNKI	117

[0242][表 12]

Bub1 A2 肽

肽	氨基酸序列	SEQ ID NO
Bub1_781(9)	QLGSKLYYV	118
Bub1_998(9)	MLFGTYMKV	119
Bub1_1026(9)	DMWNEFFHV	120
Bub1_392(9)	TVTDSMFAV	121
Bub1_900(9)	RMLYMJEQV	122
Bub1_1017(9)	GLFRRLPHL	123
Bub1_903(9)	YMIEQVHDC	124
Bub1_750(9)	LLSGLSKPV	125
Bub1_44(9)	YLITLLEHL	126
Bub1_88(9)	FLYNHGIGT	127
Bub1_292(9)	KMDELHKKL	128
Bub1_785(9)	KLVYVHLL	129
Bub1_268(9)	WVNEDRHYM	130
Bub1_699(9)	RLTDTDAAI	131
Bub1_943(9)	ALIDLQSI	132
Bub1_746(9)	LIFKLLSGL	133
Bub1_235(9)	VMYCKEKLI	134
Bub1_68(9)	FITYCLKFA	135
Bub1_613(9)	KLPVESVHI	136
Bub1_471(9)	IMNMFQAPT	137
Bub1_780(10)	FQLGSKLVYV	138
Bub1_575(10)	FLDDSTVWGI	139
Bub1_792(10)	LLGEGAFQV	140
Bub1_749(10)	KLLSGLSKPV	141
Bub1_25(10)	GEWERYIQWV	142
Bub1_954(10)	KLFPKGTIFT	143
Bub1_997(10)	CMLFGTYMKV	144
Bub1_745(10)	KLIFKLLSGL	145
Bub1_686(10)	VLTCEAELGV	146
Bub1_879(10)	NLYKNTPEKV	147
Bub1_88(10)	FLYNHGIGTL	148
Bub1_287(10)	QLLKQKMDL	149
Bub1_836(10)	QLMERLKPSM	150
Bub1_613(10)	KLPVESVHIL	151
Bub1_722(10)	WMQMSSLGTV	152
Bub1_470(10)	FIMNMFQAPT	153
Bub1_141(10)	RLFQTRLTET	154
Bub1_982(10)	WNYQIDYFGV	155
Bub1_447(10)	GMVQATPSKV	156
Bub1_1048(10)	DLLRQKLKKV	157

[0243][表 13]

C10orf3 A24 肽

肽	氨基酸序列	SEQ ID NO
C10orf3_193(10)	VYVKGLLAKI	158
C10orf3_446(10)	QYPATEHRDL	159
C10orf3_169(10)	EMEIQLKDAL	160
C10orf3_355(9)	QMQACTLDF	161
C10orf3_193(10)	VYVKGLLAKI	162
C10orf3_446(10)	QYPATEHRDL	163
C10orf3_169(10)	EMEIQLKDAL	164
C10orf3_355(9)	QMQACTLDF	165

[0244][表 14]

C10orf3 A2 肽

肽	氨基酸序列	SEQ ID NO
C10orf3_197(9)	GLLAKIFEL	166
C10orf3_376(9)	QLLVILKEL	167
C10orf3_99(9)	ALLEQLEET	168
C10orf3_341(9)	KQQEEQTRV	169
C10orf3_228(9)	YLQEEKQKC	170
C10orf3_392(9)	TQLESLKQL	171
C10orf3_282(10)	LLYSQRRADV	172
C10orf3_50(10)	KLTDKERHRL	173
C10orf3_350(10)	ALLEQQMQAC	174
C10orf3_328(10)	LLSQVQFLYT	175

[0245][表 15]

HIFPH3 A24 肽

肽	氨基酸序列	SEQ ID NO
HIFPH3_6(9)	IMRLDLEKI	176
HIFPH3_27(9)	GFCYLDNFL	177
HIFPH3_92(10)	SFLLSLIDRL	178
HIFPH3_112(10)	YYVKERSKAM	179
HIFPH3_155(10)	NWDAKLHGGI	180
HIFPH3_167(9)	IFPEGKFSI	181
HIFPH3_173(9)	SFIADVEPI	182
HIFPH3_295(9)	RYAMTVWYF	183

[0246][表 16]

HIFPH3 A2 肽

肽	氨基酸序列	SEQ ID NO
HIFPH3_93(9)	FLLSLIDRL	184
HIFPH3_30(9)	YLDNFLGEV	185
HIFPH3_18(9)	YIVPCLHEV	186
HIFPH3_159(9)	KLHGGILRI	187
HIFPH3_60(9)	QLAGPRAGV	188
HIFPH3_42(9)	CVLERVKQL	189
HIFPH3_22(10)	CLHEVGFCYL	190
HIFPH3_34(10)	FLGEVVGDCV	191
HIFPH3_96(10)	SLIDRLVLYC	192
HIFPH3_93(10)	FLLSLIDRLV	193
HIFPH3_151(10)	YLNKNWDAKL	194

[0247] (2)肽合成

在如上所述的肽(SEQ ID NO: 13-201)中, 通过 Fmoc 法合成得自 Lengsin 的 26 种肽(表 5)、得自 BJ-TSA-9 的 7 种肽(表 7)、得自 C20orf42 的 20 种肽(表 9)和得自 C10orf3 的 4 种肽(表 13)。

实施例 6

[0248] 评价得自不同抗原的肽对 HLA-A*2402 和 HLA-A*0201 的结合亲和性

通过如在文献(J. Immunol. 164:2565, 2000)中所述的方法, 测定在实施例 1 中合成的肽对 HLA-A*2402 的结合亲和性。通过将由 HLA-A*2402 和 H-2Kb 组成的嵌合 MHC 基因稳定引入到没有 MHC I 类分子的小鼠淋巴瘤细胞系 RMA-S 中获得细胞系 RMA-S-A*2402 细胞, 将该细胞于 26°C 温育 18 小时。用 PBS 溶液洗涤 RMA-S-A*2402 细胞, 悬浮在含有 3 μ L/mL 人 β_2 -微球蛋白和 100 μ L/mL 每种肽的培养溶液 OPTI-MEM (Invitrogen) 中, 并于 26°C 温育 3 小时和于 37°C 温育 3 小时。用 PBS 溶液洗涤细胞, 并用抗-HLA-A24 抗体于 4°C 处理 30 分钟。此外, 用 PBS 溶液洗涤细胞, 并用 PE-标记的抗小鼠 IgG 抗体于 4°C 处理 30 分钟。洗涤细胞, 并悬浮在含有 1%福尔马林的 1 ml

PBS 溶液中进行固定。通过用于流式细胞术的装置 FACScan (BD Bioscience)检测细胞,由平均荧光强度获得肽的结合亲和力。在图 7-9 中显示了所检验的 31 种肽的结合亲和力。

[0249] EB 病毒来源的肽(EBV)和 HIV 病毒来源的肽(HIV)已被报告结合 HLA-A*2402 (分别为 J. Immunol. 158:3325, 1997 和 J. Immunol. 164:2565, 2000), 用作阳性对照,表现出强结合活性。卵清蛋白来源的肽(SL8)已被报告结合 H2-Kb (Eur J Immunol. 21:2891, 1991), 用作阴性对照,显示出弱结合活性。所检验的 31 种肽都被证实全部以高度优先的方式结合 HLA-A*2402。因此,发现这 31 种肽为肿瘤抗原肽,得自这些肽的蛋白 BJ-TSA-9、C20orf42 和 C10orf3 为肿瘤抗原蛋白。

[0250] 同样,得自 Lengsin 的 23 种肽对 HLA-A*2402 的结合亲和性示于图 10 和 11。SEQ ID NO: 15、16、21、22、23、24、25、26、27、30、195、197、198、199、200 和 201 的肽显示出对 HLA-A*2402 的显著结合亲和性。因此,发现这些肽为肿瘤抗原肽,得自这些肽的蛋白 Lengsin 为肿瘤抗原蛋白。

对 HLA-A*0201 的结合亲和性可以通过使用人淋巴瘤细胞系 T2 以相同方式证实。另外,可类似地证实得自 BUB1 和 HIFPH3 的肽对 HLA 的结合亲和性。

实施例 7

[0251] 肿瘤抗原来源的肽的 CTL 诱导

通过评价在实施例 6 中显示出对 HLA 的结合亲和性的肽诱导的 CTL 活性,可以证实以下事实:蛋白 BUB1、C10orf3、C20orf42、Lengsin、HIFPH3 和 BJ-TSA-9 为肿瘤抗原蛋白,其部分肽为肿瘤抗原肽。肽的 CTL 诱导和所诱导 CTL 的活性评价可以如在文献(J. Immunol. 169:1611, 2002, WO 02/47474, Int J. Cancer:100, 565-570 (2002))中所述使用人外周血单核细胞或用于人的模型动物按照已知

测定进行。就使用人外周血单核细胞的测定而言，在得到知情同意后由癌症患者或健康个体收集外周血，通过密度梯度离心法分离单核细胞，并在 AIM-V 培养溶液(Invitrogen)中培养。在 24 小时培养后，收集非粘附细胞，并在含 100 U/mL IL-2 的 AIM-V 中培养。为了制备抗原提呈细胞，在含有 1000 U/mL IL-4 和 1000 U/mL GM-CSF 的 AIM-V 培养物溶液中培养粘附细胞达 5 天，在加入每种不同的肿瘤抗原肽后再培养 1 天，在加入 10 ng/mL TNF 和 1000 U/mL IFN- α 后再培养 1 天。非粘附细胞与用所述肽脉冲的抗原提呈细胞一起培养。在通过使用肽脉冲的抗原提呈细胞进行第一次肽刺激后的第 7 和 14 天，含有肽脉冲的 CTL 的非粘附细胞接受第二次和第三次肽刺激。在第三次刺激后 1 周，将细胞与用放射性物质如 ^{51}Cr 或非放射性物质标记的靶细胞(表达与要给予的肽结合的 HLA 此外还提供要给予的肽的肿瘤抗原或要给予的肽自身的细胞，或者预先用对应肽脉冲的细胞)一起共培养，然后由在条件培养基中的放射性或非放射性物质的量评价 CTL 对靶细胞的裂解，以确定 CTL 活性。还可以使用 ELISA 或 ELISPOT，通过检测在条件培养基中作为 CTL 和靶细胞之间响应的结果而产生的 IFN- γ ，评价 CTL 活性。

[0252] 在使用用于人的模型动物评价该活性时，将稳定形成的肿瘤抗原肽给予表达人 HLA 的转基因小鼠，在约 1 周后，获得脾细胞或其它淋巴组织。用与给予肽相同的肽在体外刺激获得的细胞，并培养约 5 天。所获细胞被视为对应于 CTL 群体，CTL 活性可如上所述评价。

实施例 8

[0253] C10orf3 来源的肽的 CTL 诱导

如在实施例 7 中所述，用 C10orf3 来源的肽 C10orf3_193(10)刺激 HLA-A*2402-阳性的癌症患者的外周血单核细胞(PBMC)，以诱导 CTL。通过使用 T2A24 细胞(通过将 HLA-A*2402 基因导入不具有该

基因的 T2 细胞(ATCC 编号 CRL-1992)中产生)以及对 HLA-A*2402 和 C10orf3 为阳性的癌细胞系 Sw480 作为靶细胞测定 CTL 活性,结果示于表 12。诱导的 CTL 克隆显示出肽特异性细胞毒活性。另外,诱导的 CTL 克隆杀死 HLA-A*2402 和 C10orf3 阳性癌细胞系,但不杀死 HLA 阴性的 K562 细胞。该结果证实,肽 C10orf3_193(10)(SEQ ID NOS: 158、162)是肿瘤抗原肽, C10orf3 也是肿瘤抗原蛋白。

实施例 9

[0254] Lengsin 来源的肽的 CTL 诱导

混合一些潜在结合 HLA-A*2402 的合成肽,并如在实施例 7 中所述,用于刺激 HLA-A*2402 阳性的癌症患者的 PBMC,以诱导 CTL。在诱导后,利用 ELISPOT 通过检测 IFN- γ 来测量对每种肽的响应,结果示于图 13。SEQ ID NO: 22、23、26、27、198、200 和 201 的肽诱导肽特异性 CTL。这些结果证实,SEQ ID NO: 22、23、26、27、198、200 和 201 的肽为肿瘤抗原肽,还证实 Lengsin 为肿瘤抗原蛋白。

然后,用 SEQ ID NO: 22 的肽刺激癌症患者的 PBMC,以诱导 CTL,并检测作为细胞毒活性的 CTL 诱导活性。结果示于图 14。CTL 显示出特异性针对靶细胞的细胞毒活性,所述靶细胞用与用于刺激相同的肽脉冲。

实施例 10

[0255] HIFPH3 来源的肽的 CTL 诱导

混合 SEQ ID NO: 177、178、179 和 183 的肽,并如在实施例 7 中所述,用于刺激 HLA-A*2402 阳性的癌症患者的 PBMC,以诱导 CTL。在诱导后,检测针对不同肾癌细胞系的细胞毒活性,结果示于图 15。用所述肽刺激的淋巴细胞显示出针对 HLA-A24⁺ HIFPH3⁺ SMKT R-1 细胞但不针对 HLA-A24⁻ HIFPH3⁺ SMKT R-4 细胞、HLA-A24⁺ HIFPH3⁻ Caki-1 细胞、HLA-A24⁻ HIFPH3⁻ ACHN 细胞和

HLA` K562 细胞的细胞毒活性。这些结果表明，针对 HIFPH3 的 HLA-A24-限制性的 CTL 通过所述肽刺激诱导，然后证实肽 HIFPH3_27(9)(SEQ ID NO: 177)、HIFPH3_92(10)(SEQ ID NO: 178)、HIFPH3_112(10)(SEQ ID NO: 179)和 HIFPH3_295(9)(SEQ ID NO: 183) 为肿瘤抗原肽，HIFPH3 为肿瘤抗原蛋白。

实施例 11

[0256] C20orf42 来源的肽的 CTL 诱导

将 20 种针对 HLA-A*2402 合成的肽分为 5 组，每组含有 4 种肽。将 HLA-A*2402 阳性的癌症患者的 PBMC 铺板在 96 孔板中，并按照实施例 7 用各组肽的每一种刺激，以诱导 CTL。当用各组刺激在 384 孔中的 PBMC 时，检查针对用于刺激的各种肽的细胞毒活性，检测 8-14 个孔中对所述肽特异性的细胞毒活性。该结果证实，C20orf42 来源的肽为肿瘤抗原肽，C20orf42 为肿瘤抗原蛋白。

工业适用性

[0257] 本发明提供新的肿瘤抗原蛋白和肽及其在癌症免疫领域中的用途。肿瘤抗原蛋白和由其获得的的肽可以用于治疗表达所述肿瘤抗原蛋白的癌症患者。

附图简述

[0258] [图 1]图 1 显示了通过 RT-PCR 分析的 Lengsin 基因的 mRNA 在主要的不同正常组织和肺癌细胞系中的表达水平。在该图中，Lengsin 和 G3PDH (阳性对照)的箭头表示分别显示所述基因的 mRNA 水平的条带位置。不同的正常组织或癌症组织如下：心脏、脑、胎盘、肺、肝、骨骼肌、肾脏、胰腺、脾脏、胸腺、前列腺、睾丸、卵巢、小肠、大肠、PBMC、腺癌、鳞状细胞癌和小细胞癌。以下为不同肺癌细胞系的名称：LNY-1、A549、LHK2、1-87、LK79、Sq-1、LC817 和 Lu65。

[0259] [图 2]图 2 显示了通过 RT-PCR 分析的 BJ-TSA-9 基因的 mRNA 在主要的不同正常组织以及肺、口腔和其它癌细胞系中的表达水平。在该图中, BJ-TSA-9 和 G3PDH (阳性对照)的箭头表示分别显示所述基因的 mRNA 水平的条带位置。不同的正常组织或癌症组织如下: 心脏、脑、胎盘、肺、肝、骨骼肌、肾脏、胰腺、脾脏、胸腺、前列腺、睾丸、卵巢、小肠、大肠、PBMC、腺癌、鳞状细胞癌和小细胞癌。以下为起源于不同器官的癌细胞系的名称: LNY-1、A549、LHK2、1-87、LK79、Sq-1、LC817、Lu65、OSC19、OSC20、OSC30、OSC40、OSC70、HSC2、HSC3、HSC4、KOSC3、HO-1-N-1、R3、SW450、SSTW、HS776、CHC20、CHC32、C1R、T2、888mel 和 LG2mel。

[0260] [图 3]图 3 显示了通过 RT-PCR 分析的 C20orf42 基因的 mRNA 在主要的不同正常组织以及结肠和胰腺癌细胞系中的表达水平。在该图中, C20orf42 和 G3PDH (阳性对照)的箭头表示分别显示所述基因的 mRNA 水平的条带位置。不同的正常组织或癌症组织如下: 心脏、脑、胎盘、肺、肝、骨骼肌、肾脏、胰腺、脾脏、胸腺、前列腺、睾丸、卵巢、小肠、结肠、白细胞、结肠癌和胰腺癌。以下为不同的结肠和胰腺癌细胞系的名称: SW480、SW620、colo205、WiDR、CFPAC、PK8、SuSu86 和 PUN。

[0261] [图 4]图 4 显示了通过 RT-PCR 分析的 BUB1 基因的 mRNA 在主要的不同正常组织和口腔鳞状细胞癌、肾癌、大肠癌、胃癌、胰腺癌、肝癌、淋巴瘤和黑素瘤细胞系中的表达水平。在该图中, BUB1 和 G3PDH (阳性对照)的箭头表示分别显示所述基因的 mRNA 水平的条带位置。以下为起源于不同器官的癌细胞系的名称: OSC19、HSC2、Caki1、R3、SW450、SSTW、HS776T、PUN、CHC20、CHC32、C1R、T2、888mel 和 LG2MEL。

[0262] [图 5]图 5 显示了通过 RT-PCR 分析的 C10orf3 基因的 mRNA 在主要的不同正常组织和肺癌细胞系中的表达水平。在该图中, C10orf3 和 G3PDH (阳性对照)的箭头表示分别显示所述基因的

mRNA 水平的条带位置。不同的正常组织或癌症组织如下：心脏、脑、胎盘、肺、肝、骨骼肌、肾脏、胰腺、脾脏、胸腺、前列腺、睾丸、卵巢、小肠、大肠、PBMC、腺癌和鳞状细胞癌、小细胞癌。以下为起源于不同器官的癌细胞系的名称：LNY-1、A549、LHK2、1-87、LK79、Sq-1、LC817 和 Lu65。

[0263] [图 6]图 6 显示了通过 RT-PCR 分析的 HIFPH3 基因的 mRNA 在主要的不同正常组织和肾癌组织以及肾癌细胞系中的表达水平。在该图中，HIFPH3 和 G3PDH 与 β -肌动蛋白(阳性对照)的箭头表示分别显示所述基因的 mRNA 水平的条带位置。不同的正常组织或癌症组织如下：心脏、脑、胎盘、肺、肝、骨骼肌、肾脏、胰腺、脾脏、胸腺、前列腺、睾丸、卵巢、小肠、结肠、白细胞、RCC (肾细胞癌)。T 和 N 分别意指肿瘤和正常组织。以下为肾癌细胞系的名称：SMKTR-1、SMKTR-2、SMKTR-3、SMKTR-4、Caki-1 和 ACHN。

[0264] [图 7]图 7 显示了得自 BJ-TSA-9 的 8 种肽、EB 病毒来源的肽(EBV)以及 HIV 来源的肽(HIV)(阳性对照)和卵清蛋白来源的肽(SL8)(阴性对照)对 HLA-A*2402 的结合亲和性。在该图中，水平轴指示平均荧光强度(MFI, 对应于结合亲和性)，垂直轴指示肽的名称。未加入肽所获得的结果记录为(-)。

[0265] [图 8]图 8 显示了得自 C20orf42 的 20 种肽、EB 病毒来源的肽(EBV)以及 HIV 来源的肽(HIV)(阳性对照)和卵清蛋白来源的肽(SL8)(阴性对照)对 HLA-A*2402 的结合亲和性。在该图中，水平轴指示平均荧光强度(MFI, 对应于结合亲和性)，垂直轴指示肽的名称。未加入肽所获得的结果记录为(-)。

[0266] [图 9]图 9 显示了得自 C10orf3 的 4 种肽、EB 病毒来源的肽(EBV)以及 HIV 来源的肽(HIV)(阳性对照)和卵清蛋白来源的肽(SL8)(阴性对照)对 HLA-A*2402 的结合亲和性。在该图中，水平轴指示平均荧光强度(MFI, 对应于结合亲和性)，垂直轴指示肽的名称。未加入肽所获得的结果记录为(-)。

[0267] [图 10]图 10 显示了得自 Lengsin 的 16 种肽、HIV 来源的肽(HIV)(阳性对照)和卵清蛋白来源的肽(SL8)(阴性对照)对 HLA-A*2402 的结合亲和性。在该图中,水平轴指示平均荧光强度(MFI,对应于结合亲和性),垂直轴指示肽的名称。未加入肽所获得的结果记录为(-)。

[0268] [图 11]图 11 显示了得自 Lengsin 的 7 种肽、EB 病毒来源的肽(EBV)(阳性对照)和卵清蛋白来源的肽(SL8)(阴性对照)对 HLA-A*2402 的结合亲和性。在该图中,水平轴指示平均荧光强度(MFI,对应于结合亲和性),垂直轴指示肽的名称。未加入肽所获得的结果记录为(-)。

[0269] [图 12]图 12 显示了通过用 C10orf3_193(10)肽刺激 HLA-A*2402 阳性的癌症患者的 PBMC 诱导的 CTL 的细胞毒活性。在该图中,“肽 2”和“Sw480”分别表示使用以所述肽脉冲的 T2-A24 细胞以及对 HLA-A*2402 和 C10orf3 为阳性的癌细胞系 Sw480 作为靶细胞的结果。同样,“K562”和“(-)”分别表示使用 HLA-阴性的 K562 细胞和未加入肽的 T2-A24 细胞作为靶细胞的结果。

[0270] [图 13]图 13 显示了用 ELOSPOT 通过检测 INF- γ 测量用 Lengsin 来源的肽刺激 HLA-A*2402 阳性的癌症患者的 PBMC 诱导的 CTL 响应。

[0271] [图 14]图 14 显示了通过用 Lengsin_142(9)肽刺激癌症患者的 PBMC 诱导的 CTL 的细胞毒活性。在该图中,“T2A24 + 肽”和“T2A24”分别表示使用所述肽脉冲和未脉冲的 T2-A24 细胞作为靶细胞的结果。同样,“T2A24 + HIV”和“K562”分别表示使用 HIV 来源的肽和 HLA 阴性 K562 细胞脉冲的 T2-A24 细胞作为靶细胞的结果。

[0272] [图 15]图 15 图示了通过用 HIFPH3 来源的肽(SEQ ID NO: 177、178、179 和 183)的混合物刺激 HLA-A*2402 阳性的癌症患者的 PBMC 诱导 CTL 的结果。垂直轴指示细胞毒活性。在该图中,

“E:T1Cr%”、“E:T3Cr%”和“E:T10Cr%”意指所述肽刺激的PBMC和癌细胞之间的比率分别为1、3和10。

序列表自由文本

[0273] SEQ ID NO: 13-201 的氨基酸序列是合成肽。

SEQ ID NO: 202-213 的碱基序列是引物。

- <110> Sato, Noriyuki
Dainippon Sumitomo Pharma Co., Ltd. 公司
- <120> 新肿瘤抗原肽
- <130> 2007002W01
- <150> JP 2006-086820
<151> 2006-03-28
- <160> 213
- <170> PatentIn version 3.1
- <210> 1
<211> 1947
<212> DNA
<213> 人
- <220>
<221> CDS
<222> (35)..(1564)
<223>
- <400> 1
gtggtttcta acgctgaact taaaaagtgt tgag atg aat aat gaa gag gac ott 55
Met Asn Asn Glu Glu Asp Leu
1 5
- ctg cag gag gac tca aca aga gat gaa ggc aat gag act gaa gcc aac 103
Leu Gln Glu Asp Ser Thr Arg Asp Glu Gly Asn Glu Thr Glu Ala Asn
10 15 20
- agc atg aac aca tta aga agg aca agg aag aaa gtc act aaa cca tat 151
Ser Met Asn Thr Leu Arg Arg Thr Arg Lys Lys Val Thr Lys Pro Tyr
25 30 35

gtt tgt tca act gaa gtg gga gaa acg gat atg tcc aat tca aat gat	199
Val Cys Ser Thr Glu Val Gly Glu Thr Asp Met Ser Asn Ser Asn Asp	
40 45 50 55	
tgc atg agg gac agc agt caa att ttg acc cca cct caa ctc tct tct	247
Cys Met Arg Asp Ser Ser Gln Ile Leu Thr Pro Pro Gln Leu Ser Ser	
60 65 70	
aga atg aaa cac att aga caa gcc atg gcc aaa aat cgc ctc cag ttt	295
Arg Met Lys His Ile Arg Gln Ala Met Ala Lys Asn Arg Leu Gln Phe	
75 80 85	
gta cga ttt gaa gca aca gac ctc cac ggc gtg tcc agg tct aag act	343
Val Arg Phe Glu Ala Thr Asp Leu His Gly Val Ser Arg Ser Lys Thr	
90 95 100	
atc cct gca cac ttt ttt caa gag aaa gtg agc cat ggt gtt tgc atg	391
Ile Pro Ala His Phe Phe Gln Glu Lys Val Ser His Gly Val Cys Met	
105 110 115	
ccc cga ggt tat ctt gaa gtg ata cca aat cca aag gac aat gaa atg	439
Pro Arg Gly Tyr Leu Glu Val Ile Pro Asn Pro Lys Asp Asn Glu Met	
120 125 130 135	
aat aac ata aga gcc aca tgt ttt aat agc gac ata gtc cta atg cca	487
Asn Asn Ile Arg Ala Thr Cys Phe Asn Ser Asp Ile Val Leu Met Pro	
140 145 150	
gag tta tca acc ttt aga gtt ttg cca tgg gct gac aga act gca aga	535
Glu Leu Ser Thr Phe Arg Val Leu Pro Trp Ala Asp Arg Thr Ala Arg	
155 160 165	
gtg ata tgt gat acc ttc act gtg act ggt gag cct ctt ttg act tcc	583
Val Ile Cys Asp Thr Phe Thr Val Thr Gly Glu Pro Leu Leu Thr Ser	
170 175 180	
cca agg tac att gca aag agg cag ctg agc cat ctg cag gcc tct ggc	631
Pro Arg Tyr Ile Ala Lys Arg Gln Leu Ser His Leu Gln Ala Ser Gly	
185 190 195	
ttt tcc ctg ctt tct gct ttc atc tat gat ttt tgc att ttt ggt gtg	679

Phe Ser Leu Leu Ser Ala Phe Ile Tyr Asp Phe Cys Ile Phe Gly Val	
200	205 210 215
ccc gaa att tta aat tca aag att ata tet ttt cct gct tta aca ttt	727
Pro Glu Ile Leu Asn Ser Lys Ile Ile Ser Phe Pro Ala Leu Thr Phe	
	220 225 230
tta aat aac cat gat cag ccc ttc atg cag gaa ctt gtt gat ggc ttg	775
Leu Asn Asn His Asp Gln Pro Phe Met Gln Glu Leu Val Asp Gly Leu	
	235 240 245
tat cac act gga gcc aat gtc gag agt ttt tcc tcc tct acc agg cct	823
Tyr His Thr Gly Ala Asn Val Glu Ser Phe Ser Ser Ser Thr Arg Pro	
	250 255 260
ggt cag atg gaa atc tct ttc ctg cct gaa ttt ggc att agc tca gct	871
Gly Gln Met Glu Ile Ser Phe Leu Pro Glu Phe Gly Ile Ser Ser Ala	
	265 270 275
gat aat gca ttt acc ctc aga aca ggt gtc aaa gaa gtg gca agg aaa	919
Asp Asn Ala Phe Thr Leu Arg Thr Gly Val Lys Glu Val Ala Arg Lys	
	280 285 290 295
tat aat tac att gcc agc ttc ttc att gag act gga ttt tgt gat tca	967
Tyr Asn Tyr Ile Ala Ser Phe Phe Ile Glu Thr Gly Phe Cys Asp Ser	
	300 305 310
ggg att ttg tct cat agt ctc tgg gat gtc gat agg aag aaa aac atg	1015
Gly Ile Leu Ser His Ser Leu Trp Asp Val Asp Arg Lys Lys Asn Met	
	315 320 325
ttc tgc agc act tet gga act gag cag ctc acg atc act ggg aaa aaa	1063
Phe Cys Ser Thr Ser Gly Thr Glu Gln Leu Thr Ile Thr Gly Lys Lys	
	330 335 340
tgg ttg gca gga ctc ttg aag cac tct gct gcg ctc agc tgc ctg atg	1111
Trp Leu Ala Gly Leu Leu Lys His Ser Ala Ala Leu Ser Cys Leu Met	
	345 350 355
gcg cct tct gtt agc tgc cga aag cgt tat tcc aag gac agg aaa gac	1159
Ala Pro Ser Val Ser Cys Arg Lys Arg Tyr Ser Lys Asp Arg Lys Asp	
	360 365 370 375

ctg aag aag agt gtg cct aca aca tgg gga tac aat gac aac agc tgt	1207
Leu Lys Lys Ser Val Pro Thr Thr Trp Gly Tyr Asn Asp Asn Ser Cys	
380 385 390	
ata ttt aat atc aaa tgt cat gga gag aaa ggc acc cgg ata gaa aat	1255
Ile Phe Asn Ile Lys Cys His Gly Glu Lys Gly Thr Arg Ile Glu Asn	
395 400 405	
aaa cta ggc tca gca aca gca aac cct tac ttg gtg ctg gct gca act	1303
Lys Leu Gly Ser Ala Thr Ala Asn Pro Tyr Leu Val Leu Ala Ala Thr	
410 415 420	
gtt gct gcc ggc tta gat gga ctt cat agc agt aat gag gtc ttg gct	1351
Val Ala Ala Gly Leu Asp Gly Leu His Ser Ser Asn Glu Val Leu Ala	
425 430 435	
ggt cca gat gag agc aca gac ttt tac caa gtg gaa cct tct gag atc	1399
Gly Pro Asp Glu Ser Thr Asp Phe Tyr Gln Val Glu Pro Ser Glu Ile	
440 445 450 455	
cct tta aaa cta gaa gat gcc ctt gtg gca ctg gaa gaa gat caa tgc	1447
Pro Leu Lys Leu Glu Asp Ala Leu Val Ala Leu Glu Glu Asp Gln Cys	
460 465 470	
cag aga cag gct cta gga gaa acc ttt att cga tat ttt gtt gcc atg	1495
Gln Arg Gln Ala Leu Gly Glu Thr Phe Ile Arg Tyr Phe Val Ala Met	
475 480 485	
aag aaa tat gag ttg gag aat gaa gaa ata gct gca gag aga aat aaa	1543
Lys Lys Tyr Glu Leu Glu Asn Glu Glu Ile Ala Ala Glu Arg Asn Lys	
490 495 500	
ttc tta gag tat ttt att tag aatagagctc acaactactc ctttagacat	1594
Phe Leu Glu Tyr Phe Ile	
505	
gtaattgtta cttaaagcta atctacaaa acaaaaagac tgaacttttg taattaacaa	1654
cagcaacaat aacaagatta cgaatgcttt tgactttttt tttgtccat ggaatatttg	1714
acagatgaag taggactgct gagaagattc tgataacaga aacaacaac aaagttgtgg	1774

tgaagctata atatgcaaac ttaccagatc ttgtcagtc tttoctatgt gtatgttgac 1834
 ctggataaga atataccaatt ttgggtggca ataaatatat tcgcacatta agaaaaaaaa 1894
 gacttaagca ttagaaagca aaatttaa at gactaaaaaaaa aaaaaaaaaaaa aaa 1947

<210> 2

<211> 509

<212> PRT

<213> 人

<400> 2

Met Asn Asn Glu Glu Asp Leu Leu Gln Glu Asp Ser Thr Arg Asp Glu
 1 5 10 15
 Gly Asn Glu Thr Glu Ala Asn Ser Met Asn Thr Leu Arg Arg Thr Arg
 20 25 30
 Lys Lys Val Thr Lys Pro Tyr Val Cys Ser Thr Glu Val Gly Glu Thr
 35 40 45
 Asp Met Ser Asn Ser Asn Asp Cys Met Arg Asp Ser Ser Gln Ile Leu
 50 55 60
 Thr Pro Pro Gln Leu Ser Ser Arg Met Lys His Ile Arg Gln Ala Met
 65 70 75 80
 Ala Lys Asn Arg Leu Gln Phe Val Arg Phe Glu Ala Thr Asp Leu His
 85 90 95
 Gly Val Ser Arg Ser Lys Thr Ile Pro Ala His Phe Phe Gln Glu Lys
 100 105 110
 Val Ser His Gly Val Cys Met Pro Arg Gly Tyr Leu Glu Val Ile Pro
 115 120 125
 Asn Pro Lys Asp Asn Glu Met Asn Asn Ile Arg Ala Thr Cys Phe Asn
 130 135 140
 Ser Asp Ile Val Leu Met Pro Glu Leu Ser Thr Phe Arg Val Leu Pro
 145 150 155 160
 Trp Ala Asp Arg Thr Ala Arg Val Ile Cys Asp Thr Phe Thr Val Thr
 165 170 175
 Gly Glu Pro Leu Leu Thr Ser Pro Arg Tyr Ile Ala Lys Arg Gln Leu
 180 185 190
 Ser His Leu Gln Ala Ser Gly Phe Ser Leu Leu Ser Ala Phe Ile Tyr
 195 200 205
 Asp Phe Cys Ile Phe Gly Val Pro Glu Ile Leu Asn Ser Lys Ile Ile

210	215	220
Ser Phe Pro Ala Leu Thr Phe Leu Asn Asn His Asp Gln Pro Phe Met		
225	230	235
Gln Glu Leu Val Asp Gly Leu Tyr His Thr Gly Ala Asn Val Glu Ser		
	245	250
Phe Ser Ser Ser Thr Arg Pro Gly Gln Met Glu Ile Ser Phe Leu Pro		
	260	265
Glu Phe Gly Ile Ser Ser Ala Asp Asn Ala Phe Thr Leu Arg Thr Gly		
	275	280
Val Lys Glu Val Ala Arg Lys Tyr Asn Tyr Ile Ala Ser Phe Phe Ile		
	290	295
Glu Thr Gly Phe Cys Asp Ser Gly Ile Leu Ser His Ser Leu Trp Asp		
305	310	315
Val Asp Arg Lys Lys Asn Met Phe Cys Ser Thr Ser Gly Thr Glu Gln		
	325	330
Leu Thr Ile Thr Gly Lys Lys Trp Leu Ala Gly Leu Leu Lys His Ser		
	340	345
Ala Ala Leu Ser Cys Leu Met Ala Pro Ser Val Ser Cys Arg Lys Arg		
	355	360
Tyr Ser Lys Asp Arg Lys Asp Leu Lys Lys Ser Val Pro Thr Thr Trp		
	370	375
Gly Tyr Asn Asp Asn Ser Cys Ile Phe Asn Ile Lys Cys His Gly Glu		
385	390	395
Lys Gly Thr Arg Ile Glu Asn Lys Leu Gly Ser Ala Thr Ala Asn Pro		
	405	410
Tyr Leu Val Leu Ala Ala Thr Val Ala Ala Gly Leu Asp Gly Leu His		
	420	425
Ser Ser Asn Glu Val Leu Ala Gly Pro Asp Glu Ser Thr Asp Phe Tyr		
	435	440
Gln Val Glu Pro Ser Glu Ile Pro Leu Lys Leu Glu Asp Ala Leu Val		
	450	455
Ala Leu Glu Glu Asp Gln Cys Gln Arg Gln Ala Leu Gly Glu Thr Phe		
465	470	475
Ile Arg Tyr Phe Val Ala Met Lys Lys Tyr Glu Leu Glu Asn Glu Glu		
	485	490
Ile Ala Ala Glu Arg Asn Lys Phe Leu Glu Tyr Phe Ile		
	500	505

<210> 3

<211> 2946

<212> DNA

<213> 人

<220>

<221> CDS

<222> (346)..(1650)

<223>

<400> 3

```

ggaaagccgg ctcaccttcg cctccccctg cggctgggag gagaggaaat atcccatggc      60

tgactgtgcc aaggaggtgt ctgagccagc cctccccggc cgagggcagg gcaggtggcc      120

ctgagagata agccaatccc gcagctgcag atgaggagtt ctgagaagca ttgctcagga      180

cagcggtaaa tcacttcttg gaggtgccct gcacgccggt cctgggagca ggcggcctcc      240

cgggggtgcg ggagccccac tctcctgtgg tgtgttccat ttgcttccca catctggagg      300

agctgacgtg ccagcctccc ccagcaccac ccagggacgg gaggc atg agc cgg tca      357
                               Met Ser Arg Ser
                               1

agg cac ctg ggc aaa atc cgg aag cgt ctg gaa gat gtc aag agc cag      405
Arg His Leu Gly Lys Ile Arg Lys Arg Leu Glu Asp Val Lys Ser Gln
5                10                15                20

tgg gtc cgg cca gcc agg gct gac ttt agt gac aac gag agt gcc cgg      453
Trp Val Arg Pro Ala Arg Ala Asp Phe Ser Asp Asn Glu Ser Ala Arg
                25                30                35

ctg gcc acg gac gcc ctc ttg gat ggg ggt tct gaa gcc tac tgg cgg      501
Leu Ala Thr Asp Ala Leu Leu Asp Gly Gly Ser Glu Ala Tyr Trp Arg
                40                45                50

gtg ctc agc cag gaa ggc gag gtg gac ttc ttg tcc tcg gtg gag gcc      549
Val Leu Ser Gln Glu Gly Glu Val Asp Phe Leu Ser Ser Val Glu Ala
                55                60                65

cag tac atc cag gcc cag gcc agg gag ccc ccg tgt ccc cca gac acc      597
Gln Tyr Ile Gln Ala Gln Ala Arg Glu Pro Pro Cys Pro Pro Asp Thr
                70                75                80

```

ctg gga ggg gcg gaa gca ggc cct aag gga ctg gac tcc agc tcc cta Leu Gly Gly Ala Glu Ala Gly Pro Lys Gly Leu Asp Ser Ser Ser Leu 85 90 95 100	645
cag tcc ggc acc tac ttc cct gtg gcc tca gag ggc agc gag ccg gcc Gln Ser Gly Thr Tyr Phe Pro Val Ala Ser Glu Gly Ser Glu Pro Ala 105 110 115	693
cta ctg cac agc tgg gcc tca gct gag aag ccc tac ctg aag gaa aaa Leu Leu His Ser Trp Ala Ser Ala Glu Lys Pro Tyr Leu Lys Glu Lys 120 125 130	741
tcc agc gcc act gtg tac ttc cag acc gtc aag cac aac aac atc aga Ser Ser Ala Thr Val Tyr Phe Gln Thr Val Lys His Asn Asn Ile Arg 135 140 145	789
gac ctc gtc cgc cgc tgc atc acc cgg act agc cag gtc ctg gtc atc Asp Leu Val Arg Arg Cys Ile Thr Arg Thr Ser Gln Val Leu Val Ile 150 155 160	837
ctg atg gat gtg ttc acg gat gtg gag atc ttc tgt gac att cta gag Leu Met Asp Val Phe Thr Asp Val Glu Ile Phe Cys Asp Ile Leu Glu 165 170 175 180	885
gca gcc aac aag cgt ggg gtg ttc gtt tgt gtg ctc ctg gac cag gga Ala Ala Asn Lys Arg Gly Val Phe Val Cys Val Leu Leu Asp Gln Gly 185 190 195	933
ggt gtg aag ctc ttc cag gag atg tgt gac aaa gtc cag atc tct gac Gly Val Lys Leu Phe Gln Glu Met Cys Asp Lys Val Gln Ile Ser Asp 200 205 210	981
agt cac ctc aag aac att tcc atc cgg agt gtg gaa gga gag ata tac Ser His Leu Lys Asn Ile Ser Ile Arg Ser Val Glu Gly Glu Ile Tyr 215 220 225	1029
tgt gcc aag tca ggc agg aaa ttc gct ggc caa atc cgg gag aag ttc Cys Ala Lys Ser Gly Arg Lys Phe Ala Gly Gln Ile Arg Glu Lys Phe 230 235 240	1077
atc atc tcg gac tgg aga ttt gtc ctg tct gga tct tac agc ttc acc Ile Ile Ser Asp Trp Arg Phe Val Leu Ser Gly Ser Tyr Ser Phe Thr	1125

245	250	255	260	
tgg ctc tgc gga cac gtg cac cgg aac atc ctc tcc aag ttc aca ggc				1173
Trp Leu Cys Gly His Val His Arg Asn Ile Leu Ser Lys Phe Thr Gly				
	265	270	275	
cag gcg gtg gag ctg ttt gac gag gag ttc cgc cac ctc tac gcc tcc				1221
Gln Ala Val Glu Leu Phe Asp Glu Glu Phe Arg His Leu Tyr Ala Ser				
	280	285	290	
tcc aag cct gtg atg ggc ctg aag tcc ccg cgg ctg gtc gcc ccc gtc				1269
Ser Lys Pro Val Met Gly Leu Lys Ser Pro Arg Leu Val Ala Pro Val				
	295	300	305	
ccg ccc gga gca gcc ccg gcc aat ggc cgc ctt agc agc agc agt ggc				1317
Pro Pro Gly Ala Ala Pro Ala Asn Gly Arg Leu Ser Ser Ser Ser Gly				
	310	315	320	
tcc gcc agt gac cgc acg tcc tcc aac ccc ttc agc ggc cgc tcg gca				1365
Ser Ala Ser Asp Arg Thr Ser Ser Asn Pro Phe Ser Gly Arg Ser Ala				
	325	330	335	340
ggc agc cac ccc ggt acc cga agt gtg tcc cgc tct tca ggg ccc tgt				1413
Gly Ser His Pro Gly Thr Arg Ser Val Ser Ala Ser Ser Gly Pro Cys				
	345	350	355	
agc ccc gcg gcc cca cac ccg cct cca ccg ccc cgg ttc cag ccc cac				1461
Ser Pro Ala Ala Pro His Pro Pro Pro Pro Pro Arg Phe Gln Pro His				
	360	365	370	
caa ggc cct tgg gga gcc ccg agt ccc cag gcc cac ctc tcc ccg cgg				1509
Gln Gly Pro Trp Gly Ala Pro Ser Pro Gln Ala His Leu Ser Pro Arg				
	375	380	385	
ccc cac gac ggc ccg ccc gcc gct gtc tac agc aac ctg ggg gcc tac				1557
Pro His Asp Gly Pro Pro Ala Ala Val Tyr Ser Asn Leu Gly Ala Tyr				
	390	395	400	
agg ccc acg cgg ctg cag ctg gag cag ctg ggc ctg gtg ccg agg ctg				1605
Arg Pro Thr Arg Leu Gln Leu Glu Gln Leu Gly Leu Val Pro Arg Leu				
	405	410	415	420

act cca acc tgg agg ccc ttc ctg cag gcc tcc cct cac ttc tga	1650
Thr Pro Thr Trp Arg Pro Phe Leu Gln Ala Ser Pro His Phe	
425 430	
aggccccatc cctgtctgcc ctccgcaggc ccagggttgg gcactccctg agacccaaag	1710
accacacctca acgacgagtg gcgttgagcc acttcccttt gaaaagacac tcaaaatcac	1770
tgccatggtt caatgttccc aggccccagg ccatccactt gccggcccc accagttctt	1830
gggttccccg ctctagtttg acctgtgcag cacattccag aaggttccag ggagttgtg	1890
gggcagctag aggacaaaat catgaaaaca gagtccctgt ctccagaga tcatccgggg	1950
ctttaatatt aatggcccc aaaactccgt aagaagcagg aaatgcagcc caagttttac	2010
aaatgggtaa acagaggcac tgagagatag atggtagttt ggtacttctg gttcccagtg	2070
cccaggaatg gtccactccc aagaaattca ggaaagaaag actgaggaga aggtgtggga	2130
acattctgga tgtttcggga gagttgggga aactcctcct cttaggaaag gctaatacta	2190
gggtatcctt gggcccaatg aattaggggt gaggccccag aaccgttat ctatgagttg	2250
tatgggggag ccactcgaag ctgtagccac cagggatgca gctagctgag gagtttgggg	2310
tgttgggttg gacaaggcag gttagtagac tcagattctt gcttcaaaga gccttgggct	2370
ggcctggagg tccttgaggt ctagactgga cctaggagct tgagttgtca ggggccagga	2430
ctggccccac tgcagtgcc aggccagtct tgagcagcag ggagggetca gctgtcccca	2490
gatccaggtg cctctgacca gcctggtcac ctctgagga ataatgctg aacctacaa	2550
gccccatcat tcatttcttc tcaattcaca gtgccctct ttgtttctgg ggtggaacta	2610
ggctctgagg gcacagccta gctgagtgca aagaaatata ggatgcttag aaagcataca	2670
ggaggggcca ggcgtggtgg ctcatgctg taatoccaga actttgggat gccaaaggtg	2730
ttggattacc tgagatcagg tggattacct ggtctcgaga ccagcctgac caatatggtg	2790

```

aaaccccgctc tctactaaaa atacaaaaat taggctgaga caggagaatt gcttgaaccc 2850

aggaagcaga ggttgcaatg agctgagatt gcactactgc actccagcat gggcaacaaa 2910

gcaagactcc gtcacagaaa aaaaaaaaaa aaaaaa 2946

```

```

<210> 4
<211> 434
<212> PRT
<213> 人

<400> 4

```

```

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

```

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

<210> 5

<211> 5168

<212> DNA

<213> 人

<220>

<221> CDS

<222> (790).. (2823)

<223>

<400> 5

ctggagtctc cctggtcgg gagcctcagc cttctggaga ctgacaccac cctcttactc

60

agagacgaat cgttttgtcg ttgccttcac ctccctgacc acaagtgtc cggggctctt	120
tctaaggag gcagctgttc taggcgggag gctggagtct cctgggcctg gacgcacccc	180
gggtgtgag tgatgggtat gcctgaaagg aggggaagtg gcgcgcttt aatcatctgg	240
gctagtcccc ggcgggcctg ggggaacagg gaaactggag gcggccttaa agcgtcagga	300
tgagacatcg caagaggagc tgcagatact gacgctgcgc cccgggttct cgcgccttc	360
tctccgcga gcagccctc ggcaccctt tgccctaaa aatctgcaga ctgcgcctcc	420
tctccgggg agcgagacct agcaggccc gggctggcg tgccctcgcc tgccacgctg	480
cgcgctgcc tcagccgggc cgtggggcc gtgcagtga cgggcaacc cgcgccagc	540
tggggcagg caccgacct ccgtgggagg toccgaggca gcttcgctg ctgcacctg	600
ctccagccct cacctgcgc agccttagct gacgaccgc cgcactggg cgcccccgc	660
tcccacttc gccagcgcg gctcctcgcc tcggcccggg gtagtttga gggacgcgc	720
tctccacgtg cgcgactgc aggctggacc ctacgggctc ctggaaagga gacaccagca	780
tttgccaca atg ctg tca tcc act gac ttt aca ttt gct tcc tgg gag ctt	831
Met Leu Ser Ser Thr Asp Phe Thr Phe Ala Ser Trp Glu Leu	
1 5 10	
gtg gtc cgc gtt gac cat ccc aat gaa gag cag cag aaa gac gtc aca	879
Val Val Arg Val Asp His Pro Asn Glu Glu Gln Gln Lys Asp Val Thr	
15 20 25 30	
ctg aga gta tct gga gac ctt cat gtt gga gga gtg atg etc aag tta	927
Leu Arg Val Ser Gly Asp Leu His Val Gly Gly Val Met Leu Lys Leu	
35 40 45	
gta gaa cag atc aat ata tcc caa gac tgg tca gac ttt gct ctt tgg	975
Val Glu Gln Ile Asn Ile Ser Gln Asp Trp Ser Asp Phe Ala Leu Trp	
50 55 60	
tgg gaa cag aag cat tgc tgg ctt ctg aaa acc cac tgg acc ctg gac	1023

Trp Glu Gln Lys His Cys Trp Leu Leu Lys Thr His Trp Thr Leu Asp	
65	70
75	
aaa tat ggg gtc cag gca gat gca aag ctt ctc ttc acc cct cag cat	1071
Lys Tyr Gly Val Gln Ala Asp Ala Lys Leu Leu Phe Thr Pro Gln His	
80	85
90	
aaa atg ctg cgc ctt cgt ctg ccg aat ttg aag atg gtg agg ttg cga	1119
Lys Met Leu Arg Leu Arg Leu Pro Asn Leu Lys Met Val Arg Leu Arg	
95	100
105	110
gtc agc ttc tca gct gtg gtt ttt aaa gct gtc agt gat atc tgc aaa	1167
Val Ser Phe Ser Ala Val Val Phe Lys Ala Val Ser Asp Ile Cys Lys	
115	120
125	
atc ctg aat att aga aga tca gaa gag ctt tcc ttg tta aag ccg tct	1215
Ile Leu Asn Ile Arg Arg Ser Glu Glu Leu Ser Leu Leu Lys Pro Ser	
130	135
140	
ggg gac tat ttt aag aag aag aag aaa aaa gac aaa aat aat aag gaa	1263
Gly Asp Tyr Phe Lys Lys Lys Lys Lys Lys Asp Lys Asn Asn Lys Glu	
145	150
155	
ccc ata att gaa gat att cta aac ctg gag agt tct cca aca gct tca	1311
Pro Ile Ile Glu Asp Ile Leu Asn Leu Glu Ser Ser Pro Thr Ala Ser	
160	165
170	
ggg tca tca gta agt cct ggt tta tac agt aaa acc atg acc cct ata	1359
Gly Ser Ser Val Ser Pro Gly Leu Tyr Ser Lys Thr Met Thr Pro Ile	
175	180
185	190
tat gac ccc atc aat gga aca cca gca tca tcc acc atg act tgg ttc	1407
Tyr Asp Pro Ile Asn Gly Thr Pro Ala Ser Ser Thr Met Thr Trp Phe	
195	200
205	
agt gac agc cct ttg acg gaa caa aac tgc agc atc ctc gca ttc agc	1455
Ser Asp Ser Pro Leu Thr Glu Gln Asn Cys Ser Ile Leu Ala Phe Ser	
210	215
220	
caa ccc ccc cag tcc cca gaa gca ctt gcg gat atg tac cag cct egg	1503
Gln Pro Pro Gln Ser Pro Glu Ala Leu Ala Asp Met Tyr Gln Pro Arg	
225	230
235	

tct ctg gtt gat aaa gcc aag ctc aat gca ggt tgg cta gac tcc tca	1551
Ser Leu Val Asp Lys Ala Lys Leu Asn Ala Gly Trp Leu Asp Ser Ser	
240 245 250	
cgc tcc ctt atg gaa caa ggc atc caa gag gat gag cag ctg ctc tta	1599
Arg Ser Leu Met Glu Gln Gly Ile Gln Glu Asp Glu Gln Leu Leu Leu	
255 260 265 270	
cga ttt aaa tat tat tct ttc ttc gac ttg aat cct aaa tat gat gct	1647
Arg Phe Lys Tyr Tyr Ser Phe Phe Asp Leu Asn Pro Lys Tyr Asp Ala	
275 280 285	
gtc cga ata aac caa ctc tat gag caa gcc agg tgg gcc att ctc tta	1695
Val Arg Ile Asn Gln Leu Tyr Glu Gln Ala Arg Trp Ala Ile Leu Leu	
290 295 300	
gaa gaa att gat tgc aca gag gaa gaa atg ttg atc ttt gca gct cta	1743
Glu Glu Ile Asp Cys Thr Glu Glu Glu Met Leu Ile Phe Ala Ala Leu	
305 310 315	
cag tac cac att agc aaa ctg tcg ttg tct gct gaa aca cag gat ttt	1791
Gln Tyr His Ile Ser Lys Leu Ser Leu Ser Ala Glu Thr Gln Asp Phe	
320 325 330	
gca ggc gag tcc gag gtt gat gaa ata gaa gcg gcg ctt tct aat ttg	1839
Ala Gly Glu Ser Glu Val Asp Glu Ile Glu Ala Ala Leu Ser Asn Leu	
335 340 345 350	
gaa gta acc cta gaa ggt gga aaa gcg gac agc ctt ttg gag gac att	1887
Glu Val Thr Leu Glu Gly Gly Lys Ala Asp Ser Leu Leu Glu Asp Ile	
355 360 365	
act gat atc cct aaa ctt gca gat aat ctc aaa tta ttt agg ccc aag	1935
Thr Asp Ile Pro Lys Leu Ala Asp Asn Leu Lys Leu Phe Arg Pro Lys	
370 375 380	
aag tta cta cca aaa gct ttc aaa caa tat tgg ttt atc ttt aaa gac	1983
Lys Leu Leu Pro Lys Ala Phe Lys Gln Tyr Trp Phe Ile Phe Lys Asp	
385 390 395	
aca tcc ata gca tac ttt aaa aat aag gaa ctt gaa caa gga gaa cca	2031

Thr Ser Ile Ala Tyr Phe Lys Asn Lys Glu Leu Glu Gln Gly Glu Pro	
400	405 410
cta gaa aaa cta aat ctt aga ggc tgc gaa gtt gtg ccc gat gta aat	2079
Leu Glu Lys Leu Asn Leu Arg Gly Cys Glu Val Val Pro Asp Val Asn	
415	420 425 430
gta gca gga aga aaa ttt gga atc aag tta cta atc cct gtt gcc gat	2127
Val Ala Gly Arg Lys Phe Gly Ile Lys Leu Leu Ile Pro Val Ala Asp	
	435 440 445
ggt atg aat gaa atg tat ttg aga tgt gac cat gag aat caa tac gcc	2175
Gly Met Asn Glu Met Tyr Leu Arg Cys Asp His Glu Asn Gln Tyr Ala	
	450 455 460
caa tgg atg gct gcc tgc atg ttg gca tgc aag ggc aaa acc atg gca	2223
Gln Trp Met Ala Ala Cys Met Leu Ala Ser Lys Gly Lys Thr Met Ala	
	465 470 475
gac agc tcc tac cag cca gag gtc ctc aac atc ctt tca ttt ctg agg	2271
Asp Ser Ser Tyr Gln Pro Glu Val Leu Asn Ile Leu Ser Phe Leu Arg	
	480 485 490
atg aaa aac agg aac tct gca tct cag gtg gct tcc agt ctc gaa aac	2319
Met Lys Asn Arg Asn Ser Ala Ser Gln Val Ala Ser Ser Leu Glu Asn	
	495 500 505 510
atg gat atg aac cca gaa tgt ttt gtg tca cca cgg tgt gca aag aaa	2367
Met Asp Met Asn Pro Glu Cys Phe Val Ser Pro Arg Cys Ala Lys Lys	
	515 520 525
cac aaa tcc aaa cag ctg gcc gcc cgg atc ctg gag gcg cac cag aac	2415
His Lys Ser Lys Gln Leu Ala Ala Arg Ile Leu Glu Ala His Gln Asn	
	530 535 540
gtg gcc cag atg ccc ctg gtc gaa gcc aag ctg cgg ttc atc cag gcg	2463
Val Ala Gln Met Pro Leu Val Glu Ala Lys Leu Arg Phe Ile Gln Ala	
	545 550 555
tgg cag tca ctg cct gag ttt ggc ctc acc tac tac ctt gtc aga ttt	2511
Trp Gln Ser Leu Pro Glu Phe Gly Leu Thr Tyr Tyr Leu Val Arg Phe	
	560 565 570

aaa gga agc aaa aaa gat gac att ctg gga gtt tca tat aac agg ttg Lys Gly Ser Lys Lys Asp Asp Ile Leu Gly Val Ser Tyr Asn Arg Leu 575 580 585 590	2559
att aaa att gat gca gcc acc ggg att cca gtg aca aca tgg aga ttc Ile Lys Ile Asp Ala Ala Thr Gly Ile Pro Val Thr Thr Trp Arg Phe 595 600 605	2607
aca aat atc aaa cag tgg aat gta aac tgg gaa acc cgg cag gtg gtc Thr Asn Ile Lys Gln Trp Asn Val Asn Trp Glu Thr Arg Gln Val Val 610 615 620	2655
atc gag ttt gac caa aac gtc ttt act gct ttc acc tgc ctg agt gca Ile Glu Phe Asp Gln Asn Val Phe Thr Ala Phe Thr Cys Leu Ser Ala 625 630 635	2703
gat tgc aag att gtg cac gag tac att ggc ggc tac att ttc ttg tcc Asp Cys Lys Ile Val His Glu Tyr Ile Gly Gly Tyr Ile Phe Leu Ser 640 645 650	2751
acc cgc tcc aag gac cag aat gaa aca ctc gat gag gac ttg ttc cac Thr Arg Ser Lys Asp Gln Asn Glu Thr Leu Asp Glu Asp Leu Phe His 655 660 665 670	2799
aaa ttg acc ggc ggt cag gat tga aacaagcacg cgtgctcggc tcaccaaac Lys Leu Thr Gly Gly Gln Asp 675	2853
aaggcaagcc aaaggcgccc ctccccagag ggatccctaa cgtgcccagc atgtagattc	2913
tggactaaca gacaacatac attcaccgct ggtaaccagc atcctcattc aaacccactg	2973
ctggcacatc cctttcetta ctttgccctg tgetaccagc cacggaagga gcctctcttg	3033
tttttctat aaaatgggta ggcaggagaa aagcaggtgc cctaagattg ctctaagccc	3093
cageatgtgg ttacagttct ctgacttgca gaacctgcca ggtgtatggc tacaagttat	3153
cctcgtgctg atctgtctca ttactaagtc aatggagaag acagaaaggt aaaaatcacg	3213
tgtagcaaga acaactctta tttcacaaac tcaggtatga aacgaaacgc ctgtccttca	3273

tggaactgct tttagctcct gtcttttcaa aatggcagag ggagttccta cacacacttt	3333
ttccctggag gccaaggtct aggggtagaa aggggagggg tggggctacc aggtagcagt	3393
tgacaacca aggtcagagg agtggcctc agtgtcatct gtccacagtg atacctgcca	3453
agatgaccac tgaccacat ctggtcttag tcattggtct cctcagattt ctggggccac	3513
ctgcaagccc cattccattc ctacagatct ctcagccacc tgtaagtctt ttgtgaagat	3573
gtgggtgaca cagggggaca ggaaaacca tttctcaacc cagatccatg tctccactgc	3633
ttctactctg ggttgggatt caggaagaca ggcacagtcc tctctgttca tagaaacacc	3693
tgccagtgtc aaggattcca gtcaggtgtc tatcccaact ggtcagggag agaagggcag	3753
accattctc aaagaccacc atgtccaagg tctgacagct ccccactggc tgcecccaca	3813
ggggctttag gctggtctgg gtcatgggga agcgtccctc ttatcgttg tctgtgttct	3873
cctggatttg gtatctatgt tggtacgact cctggccttt tatctaaagg actttggctt	3933
ttgtaaata caagccaata atagactttt ttctcccct ctgttttttg ctgtgtcatc	3993
tctgccttga gactgccttg agacagtgtc tgccttgaga gaggtagcca attaacagct	4053
gcctgaattg tcattttcca ttttggtttg ttagaggtgg gaggggtggg ttttgagaag	4113
gtcaaaagca ataccagaag taaagggaaa taccagacaa tattttatta tttttcata	4173
gatgttctgc cacacaaaga acttgggggtg taaggataag gcaaaagctc caatccatt	4233
tttcagttct cctagatgc acccctcagg gagectggcc agagttccga ggctcgtgag	4293
cgtcagctgt tgctttattt tccatcaaag ccctctgaga agtgagacct cagcaattcc	4353
gggagccaca tagagacaga cttggcaagg gaccccctgg ttctgagcca gtagctgcca	4413
tctggaaatt cctcttttag cctctcctta gaggtgaatg tgaatgaagc ctcccaggca	4473
cccgtgaat ttctgaggcc ttgcttaaag ctcagaagtg gtttaggcat ttggaaaatc	4533

tggttcacat cataaagaac ttgatttgaa atgttttcta tagaaacaag tgctaagtgt 4593
 accgtattat acttgatggt ggtcatttct cagtcctatt tctcagttct attatttttag 4653
 aacctagtca gttctttaag attataactg gtectacatt aaaataatgc ttctcgatgt 4713
 cagattttac ctgtttgctg ctgagaacat ctctgcctaa tttaccaaaag ccagaccttc 4773
 agttcaacat gcttccttag cttttcatag ttgtctgaca tttccatgaa aacaaaggaa 4833
 ccaactttgt ttaaccaaaa ctttgtttgg ttacagtttt caggggagcg tttcttccat 4893
 gacacacagc aacatcccaa agaaataaac aagtgtgaca aaaaaaaaaa aaaacaaacc 4953
 taaatgctac tgttccaaag agcaacttga tggttttttt taatactgag tgcaaaaggt 5013
 cacccaaatt cctatgatga aattttaaat taatgggcac ctttcaacat catttgettc 5073
 cttatctaca gttgattcag aaatctgcat tttttattct tttatatgac ttttaagtaa 5133
 aagatttata tggatttgaa aaaaaaaaaa aaaaa 5168

<210> 6

<211> 677

<212> PRT

<213> 人

<400> 6

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

85	90	95
Leu Arg Leu Arg Leu Pro Asn Leu Lys Met Val Arg Leu Arg Val Ser		
100	105	110
Phe Ser Ala Val Val Phe Lys Ala Val Ser Asp Ile Cys Lys Ile Leu		
115	120	125
Asn Ile Arg Arg Ser Glu Glu Leu Ser Leu Leu Lys Pro Ser Gly Asp		
130	135	140
Tyr Phe Lys Lys Lys Lys Lys Lys Asp Lys Asn Asn Lys Glu Pro Ile		
145	150	155
Ile Glu Asp Ile Leu Asn Leu Glu Ser Ser Pro Thr Ala Ser Gly Ser		
165	170	175
Ser Val Ser Pro Gly Leu Tyr Ser Lys Thr Met Thr Pro Ile Tyr Asp		
180	185	190
Pro Ile Asn Gly Thr Pro Ala Ser Ser Thr Met Thr Trp Phe Ser Asp		
195	200	205
Ser Pro Leu Thr Glu Gln Asn Cys Ser Ile Leu Ala Phe Ser Gln Pro		
210	215	220
Pro Gln Ser Pro Glu Ala Leu Ala Asp Met Tyr Gln Pro Arg Ser Leu		
225	230	235
Val Asp Lys Ala Lys Leu Asn Ala Gly Trp Leu Asp Ser Ser Arg Ser		
245	250	255
Leu Met Glu Gln Gly Ile Gln Glu Asp Glu Gln Leu Leu Leu Arg Phe		
260	265	270
Lys Tyr Tyr Ser Phe Phe Asp Leu Asn Pro Lys Tyr Asp Ala Val Arg		
275	280	285
Ile Asn Gln Leu Tyr Glu Gln Ala Arg Trp Ala Ile Leu Leu Glu Glu		
290	295	300
Ile Asp Cys Thr Glu Glu Glu Met Leu Ile Phe Ala Ala Leu Gln Tyr		
305	310	315
His Ile Ser Lys Leu Ser Leu Ser Ala Glu Thr Gln Asp Phe Ala Gly		
325	330	335
Glu Ser Glu Val Asp Glu Ile Glu Ala Ala Leu Ser Asn Leu Glu Val		
340	345	350
Thr Leu Glu Gly Gly Lys Ala Asp Ser Leu Leu Glu Asp Ile Thr Asp		
355	360	365
Ile Pro Lys Leu Ala Asp Asn Leu Lys Leu Phe Arg Pro Lys Lys Leu		
370	375	380
Leu Pro Lys Ala Phe Lys Gln Tyr Trp Phe Ile Phe Lys Asp Thr Ser		
385	390	395
Ile Ala Tyr Phe Lys Asn Lys Glu Leu Glu Gln Gly Glu Pro Leu Glu		
405	410	415
Lys Leu Asn Leu Arg Gly Cys Glu Val Val Pro Asp Val Asn Val Ala		

	420	425	430
Gly Arg Lys Phe Gly Ile Lys Leu Leu Ile Pro Val Ala Asp Gly Met			
	435	440	445
Asn Glu Met Tyr Leu Arg Cys Asp His Glu Asn Gln Tyr Ala Gln Trp			
	450	455	460
Met Ala Ala Cys Met Leu Ala Ser Lys Gly Lys Thr Met Ala Asp Ser			
465	470	475	480
Ser Tyr Gln Pro Glu Val Leu Asn Ile Leu Ser Phe Leu Arg Met Lys			
	485	490	495
Asn Arg Asn Ser Ala Ser Gln Val Ala Ser Ser Leu Glu Asn Met Asp			
	500	505	510
Met Asn Pro Glu Cys Phe Val Ser Pro Arg Cys Ala Lys Lys His Lys			
	515	520	525
Ser Lys Gln Leu Ala Ala Arg Ile Leu Glu Ala His Gln Asn Val Ala			
530	535	540	
Gln Met Pro Leu Val Glu Ala Lys Leu Arg Phe Ile Gln Ala Trp Gln			
545	550	555	560
Ser Leu Pro Glu Phe Gly Leu Thr Tyr Tyr Leu Val Arg Phe Lys Gly			
	565	570	575
Ser Lys Lys Asp Asp Ile Leu Gly Val Ser Tyr Asn Arg Leu Ile Lys			
	580	585	590
Ile Asp Ala Ala Thr Gly Ile Pro Val Thr Thr Trp Arg Phe Thr Asn			
	595	600	605
Ile Lys Gln Trp Asn Val Asn Trp Glu Thr Arg Gln Val Val Ile Glu			
	610	615	620
Phe Asp Gln Asn Val Phe Thr Ala Phe Thr Cys Leu Ser Ala Asp Cys			
625	630	635	640
Lys Ile Val His Glu Tyr Ile Gly Gly Tyr Ile Phe Leu Ser Thr Arg			
	645	650	655
Ser Lys Asp Gln Asn Glu Thr Leu Asp Glu Asp Leu Phe His Lys Leu			
	660	665	670
Thr Gly Gly Gln Asp			
	675		

<210> 7

<211> 3486

<212> DNA

<213> 人

<220>

<221> CDS

<222> (93)..(3350)	
<223>	
<400> 7	
gagccgactg cggctgcgcg gggatttcga atcggcggcg gcttctagtt tgcggttcag	60
gtttggecgc tgccggccag cgtcctctgg cc atg gac acc ccg gaa aat gtc	113
Met Asp Thr Pro Glu Asn Val	
1 5	
ctt cag atg ctt gaa gcc cac atg cag agc tac aag ggc aat gac cct	161
Leu Gln Met Leu Glu Ala His Met Gln Ser Tyr Lys Gly Asn Asp Pro	
10 15 20	
ctt ggt gaa tgg gaa aga tac ata cag tgg gta gaa gag aat ttt cct	209
Leu Gly Glu Trp Glu Arg Tyr Ile Gln Trp Val Glu Glu Asn Phe Pro	
25 30 35	
gag aat aaa gaa tac ttg ata act tta cta gaa cat tta atg aag gaa	257
Glu Asn Lys Glu Tyr Leu Ile Thr Leu Leu Glu His Leu Met Lys Glu	
40 45 50 55	
ttt tta gat aag aag aaa tac cac aat gac cca aga ttc atc agt tat	305
Phe Leu Asp Lys Lys Lys Tyr His Asn Asp Pro Arg Phe Ile Ser Tyr	
60 65 70	
tgt tta aaa ttt gct gag tac aac agt gac ctc cat caa ttt ttt gag	353
Cys Leu Lys Phe Ala Glu Tyr Asn Ser Asp Leu His Gln Phe Phe Glu	
75 80 85	
ttt ctg tac aac cat ggg att gga acc ctg tca tcc cct ctg tac att	401
Phe Leu Tyr Asn His Gly Ile Gly Thr Leu Ser Ser Pro Leu Tyr Ile	
90 95 100	
gcc tgg gcg ggg cat ctg gaa gcc caa gga gag ctg cag cat gcc agt	449
Ala Trp Ala Gly His Leu Glu Ala Gln Gly Glu Leu Gln His Ala Ser	
105 110 115	
gct gtc ctt cag aga gga att caa aac cag gct gaa ccc aga gag ttc	497
Ala Val Leu Gln Arg Gly Ile Gln Asn Gln Ala Glu Pro Arg Glu Phe	
120 125 130 135	

ctg caa caa caa tac agg tta ttt cag aca cgc ctc act gaa acc cat	545
Leu Gln Gln Gln Tyr Arg Leu Phe Gln Thr Arg Leu Thr Glu Thr His	
140 145 150	
ttg cca gct caa gct aga acc tca gaa cct ctg cat aat gtt cag gtt	593
Leu Pro Ala Gln Ala Arg Thr Ser Glu Pro Leu His Asn Val Gln Val	
155 160 165	
tta aat caa atg ata aca tca aaa tca aat cca gga aat aac atg gcc	641
Leu Asn Gln Met Ile Thr Ser Lys Ser Asn Pro Gly Asn Asn Met Ala	
170 175 180	
tgc att tct aag aat cag ggt tca gag ctt tct gga gtg ata tct tca	689
Cys Ile Ser Lys Asn Gln Gly Ser Glu Leu Ser Gly Val Ile Ser Ser	
185 190 195	
gct tgt gat aaa gag tca aat atg gaa cga aga gtg atc acg att tct	737
Ala Cys Asp Lys Glu Ser Asn Met Glu Arg Arg Val Ile Thr Ile Ser	
200 205 210 215	
aaa tca gaa tat tct gtg cac tca tct ttg gca tcc aaa gtt gat gtt	785
Lys Ser Glu Tyr Ser Val His Ser Ser Leu Ala Ser Lys Val Asp Val	
220 225 230	
gag cag gtt gtt atg tat tgc aag gag aag ctt att cgt ggg gaa tca	833
Glu Gln Val Val Met Tyr Cys Lys Glu Lys Leu Ile Arg Gly Glu Ser	
235 240 245	
gaa ttt tcc ttt gaa gaa ttg aga gcc cag aaa tac aat caa cgg aga	881
Glu Phe Ser Phe Glu Glu Leu Arg Ala Gln Lys Tyr Asn Gln Arg Arg	
250 255 260	
aag cat gag caa tgg gta aat gaa gac aga cat tat atg aaa agg aaa	929
Lys His Glu Gln Trp Val Asn Glu Asp Arg His Tyr Met Lys Arg Lys	
265 270 275	
gaa gca aat gct ttt gaa gaa cag cta tta aaa cag aaa atg gat gaa	977
Glu Ala Asn Ala Phe Glu Glu Gln Leu Leu Lys Gln Lys Met Asp Glu	
280 285 290 295	
ctt cat aag aag ttg cat cag gtg gtg gag aca tcc cat gag gat ctg	1025
Leu His Lys Lys Leu His Gln Val Val Glu Thr Ser His Glu Asp Leu	

	300	305	310	
ccc gct tcc cag gaa agg tcc gag gtt aat cca gca cgt atg ggg cca				1073
Pro Ala Ser Gln Glu Arg Ser Glu Val Asn Pro Ala Arg Met Gly Pro				
	315	320	325	
agt gta ggc tcc cag cag gaa ctg aga gcg cca tgt ctt cca gta acc				1121
Ser Val Gly Ser Gln Gln Glu Leu Arg Ala Pro Cys Leu Pro Val Thr				
	330	335	340	
tat cag cag aca cca gtg aac atg gaa aag aac cca aga gag gca cct				1169
Tyr Gln Gln Thr Pro Val Asn Met Glu Lys Asn Pro Arg Glu Ala Pro				
	345	350	355	
cct gtt gtt cct cct ttg gca aat gct att tct gca gct ttg gtg tcc				1217
Pro Val Val Pro Pro Leu Ala Asn Ala Ile Ser Ala Ala Leu Val Ser				
	360	365	370	375
cca gcc acc agc cag agc att gct cct cct gtt cct ttg aaa gcc cag				1265
Pro Ala Thr Ser Gln Ser Ile Ala Pro Pro Val Pro Leu Lys Ala Gln				
	380	385	390	
aca gta aca gac tcc atg ttt gca gtg gcc agc aaa gat gct gga tgt				1313
Thr Val Thr Asp Ser Met Phe Ala Val Ala Ser Lys Asp Ala Gly Cys				
	395	400	405	
gtg aat aag agt act cat gaa ttc aag cca cag agt gga gca gag atc				1361
Val Asn Lys Ser Thr His Glu Phe Lys Pro Gln Ser Gly Ala Glu Ile				
	410	415	420	
aaa gaa ggg tgt gaa aca cat aag gtt gcc aac aca agt tct ttt cac				1409
Lys Glu Gly Cys Glu Thr His Lys Val Ala Asn Thr Ser Ser Phe His				
	425	430	435	
aca act cca aac aca tca ctg gga atg gtt cag gca acg cca tcc aaa				1457
Thr Thr Pro Asn Thr Ser Leu Gly Met Val Gln Ala Thr Pro Ser Lys				
	440	445	450	455
gtg cag cca tca ccc acc gtg cac aca aaa gaa gca tta ggt ttc atc				1505
Val Gln Pro Ser Pro Thr Val His Thr Lys Glu Ala Leu Gly Phe Ile				
	460	465	470	

atg aat atg ttt cag gct cct aca ctt cct gat att tct gat gac aaa Met Asn Met Phe Gln Ala Pro Thr Leu Pro Asp Ile Ser Asp Asp Lys 475 480 485	1553
gat gaa tgg caa tct cta gat caa aat gaa gat gca ttt gaa gcc cag Asp Glu Trp Gln Ser Leu Asp Gln Asn Glu Asp Ala Phe Glu Ala Gln 490 495 500	1601
ttt caa aaa aat gta agg tca tct ggg gct tgg gga gtc aat aag atc Phe Gln Lys Asn Val Arg Ser Ser Gly Ala Trp Gly Val Asn Lys Ile 505 510 515	1649
atc tct tct ttg tca tct gct ttt cat gtg ttt gaa gat gga aac aaa Ile Ser Ser Leu Ser Ser Ala Phe His Val Phe Glu Asp Gly Asn Lys 520 525 530 535	1697
gaa aat tat gga tta cca cag cct aaa aat aaa ccc aca gga gcc agg Glu Asn Tyr Gly Leu Pro Gln Pro Lys Asn Lys Pro Thr Gly Ala Arg 540 545 550	1745
acc ttt gga gaa cgc tct gtc agc aga ctt cct tca aaa cca aag gag Thr Phe Gly Glu Arg Ser Val Ser Arg Leu Pro Ser Lys Pro Lys Glu 555 560 565	1793
gaa gtg cct cat gct gaa gag ttt ttg gat gac tca act gta tgg ggt Glu Val Pro His Ala Glu Glu Phe Leu Asp Asp Ser Thr Val Trp Gly 570 575 580	1841
att cgc tgc aac aaa acc ctg gca ccc agt cct aag agc cca gga gac Ile Arg Cys Asn Lys Thr Leu Ala Pro Ser Pro Lys Ser Pro Gly Asp 585 590 595	1889
ttc aca tct gct gca caa ctt gcg tct aca cca ttc cac aag ctt cca Phe Thr Ser Ala Ala Gln Leu Ala Ser Thr Pro Phe His Lys Leu Pro 600 605 610 615	1937
gtg gag tca gtg cac att tta gaa gat aaa gaa aat gtg gta gca aaa Val Glu Ser Val His Ile Leu Glu Asp Lys Glu Asn Val Val Ala Lys 620 625 630	1985
cag tgt acc cag gog act ttg gat tct tgt gag gaa aac atg gtg gtg Gln Cys Thr Gln Ala Thr Leu Asp Ser Cys Glu Glu Asn Met Val Val	2033

635	640	645	
cct tca agg gat gga aaa ttc agt cca att caa gag aaa agc cca aaa			2081
Pro Ser Arg Asp Gly Lys Phe Ser Pro Ile Gln Glu Lys Ser Pro Lys			
650	655	660	
cag gcc ttg teg tct cac atg tat tca gca tcc tta ctt cgt ctg agc			2129
Gln Ala Leu Ser Ser His Met Tyr Ser Ala Ser Leu Leu Arg Leu Ser			
665	670	675	
cag cct gct gca ggt ggg gta ctt acc tgt gag gca gag ttg ggc gtt			2177
Gln Pro Ala Ala Gly Gly Val Leu Thr Cys Glu Ala Glu Leu Gly Val			
680	685	690	695
gag gct tgc aga ctc aca gac act gac gct gcc att gca gaa gat cca			2225
Glu Ala Cys Arg Leu Thr Asp Thr Asp Ala Ala Ile Ala Glu Asp Pro			
700	705	710	
cca gat gct att gct ggg ctc caa gca gaa tgg atg cag atg agt tca			2273
Pro Asp Ala Ile Ala Gly Leu Gln Ala Glu Trp Met Gln Met Ser Ser			
715	720	725	
ctt ggg act gtt gat gct cca aac ttc att gtt ggg aac cca tgg gat			2321
Leu Gly Thr Val Asp Ala Pro Asn Phe Ile Val Gly Asn Pro Trp Asp			
730	735	740	
gat aag ctg att ttc aaa ctt tta tct ggg ctt tct aaa cca gtg agt			2369
Asp Lys Leu Ile Phe Lys Leu Leu Ser Gly Leu Ser Lys Pro Val Ser			
745	750	755	
tcc tat cca aat act ttt gaa tgg caa tgt aaa ctt cca gcc atc aag			2417
Ser Tyr Pro Asn Thr Phe Glu Trp Gln Cys Lys Leu Pro Ala Ile Lys			
760	765	770	775
ccc aag act gaa ttt caa ttg ggt tct aag ctg gtc tat gtc cat cac			2465
Pro Lys Thr Glu Phe Gln Leu Gly Ser Lys Leu Val Tyr Val His His			
780	785	790	
ctt ctt gga gaa gga gcc ttt gcc cag gtg tac gaa gct acc cag gga			2513
Leu Leu Gly Glu Gly Ala Phe Ala Gln Val Tyr Glu Ala Thr Gln Gly			
795	800	805	

gat ctg aat gat gct aaa aat aaa cag aaa ttt gtt tta aag gtc caa	2561
Asp Leu Asn Asp Ala Lys Asn Lys Gln Lys Phe Val Leu Lys Val Gln	
810 815 820	
aag cct gcc aac ccc tgg gaa ttc tac att ggg acc cag ttg atg gaa	2609
Lys Pro Ala Asn Pro Trp Glu Phe Tyr Ile Gly Thr Gln Leu Met Glu	
825 830 835	
aga cta aag cca tct atg cag cac atg ttt atg aag ttc tat tct gcc	2657
Arg Leu Lys Pro Ser Met Gln His Met Phe Met Lys Phe Tyr Ser Ala	
840 845 850 855	
cac tta ttc cag aat ggc agt gta tta gta gga gag ctc tac agc tat	2705
His Leu Phe Gln Asn Gly Ser Val Leu Val Gly Glu Leu Tyr Ser Tyr	
860 865 870	
gga aca tta tta aat gcc att aac ctc tat aaa aat acc cct gaa aaa	2753
Gly Thr Leu Leu Asn Ala Ile Asn Leu Tyr Lys Asn Thr Pro Glu Lys	
875 880 885	
gtg atg cct caa ggt ctt gtc atc tct ttt gct atg aga atg ctt tac	2801
Val Met Pro Gln Gly Leu Val Ile Ser Phe Ala Met Arg Met Leu Tyr	
890 895 900	
atg att gag caa gtg cat gac tgt gaa atc att cat gga gac att aaa	2849
Met Ile Glu Gln Val His Asp Cys Glu Ile Ile His Gly Asp Ile Lys	
905 910 915	
cca gac aat ttc ata ctt gga aac gga ttt ttg gaa cag gat gat gaa	2897
Pro Asp Asn Phe Ile Leu Gly Asn Gly Phe Leu Glu Gln Asp Asp Glu	
920 925 930 935	
gat gat tta tct gct ggc ttg gca ctg att gac ctg ggt cag agt ata	2945
Asp Asp Leu Ser Ala Gly Leu Ala Leu Ile Asp Leu Gly Gln Ser Ile	
940 945 950	
gat atg aaa ctt ttt cca aaa gga act ata ttc aca gca aag tgt gaa	2993
Asp Met Lys Leu Phe Pro Lys Gly Thr Ile Phe Thr Ala Lys Cys Glu	
955 960 965	
aca tct ggt ttt cag tgt gtt gag atg ctc agc aac aaa cca tgg aac	3041
Thr Ser Gly Phe Gln Cys Val Glu Met Leu Ser Asn Lys Pro Trp Asn	

970	975	980	
tac cag atc gat tac ttt ggg gtt gct gca aca gta tat tgc atg ctc			3089
Tyr Gln Ile Asp Tyr Phe Gly Val Ala Ala Thr Val Tyr Cys Met Leu			
985	990	995	
ttt ggc act tac atg aaa gtg aaa aat gaa gga gga gag tgt aag			3134
Phe Gly Thr Tyr Met Lys Val Lys Asn Glu Gly Gly Glu Cys Lys			
1000	1005	1010	
cct gaa ggt ctt ttt aga agg ctt cct cat ttg gat atg tgg aat			3179
Pro Glu Gly Leu Phe Arg Arg Leu Pro His Leu Asp Met Trp Asn			
1015	1020	1025	
gaa ttt ttt cat gtt atg ttg aat att cca gat tgt cat cat ctt			3224
Glu Phe Phe His Val Met Leu Asn Ile Pro Asp Cys His His Leu			
1030	1035	1040	
cca tct ttg gat ttg tta agg caa aag ctg aag aaa gta ttt caa			3269
Pro Ser Leu Asp Leu Leu Arg Gln Lys Leu Lys Lys Val Phe Gln			
1045	1050	1055	
caa cac tat act aac aag att agg gcc cta cgt aat agg cta att			3314
Gln His Tyr Thr Asn Lys Ile Arg Ala Leu Arg Asn Arg Leu Ile			
1060	1065	1070	
gta ctg ctc tta gaa tgt aag cgt tca cga aaa taa aatttgata			3360
Val Leu Leu Leu Glu Cys Lys Arg Ser Arg Lys			
1075	1080	1085	
tagacagtcc ttaaaaatca cactgtaa atgaatctgc tcaactttaa cctgtttttt			3420
tttcatttat tgtttatgta aatgtttgtt aaaaataaat cccatggaat atttccatgt			3480
aaaaaa			3486
<210> 8			
<211> 1085			
<212> PRT			
<213> 人			

<400> 8

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

Asn Pro Ala Arg Met Gly Pro Ser Val Gly Ser Gln Gln Glu Leu Arg
 325 330 335
 Ala Pro Cys Leu Pro Val Thr Tyr Gln Gln Thr Pro Val Asn Met Glu
 340 345 350
 Lys Asn Pro Arg Glu Ala Pro Pro Val Val Pro Pro Leu Ala Asn Ala
 355 360 365
 Ile Ser Ala Ala Leu Val Ser Pro Ala Thr Ser Gln Ser Ile Ala Pro
 370 375 380
 Pro Val Pro Leu Lys Ala Gln Thr Val Thr Asp Ser Met Phe Ala Val
 385 390 395 400
 Ala Ser Lys Asp Ala Gly Cys Val Asn Lys Ser Thr His Glu Phe Lys
 405 410 415
 Pro Gln Ser Gly Ala Glu Ile Lys Glu Gly Cys Glu Thr His Lys Val
 420 425 430
 Ala Asn Thr Ser Ser Phe His Thr Thr Pro Asn Thr Ser Leu Gly Met
 435 440 445
 Val Gln Ala Thr Pro Ser Lys Val Gln Pro Ser Pro Thr Val His Thr
 450 455 460
 Lys Glu Ala Leu Gly Phe Ile Met Asn Met Phe Gln Ala Pro Thr Leu
 465 470 475 480
 Pro Asp Ile Ser Asp Asp Lys Asp Glu Trp Gln Ser Leu Asp Gln Asn
 485 490 495
 Glu Asp Ala Phe Glu Ala Gln Phe Gln Lys Asn Val Arg Ser Ser Gly
 500 505 510
 Ala Trp Gly Val Asn Lys Ile Ile Ser Ser Leu Ser Ser Ala Phe His
 515 520 525
 Val Phe Glu Asp Gly Asn Lys Glu Asn Tyr Gly Leu Pro Gln Pro Lys
 530 535 540
 Asn Lys Pro Thr Gly Ala Arg Thr Phe Gly Glu Arg Ser Val Ser Arg
 545 550 555 560
 Leu Pro Ser Lys Pro Lys Glu Glu Val Pro His Ala Glu Glu Phe Leu
 565 570 575
 Asp Asp Ser Thr Val Trp Gly Ile Arg Cys Asn Lys Thr Leu Ala Pro
 580 585 590
 Ser Pro Lys Ser Pro Gly Asp Phe Thr Ser Ala Ala Gln Leu Ala Ser
 595 600 605
 Thr Pro Phe His Lys Leu Pro Val Glu Ser Val His Ile Leu Glu Asp
 610 615 620
 Lys Glu Asn Val Val Ala Lys Gln Cys Thr Gln Ala Thr Leu Asp Ser
 625 630 635 640
 Cys Glu Glu Asn Met Val Val Pro Ser Arg Asp Gly Lys Phe Ser Pro
 645 650 655

Ile Gln Glu Lys Ser Pro Lys Gln Ala Leu Ser Ser His Met Tyr Ser
 660 665 670
 Ala Ser Leu Leu Arg Leu Ser Gln Pro Ala Ala Gly Gly Val Leu Thr
 675 680 685
 Cys Glu Ala Glu Leu Gly Val Glu Ala Cys Arg Leu Thr Asp Thr Asp
 690 695 700
 Ala Ala Ile Ala Glu Asp Pro Pro Asp Ala Ile Ala Gly Leu Gln Ala
 705 710 715 720
 Glu Trp Met Gln Met Ser Ser Leu Gly Thr Val Asp Ala Pro Asn Phe
 725 730 735
 Ile Val Gly Asn Pro Trp Asp Asp Lys Leu Ile Phe Lys Leu Leu Ser
 740 745 750
 Gly Leu Ser Lys Pro Val Ser Ser Tyr Pro Asn Thr Phe Glu Trp Gln
 755 760 765
 Cys Lys Leu Pro Ala Ile Lys Pro Lys Thr Glu Phe Gln Leu Gly Ser
 770 775 780
 Lys Leu Val Tyr Val His His Leu Leu Gly Glu Gly Ala Phe Ala Gln
 785 790 795 800
 Val Tyr Glu Ala Thr Gln Gly Asp Leu Asn Asp Ala Lys Asn Lys Gln
 805 810 815
 Lys Phe Val Leu Lys Val Gln Lys Pro Ala Asn Pro Trp Glu Phe Tyr
 820 825 830
 Ile Gly Thr Gln Leu Met Glu Arg Leu Lys Pro Ser Met Gln His Met
 835 840 845
 Phe Met Lys Phe Tyr Ser Ala His Leu Phe Gln Asn Gly Ser Val Leu
 850 855 860
 Val Gly Glu Leu Tyr Ser Tyr Gly Thr Leu Leu Asn Ala Ile Asn Leu
 865 870 875 880
 Tyr Lys Asn Thr Pro Glu Lys Val Met Pro Gln Gly Leu Val Ile Ser
 885 890 895
 Phe Ala Met Arg Met Leu Tyr Met Ile Glu Gln Val His Asp Cys Glu
 900 905 910
 Ile Ile His Gly Asp Ile Lys Pro Asp Asn Phe Ile Leu Gly Asn Gly
 915 920 925
 Phe Leu Glu Gln Asp Asp Glu Asp Asp Leu Ser Ala Gly Leu Ala Leu
 930 935 940
 Ile Asp Leu Gly Gln Ser Ile Asp Met Lys Leu Phe Pro Lys Gly Thr
 945 950 955 960
 Ile Phe Thr Ala Lys Cys Glu Thr Ser Gly Phe Gln Cys Val Glu Met
 965 970 975
 Leu Ser Asn Lys Pro Trp Asn Tyr Gln Ile Asp Tyr Phe Gly Val Ala
 980 985 990

Ala Thr Val Tyr Cys Met Leu Phe Gly Thr Tyr Met Lys Val Lys Asn
 995 1000 1005
 Glu Gly Gly Glu Cys Lys Pro Glu Gly Leu Phe Arg Arg Leu Pro
 1010 1015 1020
 His Leu Asp Met Trp Asn Glu Phe Phe His Val Met Leu Asn Ile
 1025 1030 1035
 Pro Asp Cys His His Leu Pro Ser Leu Asp Leu Leu Arg Gln Lys
 1040 1045 1050
 Leu Lys Lys Val Phe Gln Gln His Tyr Thr Asn Lys Ile Arg Ala
 1055 1060 1065
 Leu Arg Asn Arg Leu Ile Val Leu Leu Leu Glu Cys Lys Arg Ser
 1070 1075 1080
 Arg Lys
 1085

<210> 9

<211> 2635

<212> DNA

<213> 人

<220>

<221> CDS

<222> (285)..(1679)

<223>

<400> 9

ggccgacgcg agcgccgcgc ttgccttcag ctgctagctg gccaagggga ggcgaccgcg 60
 gagggtggcg aggggcggcc aggaccgcga gccccggggc cgggcccggtc cggaccgcca 120
 gggagggcag gtcagtgggc agatcgcgtc cgcgggattc aatctctgcc cgctctgata 180
 acagtccttt tccttggcgc tcaacttcgtg cctggcaccg ggctggggcg ctaagaccg 240
 ttgtctcttc gatcgttctt ttggacttgg cgaccatttc agag atg tct tcc aga 296
 Met Ser Ser Arg
 1
 agt acc aaa gat tta att aaa agt aag tgg gga tcg aag cct agt aac 344
 Ser Thr Lys Asp Leu Ile Lys Ser Lys Trp Gly Ser Lys Pro Ser Asn
 5 10 15 20

tcc aaa tcc gaa act aca tta gaa aaa tta aag gga gaa att gca cac	392
Ser Lys Ser Glu Thr Thr Leu Glu Lys Leu Lys Gly Glu Ile Ala His	
25 30 35	
tta aag aca tca gtg gat gaa atc aca agt ggg aaa gga aag ctg act	440
Leu Lys Thr Ser Val Asp Glu Ile Thr Ser Gly Lys Gly Lys Leu Thr	
40 45 50	
gat aaa gag aga cac aga ctt ttg gag aaa att cga gtc ctt gag gct	488
Asp Lys Glu Arg His Arg Leu Leu Glu Lys Ile Arg Val Leu Glu Ala	
55 60 65	
gag aag gag aag aat gct tat caa ctc aca gag aag gac aaa gaa ata	536
Glu Lys Glu Lys Asn Ala Tyr Gln Leu Thr Glu Lys Asp Lys Glu Ile	
70 75 80	
cag cga ctg aga gac caa ctg aag gcc aga tat agt act acc gca ttg	584
Gln Arg Leu Arg Asp Gln Leu Lys Ala Arg Tyr Ser Thr Thr Ala Leu	
85 90 95 100	
ctt gaa cag ctg gaa gag aca acg aga gaa gga gaa agg agg gag cag	632
Leu Glu Gln Leu Glu Glu Thr Thr Arg Glu Gly Glu Arg Arg Glu Gln	
105 110 115	
gtg ttg aaa gcc tta tct gaa gag aaa gac gta ttg aaa caa cag ttg	680
Val Leu Lys Ala Leu Ser Glu Glu Lys Asp Val Leu Lys Gln Gln Leu	
120 125 130	
tct gct gca acc tca cga att gct gaa ctt gaa agc aaa acc aat aca	728
Ser Ala Ala Thr Ser Arg Ile Ala Glu Leu Glu Ser Lys Thr Asn Thr	
135 140 145	
ctc cgt tta tca cag act gtg gct cca aac tgc ttc aac tca tca ata	776
Leu Arg Leu Ser Gln Thr Val Ala Pro Asn Cys Phe Asn Ser Ser Ile	
150 155 160	
aat aat att cat gaa atg gaa ata cag ctg aaa gat gct ctg gag aaa	824
Asn Asn Ile His Glu Met Glu Ile Gln Leu Lys Asp Ala Leu Glu Lys	
165 170 175 180	
aat cag cag tgg ctc gtg tat gat cag cag cgg gaa gtc tat gta aaa	872

Asn Gln Gln Trp Leu Val Tyr Asp Gln Gln Arg Glu Val Tyr Val Lys	
185	190
195	
gga ctt tta gca aag atc ttt gag ttg gaa aag aaa acg gaa aca gct	920
Gly Leu Leu Ala Lys Ile Phe Glu Leu Glu Lys Lys Thr Glu Thr Ala	
200	205
210	
gct cat tca ctc cca cag cag aca aaa aag cct gaa tca gaa ggt tat	968
Ala His Ser Leu Pro Gln Gln Thr Lys Lys Pro Glu Ser Glu Gly Tyr	
215	220
225	
ctt caa gaa gag aag cag aaa tgt tac aac gat ctc ttg gca agt gca	1016
Leu Gln Glu Glu Lys Gln Lys Cys Tyr Asn Asp Leu Leu Ala Ser Ala	
230	235
240	
aaa aaa gat ctt gag gtt gaa cga caa acc ata act cag ctg agt ttt	1064
Lys Lys Asp Leu Glu Val Glu Arg Gln Thr Ile Thr Gln Leu Ser Phe	
245	250
255	260
gaa ctg agt gaa ttt cga aga aaa tat gaa gaa acc caa aaa gaa gtt	1112
Glu Leu Ser Glu Phe Arg Arg Lys Tyr Glu Glu Thr Gln Lys Glu Val	
265	270
275	
cac aat tta aat cag ctg ttg tat tca caa aga agg gca gat gtg caa	1160
His Asn Leu Asn Gln Leu Leu Tyr Ser Gln Arg Arg Ala Asp Val Gln	
280	285
290	
cat ctg gaa gat gat agg cat aaa aca gag aag ata caa aaa ctc agg	1208
His Leu Glu Asp Asp Arg His Lys Thr Glu Lys Ile Gln Lys Leu Arg	
295	300
305	
gaa gag aat gat att gct agg gga aaa ctt gaa gaa gag aag aag aga	1256
Glu Glu Asn Asp Ile Ala Arg Gly Lys Leu Glu Glu Glu Lys Lys Arg	
310	315
320	
tcc gaa gag ctc tta tct cag gtc cag ttt ctt tac aca tct ctg cta	1304
Ser Glu Glu Leu Leu Ser Gln Val Gln Phe Leu Tyr Thr Ser Leu Leu	
325	330
335	340
aag cag caa gaa gaa caa aca agg gta gct ctg ttg gaa caa cag atg	1352
Lys Gln Gln Glu Glu Gln Thr Arg Val Ala Leu Leu Glu Gln Gln Met	
345	350
355	

cag gca tgt act tta gac ttt gaa aat gaa aaa ctc gac cgt caa cat	1400
Gln Ala Cys Thr Leu Asp Phe Glu Asn Glu Lys Leu Asp Arg Gln His	
360 365 370	
gtg cag cat caa ttg ctt gta att ctt aag gag ctc cga aaa gca aga	1448
Val Gln His Gln Leu Leu Val Ile Leu Lys Glu Leu Arg Lys Ala Arg	
375 380 385	
aat caa ata aca cag ttg gaa tcc ttg aaa cag ctt cat gag ttt gcc	1496
Asn Gln Ile Thr Gln Leu Glu Ser Leu Lys Gln Leu His Glu Phe Ala	
390 395 400	
atc aca gag cca tta gtc act ttc caa gga gag act gaa aac aga gaa	1544
Ile Thr Glu Pro Leu Val Thr Phe Gln Gly Glu Thr Glu Asn Arg Glu	
405 410 415 420	
aaa gtt gcc gcc tca cca aaa agt ccc act gct gca ctc aat gaa agc	1592
Lys Val Ala Ala Ser Pro Lys Ser Pro Thr Ala Ala Leu Asn Glu Ser	
425 430 435	
ctg gtg gaa tgt ccc aag tgc aat ata cag tat cca gcc act gag cat	1640
Leu Val Glu Cys Pro Lys Cys Asn Ile Gln Tyr Pro Ala Thr Glu His	
440 445 450	
cgc gat ctg ctt gtc cat gtg gaa tac tgt tca aag tag caaataagt	1689
Arg Asp Leu Leu Val His Val Glu Tyr Cys Ser Lys	
455 460	
atgtgtttg atattaaaag attcaatact gtattttctg ttagcttgtg ggcattttga	1749
attatatatt tcacattttg cataaaaactg cctatctacc tttgacactc cagcatgcta	1809
gtgaatcatg tatcttttag gctgctgtgc atttctcttg gcagtgatac ctcctgaca	1869
tggttcatca tcaggctgca atgacagaat gtggtgagca gcgtctactg agactactaa	1929
cattttgcac tgtcaaaaata cttggtgagg aaaagatagc tcaggttatt gctaattgggt	1989
taatgcacca gcaagcaaaa tattttatgt tttgggggtt tgaaaaatca aagataatta	2049
accaaggatc ttaactgtgt tcgcattttt tatccaagca cttagaaaac ctacaatcct	2109

aattttgatg tccattgtta agaggtggg atagatacta ttttttttt catattgtat 2169
agcgggtatt agaaaagttg gggattttct tgatctttat tgctgcttac cattgaaact 2229
taaccagct gtgttcccca actctgttct gcgcaagaaa cagtatctgt ttgaggcata 2289
atcttaagtg gccacacaca atgttttctc ttatgttate tggcagtaac tgtaacttga 2349
attacattag cacattctgc ttagctaaaa ttgttaaaat aaactttaat aaacctatgt 2409
agccctctca tttgattgac agtattttag ttatttttgg cattcttaaa gctgggcaat 2469
gtaatgatca gatctttggt tgcctgaaca ggtattttta tacatgcttt ttgtaaacca 2529
aaaactttta aatttctca ggttttctaa catgcttacc actgggctac tgtaaatgag 2589
aaaagaataa aattatttaa tgttttaaaa aaaaaaaaa aaaaaa 2635

<210> 10
<211> 464
<212> PRT
<213> 人

<400> 10

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

115	120	125
Lys Gln Gln Leu Ser Ala Ala Thr Ser Arg Ile Ala Glu Leu Glu Ser		
130	135	140
Lys Thr Asn Thr Leu Arg Leu Ser Gln Thr Val Ala Pro Asn Cys Phe		
145	150	155
Asn Ser Ser Ile Asn Asn Ile His Glu Met Glu Ile Gln Leu Lys Asp		
165	170	175
Ala Leu Glu Lys Asn Gln Gln Trp Leu Val Tyr Asp Gln Gln Arg Glu		
180	185	190
Val Tyr Val Lys Gly Leu Leu Ala Lys Ile Phe Glu Leu Glu Lys Lys		
195	200	205
Thr Glu Thr Ala Ala His Ser Leu Pro Gln Gln Thr Lys Lys Pro Glu		
210	215	220
Ser Glu Gly Tyr Leu Gln Glu Glu Lys Gln Lys Cys Tyr Asn Asp Leu		
225	230	235
Leu Ala Ser Ala Lys Lys Asp Leu Glu Val Glu Arg Gln Thr Ile Thr		
245	250	255
Gln Leu Ser Phe Glu Leu Ser Glu Phe Arg Arg Lys Tyr Glu Glu Thr		
260	265	270
Gln Lys Glu Val His Asn Leu Asn Gln Leu Leu Tyr Ser Gln Arg Arg		
275	280	285
Ala Asp Val Gln His Leu Glu Asp Asp Arg His Lys Thr Glu Lys Ile		
290	295	300
Gln Lys Leu Arg Glu Glu Asn Asp Ile Ala Arg Gly Lys Leu Glu Glu		
305	310	315
Glu Lys Lys Arg Ser Glu Glu Leu Leu Ser Gln Val Gln Phe Leu Tyr		
325	330	335
Thr Ser Leu Leu Lys Gln Gln Glu Glu Gln Thr Arg Val Ala Leu Leu		
340	345	350
Glu Gln Gln Met Gln Ala Cys Thr Leu Asp Phe Glu Asn Glu Lys Leu		
355	360	365
Asp Arg Gln His Val Gln His Gln Leu Leu Val Ile Leu Lys Glu Leu		
370	375	380
Arg Lys Ala Arg Asn Gln Ile Thr Gln Leu Glu Ser Leu Lys Gln Leu		
385	390	395
His Glu Phe Ala Ile Thr Glu Pro Leu Val Thr Phe Gln Gly Glu Thr		
405	410	415
Glu Asn Arg Glu Lys Val Ala Ala Ser Pro Lys Ser Pro Thr Ala Ala		
420	425	430
Leu Asn Glu Ser Leu Val Glu Cys Pro Lys Cys Asn Ile Gln Tyr Pro		
435	440	445
Ala Thr Glu His Arg Asp Leu Leu Val His Val Glu Tyr Cys Ser Lys		

450	455	460	
<210>	11		
<211>	2765		
<212>	DNA		
<213>	人		
<220>			
<221>	CDS		
<222>	(329)..(1048)		
<223>			
<400>	11		
oggagtctgg ccgcagtcgc gccagtggtg gcttcccatc cccaaaaggc gcctccgac			60
tctttgcgcc gcaactgctcg ccgggccagt ccggaaacgg gtcgtggagc tccgcaccac			120
tcccgtggt tcccgaagge agatcccttc tcccagagat tgcgagaaac tttcccttgt			180
ccccgacgct gcagcggctc gggtagcctg gcagccgcag gtttctgaac cccgggccac			240
gtcccccgcg cctcgcttc gcgctcgtgt agatcgttcc ctctctggtt gcacgtggg			300
gatcccgac ctcgattctg cgggcgag atg ccc ctg gga cac atc atg agg			352
		Met Pro Leu Gly His Ile Met Arg	
		1 5	
ctg gac ctg gag aaa att gcc ctg gag tac atc gtg ccc tgt ctg cac			400
Leu Asp Leu Glu Lys Ile Ala Leu Glu Tyr Ile Val Pro Cys Leu His			
10 15 20			
gag gtg ggc ttc tgc tac ctg gac aac ttc ctg ggc gag gtg gtg ggc			448
Glu Val Gly Phe Cys Tyr Leu Asp Asn Phe Leu Gly Glu Val Val Gly			
25 30 35 40			
gac tgc gtc ctg gag cgc gtc aag cag ctg cac tgc acc ggg gcc ctg			496
Asp Cys Val Leu Glu Arg Val Lys Gln Leu His Cys Thr Gly Ala Leu			
45 50 55			
ogg gac ggc cag ctg gcg ggg ccg cgc gcc ggc gtc tcc aag cga cac			544
Arg Asp Gly Gln Leu Ala Gly Pro Arg Ala Gly Val Ser Lys Arg His			

60	65	70	
ctg cgg ggc gac cag atc acg tgg atc ggg ggc aac gag gag ggc tgc			592
Leu Arg Gly Asp Gln Ile Thr Trp Ile Gly Gly Asn Glu Glu Gly Cys			
75	80	85	
gag gcc atc agc ttc ctc ctg tcc ctc atc gac agg ctg gtc ctc tac			640
Glu Ala Ile Ser Phe Leu Leu Ser Leu Ile Asp Arg Leu Val Leu Tyr			
90	95	100	
tgc ggg agc cgg ctg ggc aaa tac tac gtc aag gag agg tct aag gca			688
Cys Gly Ser Arg Leu Gly Lys Tyr Tyr Val Lys Glu Arg Ser Lys Ala			
105	110	115	120
atg gtg gct tgc tat ccg gga aat gga aca ggt tat gtt cgc cac gtg			736
Met Val Ala Cys Tyr Pro Gly Asn Gly Thr Gly Tyr Val Arg His Val			
125	130	135	
gac aac ccc aac ggt gat ggt cgc tgc atc acc tgc atc tac tat ctg			784
Asp Asn Pro Asn Gly Asp Gly Arg Cys Ile Thr Cys Ile Tyr Tyr Leu			
140	145	150	
aac aag aat tgg gat gcc aag cta cat ggt ggg atc ctg cgg ata ttt			832
Asn Lys Asn Trp Asp Ala Lys Leu His Gly Gly Ile Leu Arg Ile Phe			
155	160	165	
cca gag ggg aaa tca ttc ata gca gat gtg gag ccc att ttt gac aga			880
Pro Glu Gly Lys Ser Phe Ile Ala Asp Val Glu Pro Ile Phe Asp Arg			
170	175	180	
ctc ctg ttc ttc tgg tca gat cgt agg aac cca cac gaa gtg cag ccc			928
Leu Leu Phe Phe Trp Ser Asp Arg Arg Asn Pro His Glu Val Gln Pro			
185	190	195	200
tct tac gca acc aga tat gct atg act gtc tgg tac ttt gat gct gaa			976
Ser Tyr Ala Thr Arg Tyr Ala Met Thr Val Trp Tyr Phe Asp Ala Glu			
205	210	215	
gaa agg gca gaa gcc aaa aag aaa ttc agg aat tta act agg aaa act			1024
Glu Arg Ala Glu Ala Lys Lys Lys Phe Arg Asn Leu Thr Arg Lys Thr			
220	225	230	

gaa tet gcc etc act gaa gac tga cegtctctg aaatctgctg gccttgttca Glu Ser Ala Leu Thr Glu Asp 235	1078
ttttagtaac ggttcctgaa ttctcttaaa ttctttgaga tccaaagatg gcctcttcag	1138
tgacaacaat ctccctgcta cttcttgcac ccttcacatc cctgtcttgt gtgtgttact	1198
tcattgtttc ttgccaagac tgtgttgatc ttcagatact ctctttgcca gatgaagtta	1258
cttgctaact ccagaaatc ctgcagacat cctactcggc cagcggttta cctgatagat	1318
tcggtaatac tatcaagaga agagcctagg agcacagcga gggaatgaac cttacttgca	1378
ctttatgtat acttctgat ttgaaaggag gaggtttgaa aagaaaaaaaa tggaggtggt	1438
agatgccaca gagaggcatc acggaagcct taacagcagg aaacagagaa atttgtgtca	1498
tctgaacaat ttccagatgt tcttaatcca gggtcttgg ggtttctgga gaattatcac	1558
aacctaatac cattaatacc tctagaaagg gctgctgtca tagtgaacaa tttataagtg	1618
tcccatgggg cagacactcc tttttccca gtctgcaac ctggattttc tgcctcagcc	1678
ccattttgct gaaaataatg actttctgaa taaagatggc aacacaattt tttctccatt	1738
ttcagttctt acctgggaac ctaattcccc agaagctaaa aaactagaca ttagttgttt	1798
tggttgcttt gttggaatgg aatttaaatt taaatgaaag gaaaaatata tccctggtag	1858
ttttgtgtta accactgata actgtggaaa gagctaggtc tactgatata caataacat	1918
gtgtgcatct tgaacaattt gagaggggag gtggagtgg aatgtgggt gttcctgttt	1978
ttttttttt tttttttta gttttcttt ttaatgagct caccotttta cacaaaaaaaa	2038
gcagggtgat gtattttaaa aaaggaagtg gaaataaaaa aatctcaaag ctatttgagt	2098
tctcgtctgt cctagcagt ctttcttcag ctcaactggc tctctagatc caactgtggtt	2158
ggcagtatga ccagaatcat ggaatttctt agaactgtgg aagcttctac tcctgcagta	2218

agcacagatc gcaactgcctc aataacttgg tattgagcac gtatittgca aaagctactt 2278
 ttccatgatt tcagtattac tttcatgttt taaaaatccc ttaatttct tgcttgaaaa 2338
 tccatgaac attaaagagc cagaaatatt ttcctttggt atgtacggat atatatatat 2398
 atagtcttcc aagatagaag tttacttttt cctcttctgg ttttgaaaa tttccagata 2458
 agacatgtca ccattaattc tcaacgactg ctctattttg ttgtacggta atagttatca 2518
 ccttctaaat tactatgtaa tttactcact tattatgttt attgtcttgt atcctttctc 2578
 tggagtgtaa gcacaatgaa gacaggaatt ttgtatattt ttaaccaatg caacatactc 2638
 tcagcaccta aaatagtgcc gggaacatag taagggtca gtaaatactt gttgaataaa 2698
 ctcagtctcc tacattagca ttctaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2758
 aaaaaag 2765

<210> 12

<211> 239

<212> PRT

<213> 人

<400> 12

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

Arg Tyr Ser Lys Asp Arg Lys Asp Leu

1 5

<210> 15

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 15

Phe Tyr Gln Val Glu Pro Ser Glu Ile

1 5

<210> 16

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 16

Lys Tyr Glu Leu Glu Asn Glu Glu Ile

1 5

<210> 17

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 17

Asp Met Ser Asn Ser Asn Asp Cys Met

1 5

<210> 18

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 18

Cys Met Arg Asp Ser Ser Gln Ile Leu

1 5

<210> 19

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 19

Arg Met Lys His Ile Arg Gln Ala Met

1 5

<210> 20

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 20

Ile Leu Ser His Ser Leu Trp Asp Val

1 5

<210> 33

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 33

Phe Met Gln Glu Leu Val Asp Gly Leu

1 5

<210> 34

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 34

Gly Leu Tyr His Thr Gly Ala Asn Val

1 5

<210> 35

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 35

Phe Ile Tyr Asp Phe Cys Ile Phe Gly Val
1 5 10

<210> 36

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 36

Leu Met Pro Glu Leu Ser Thr Phe Arg Val
1 5 10

<210> 37

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 37

Ala Met Ala Lys Asn Arg Leu Gln Phe Val
1 5 10

<210> 38

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 38

Lys Phe Thr Gly Gln Ala Val Glu Leu

1 5

<210> 48

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 48

Leu Phe Asp Glu Glu Phe Arg His Leu

1 5

<210> 49

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 49

Leu Tyr Ala Ser Ser Lys Pro Val Met

1 5

<210> 50

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 50

Ile Leu Met Asp Val Phe Thr Asp Val

1 5

<210> 51

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 51

Val Leu Ser Gly Ser Tyr Ser Phe Thr

1 5

<210> 52

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 52

Leu Gln Ser Gly Thr Tyr Phe Pro Val

1 5

<210> 53

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 53

Leu Tyr Glu Gln Ala Arg Trp Ala Ile Leu
1 5 10

<210> 63
<211> 9
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 63

Gln Tyr His Ile Ser Lys Leu Ser Leu
1 5

<210> 64
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 64

Ser Tyr Gln Pro Glu Val Leu Asn Ile Leu
1 5 10

<210> 65
<211> 9
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 65

Glu Tyr Ile Gly Gly Tyr Ile Phe Leu

1 5

<210> 66

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 66

Val Met Leu Lys Leu Val Glu Gln Ile

1 5

<210> 67

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 67

Cys Trp Leu Leu Lys Thr His Trp Thr Leu

1 5 10

<210> 68

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 68

Leu Phe Thr Pro Gln His Lys Met Leu

1 5

<210> 69

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 69

Thr Met Thr Trp Phe Ser Asp Ser Pro Leu

1 5 10

<210> 70

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 70

Gly Trp Leu Asp Ser Ser Arg Ser Leu

1 5

<210> 71

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 71

Asn Met Asp Met Asn Pro Glu Cys Phe

1 5

<210> 75

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 75

Gln Met Pro Leu Val Glu Ala Lys Leu

1 5

<210> 76

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 76

Arg Phe Ile Gln Ala Trp Gln Ser Leu

1 5

<210> 77

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 77

Ser Leu Leu Glu Asp Ile Thr Asp Ile
1 5

<210> 81

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 81

Lys Leu Phe Arg Pro Lys Lys Leu Leu
1 5

<210> 82

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 82

Lys Leu Asn Leu Arg Gly Cys Glu Val
1 5

<210> 83

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 83

Ile Leu Leu Glu Glu Ile Asp Cys Thr

1 5

<210> 84

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 84

Lys Gln Trp Asn Val Asn Trp Glu Thr

1 5

<210> 85

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 85

Val Val Ile Glu Phe Asp Gln Asn Val

1 5

<210> 86

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 86

Gln Leu Tyr Glu Gln Ala Arg Trp Ala Ile
1 5 10

<210> 87

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 87

Lys Met Leu Arg Leu Arg Leu Pro Asn Leu
1 5 10

<210> 88

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 88

Leu Leu Phe Thr Pro Gln His Lys Met Leu
1 5 10

<210> 89

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 89

Val Tyr Glu Ala Thr Gln Gly Asp Leu

1 5

<210> 96

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 96

Ser Tyr Gly Thr Leu Leu Asn Ala Ile

1 5

<210> 97

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 97

His Tyr Thr Asn Lys Ile Arg Ala Leu

1 5

<210> 98

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 98

Val Met Tyr Cys Lys Glu Lys Leu Ile

1 5

<210> 102

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 102

Lys Met Asp Glu Leu His Lys Lys Leu

1 5

<210> 103

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 103

Thr Tyr Gln Gln Thr Pro Val Asn Met

1 5

<210> 104

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 104

Glu Phe Leu Asp Asp Ser Thr Val Trp

1 5

<210> 108

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 108

Val Trp Gly Ile Arg Cys Asn Lys Thr Leu

1 5 10

<210> 109

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 109

Pro Trp Asp Asp Lys Leu Ile Phe Lys Leu

1 5 10

<210> 110

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 110

Leu Met Glu Arg Leu Lys Pro Ser Met

1 5

<210> 111

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 111

Leu Tyr Lys Asn Thr Pro Glu Lys Val Met

1 5 10

<210> 112

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 112

Asn Phe Ile Leu Gly Asn Gly Phe Leu

1 5

<210> 113

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 113

Met Leu Phe Gly Thr Tyr Met Lys Val

1 5

<210> 120

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 120

Asp Met Trp Asn Glu Phe Phe His Val

1 5

<210> 121

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 121

Thr Val Thr Asp Ser Met Phe Ala Val

1 5

<210> 122

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 122

Arg Met Leu Tyr Met Ile Glu Gln Val

1 5

<210> 123

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 123

Gly Leu Phe Arg Arg Leu Pro His Leu

1 5

<210> 124

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 124

Tyr Met Ile Glu Gln Val His Asp Cys

1 5

<210> 125

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 125

Leu Leu Ser Gly Leu Ser Lys Pro Val

1 5

<210> 126

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 126

Tyr Leu Ile Thr Leu Leu Glu His Leu

1 5

<210> 127

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 127

Phe Leu Tyr Asn His Gly Ile Gly Thr

1 5

<210> 128

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 128

Lys Met Asp Glu Leu His Lys Lys Leu

1 5

<210> 129

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 129

Lys Leu Val Tyr Val His His Leu Leu

1 5

<210> 130

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 130

Trp Val Asn Glu Asp Arg His Tyr Met

1 5

<210> 131

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 131

Arg Leu Thr Asp Thr Asp Ala Ala Ile

1 5

<210> 132

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 132

Ala Leu Ile Asp Leu Gly Gln Ser Ile

1 5

<210> 133

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 133

Leu Ile Phe Lys Leu Leu Ser Gly Leu

1 5

<210> 134

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 134

Val Met Tyr Cys Lys Glu Lys Leu Ile

1 5

<210> 135

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 135

Phe Ile Thr Tyr Cys Leu Lys Phe Ala

1 5

<210> 136

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 136

Lys Leu Pro Val Glu Ser Val His Ile

1 5

<210> 137

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 137

Leu Leu Gly Glu Gly Ala Phe Ala Gln Val
1 5 10

<210> 141
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 141

Lys Leu Leu Ser Gly Leu Ser Lys Pro Val
1 5 10

<210> 142
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 142

Gly Glu Trp Glu Arg Tyr Ile Gln Trp Val
1 5 10

<210> 143
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 143

Val Leu Thr Cys Glu Ala Glu Leu Gly Val
1 5 10

<210> 147
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 147

Asn Leu Tyr Lys Asn Thr Pro Glu Lys Val
1 5 10

<210> 148
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 148

Phe Leu Tyr Asn His Gly Ile Gly Thr Leu
1 5 10

<210> 149
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 149

Trp Met Gln Met Ser Ser Leu Gly Thr Val
1 5 10

<210> 153

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 153

Phe Ile Met Asn Met Phe Gln Ala Pro Thr
1 5 10

<210> 154

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 154

Arg Leu Phe Gln Thr Arg Leu Thr Glu Thr
1 5 10

<210> 155

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 155

Trp Asn Tyr Gln Ile Asp Tyr Phe Gly Val
1 5 10

<210> 156
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 156

Gly Met Val Gln Ala Thr Pro Ser Lys Val
1 5 10

<210> 157
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 157

Asp Leu Leu Arg Gln Lys Leu Lys Lys Val
1 5 10

<210> 158
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 158

Val Tyr Val Lys Gly Leu Leu Ala Lys Ile
1 5 10

<210> 159
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 159

Gln Tyr Pro Ala Thr Glu His Arg Asp Leu
1 5 10

<210> 160
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 160

Glu Met Glu Ile Gln Leu Lys Asp Ala Leu
1 5 10

<210> 161
<211> 9
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 161

Gln Met Gln Ala Cys Thr Leu Asp Phe

1 5

<210> 162

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 162

Val Tyr Val Lys Gly Leu Leu Ala Lys Ile

1 5 10

<210> 163

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 163

Gln Tyr Pro Ala Thr Glu His Arg Asp Leu

1 5 10

<210> 164

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 164

Gln Leu Leu Val Ile Leu Lys Glu Leu

1 5

<210> 168

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 168

Ala Leu Leu Glu Gln Leu Glu Glu Thr

1 5

<210> 169

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 169

Lys Gln Gln Glu Glu Gln Thr Arg Val

1 5

<210> 170

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 170

Tyr Leu Gln Glu Glu Lys Gln Lys Cys

1 5

<210> 171

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 171

Thr Gln Leu Glu Ser Leu Lys Gln Leu

1 5

<210> 172

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 172

Leu Leu Tyr Ser Gln Arg Arg Ala Asp Val

1 5 10

<210> 173

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 173

Lys Leu Thr Asp Lys Glu Arg His Arg Leu
1 5 10

<210> 174
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 174

Ala Leu Leu Glu Gln Gln Met Gln Ala Cys
1 5 10

<210> 175
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 175

Leu Leu Ser Gln Val Gln Phe Leu Tyr Thr
1 5 10

<210> 176
<211> 9
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 176

Tyr Tyr Val Lys Glu Arg Ser Lys Ala Met
1 5 10

<210> 180
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 180

Asn Trp Asp Ala Lys Leu His Gly Gly Ile
1 5 10

<210> 181
<211> 9
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 181

Ile Phe Pro Glu Gly Lys Ser Phe Ile
1 5

<210> 182
<211> 9
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 182

Ser Phe Ile Ala Asp Val Glu Pro Ile

1 5

<210> 183

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 183

Arg Tyr Ala Met Thr Val Trp Tyr Phe

1 5

<210> 184

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 184

Phe Leu Leu Ser Leu Ile Asp Arg Leu

1 5

<210> 185

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 185

Tyr Leu Asp Asn Phe Leu Gly Glu Val

1 5

<210> 186

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 186

Tyr Ile Val Pro Cys Leu His Glu Val

1 5

<210> 187

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 187

Lys Leu His Gly Gly Ile Leu Arg Ile

1 5

<210> 188

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 188

Gln Leu Ala Gly Pro Arg Ala Gly Val

1 5

<210> 189

<211> 9

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 189

Cys Val Leu Glu Arg Val Lys Gln Leu

1 5

<210> 190

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 190

Cys Leu His Glu Val Gly Phe Cys Tyr Leu

1 5 10

<210> 191

<211> 10

<212> PRT

<213> 人工的

<220>

<223> 合成肽

<400> 191

Ala Met Lys Lys Tyr Glu Leu Glu Asn Glu Glu Ile
1 5 10

<210> 198
<211> 11
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 198

Arg Tyr Ile Ala Lys Arg Gln Leu Ser His Leu
1 5 10

<210> 199
<211> 11
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 199

Leu Tyr His Thr Gly Ala Asn Val Glu Ser Phe
1 5 10

<210> 200
<211> 11
<212> PRT
<213> 人工的

<220>
<223> 合成肽

<400> 200

<210> 204
 <211> 20
 <212> DNA
 <213> 人工的

<220>
 <223> 引物

<400> 204
 cagatccagg tgctctgac 20

<210> 205
 <211> 21
 <212> DNA
 <213> 人工的

<220>
 <223> 引物

<400> 205
 tcaghtaate caaccacett g 21

<210> 206
 <211> 20
 <212> DNA
 <213> 人工的

<220>
 <223> 引物

<400> 206
 tgaatttctg aggccttgct 20

<210> 207
 <211> 20
 <212> DNA
 <213> 人工的

<220>

<223> 引物

<400> 207

tgttctcagc agcaaacagg

20

<210> 208

<211> 20

<212> DNA

<213> 人工的

<220>

<223> 引物

<400> 208

ttatctgetg gcttggcact

20

<210> 209

<211> 21

<212> DNA

<213> 人工的

<220>

<223> 引物

<400> 209

gcttttgect taacaaatcc a

21

<210> 210

<211> 22

<212> DNA

<213> 人工的

<220>

<223> 引物

<400> 210

tgtccattgt taagaggtag tg

22

<210> 211

<211> 20

<212> DNA

<213> 人工的

<220>

<223> 引物

<400> 211

tgagagggt acatgggttt

20

<210> 212

<211> 20

<212> DNA

<213> 人工的

<220>

<223> 引物

<400> 212

catccctgtc ttgtgtgtgg

20

<210> 213

<211> 20

<212> DNA

<213> 人工的

<220>

<223> 引物

<400> 213

ccaacagccc tggattaaga

20

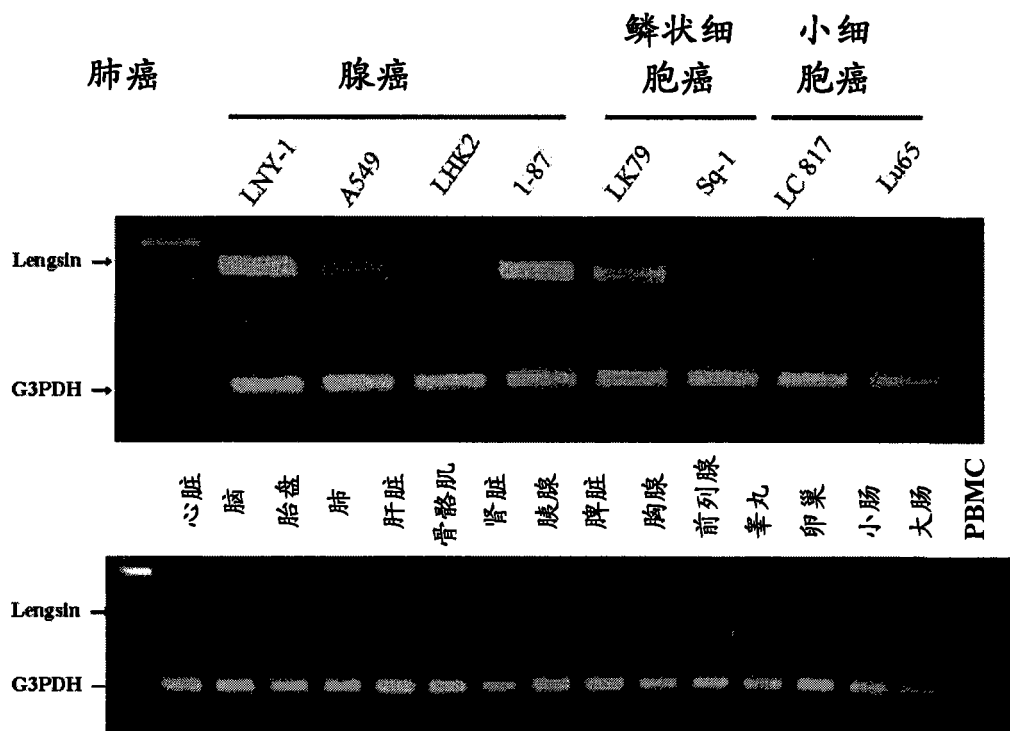


图 1

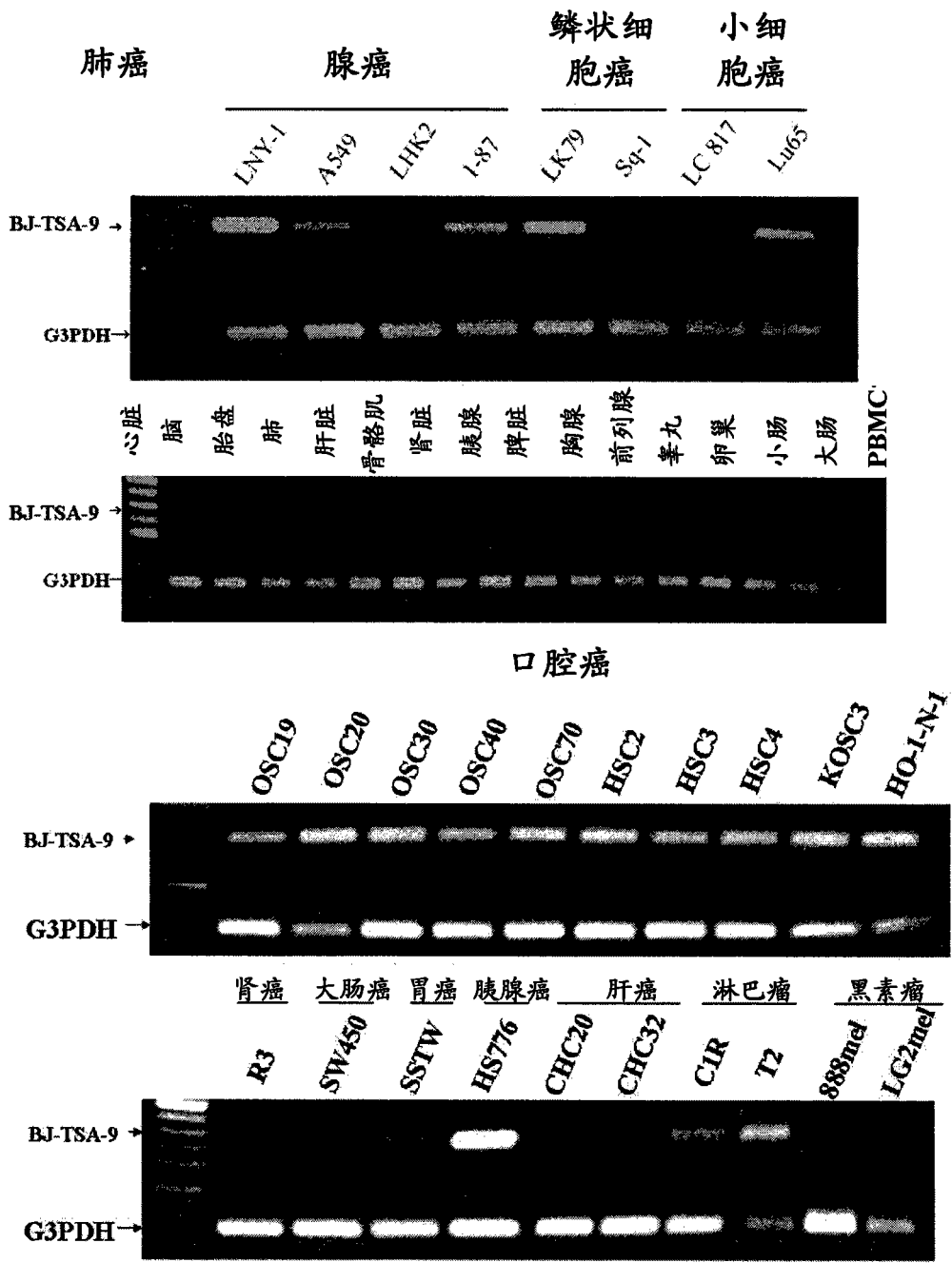


图 2

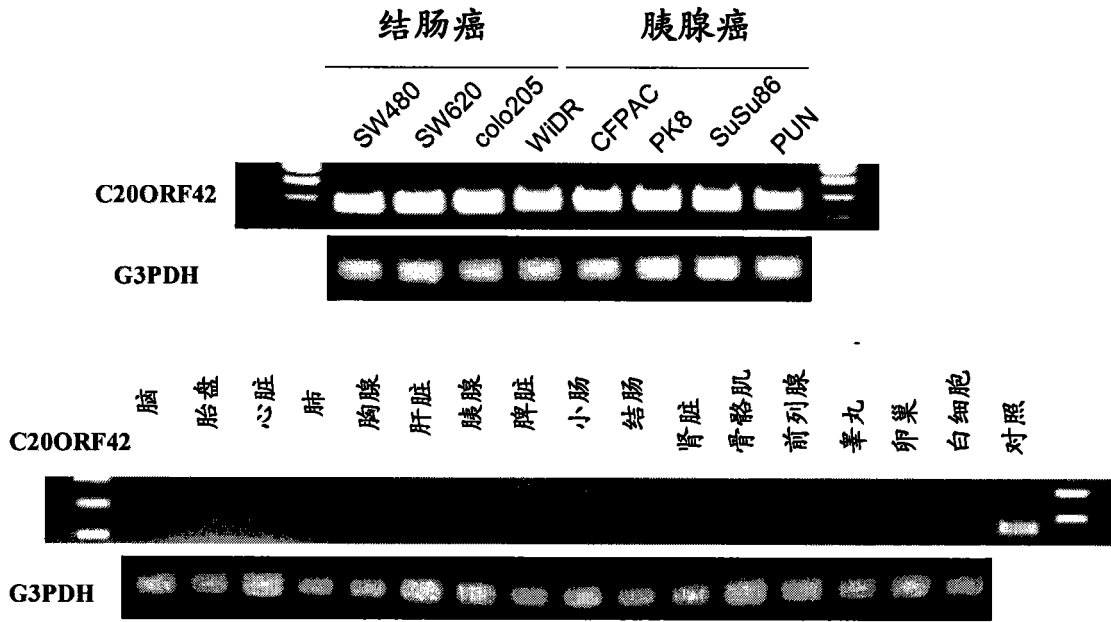


图 3

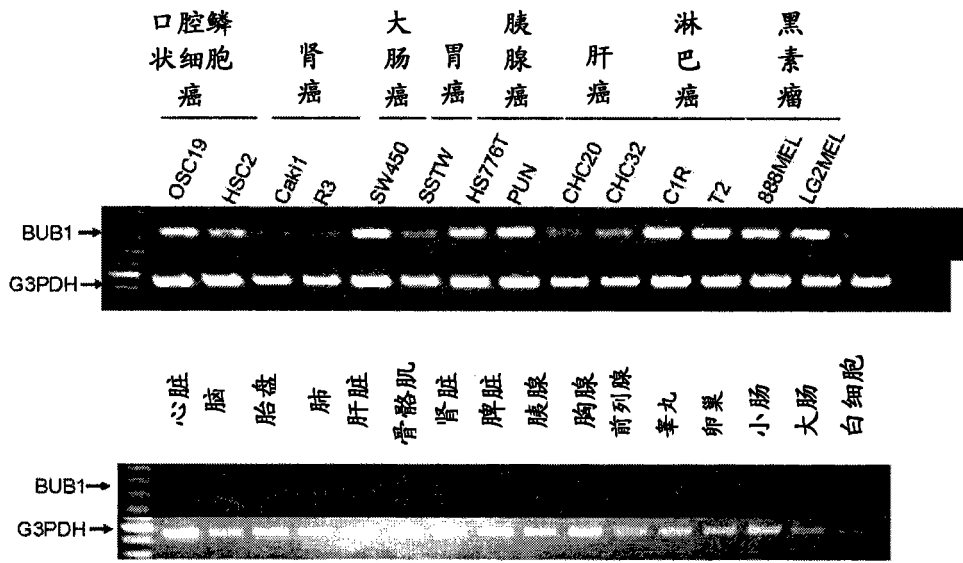


图 4

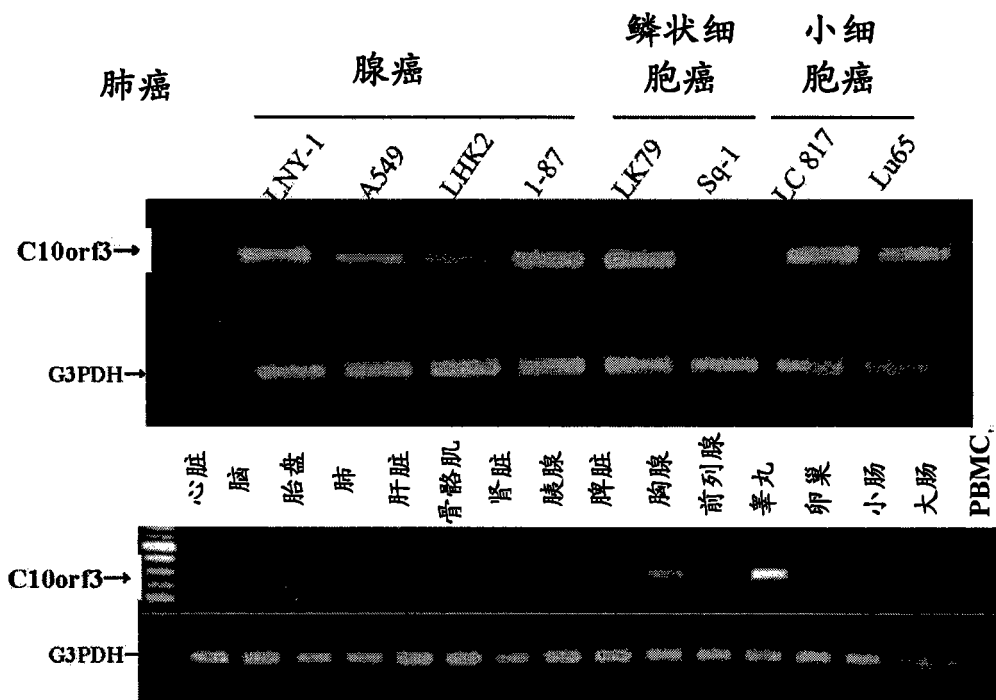


图 5

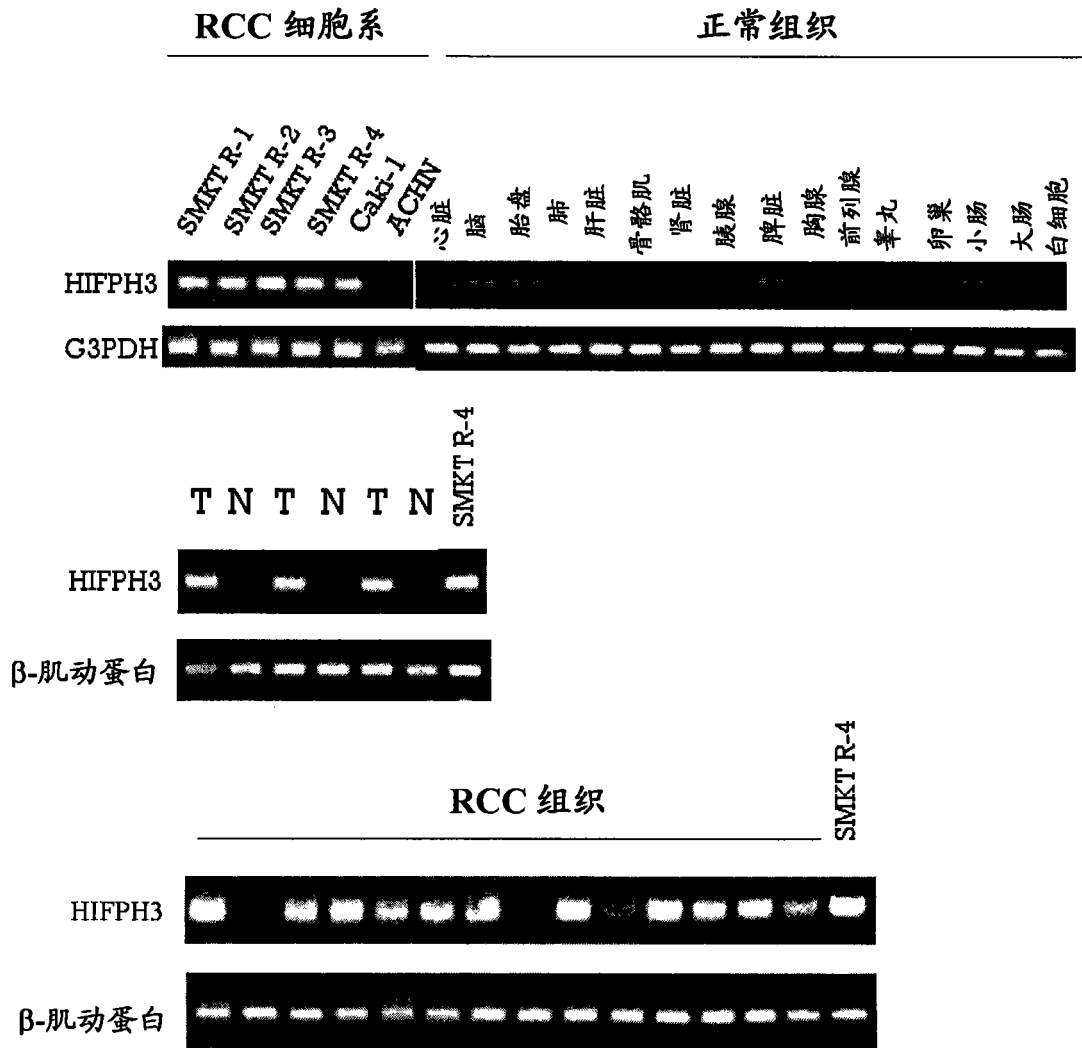


图 6

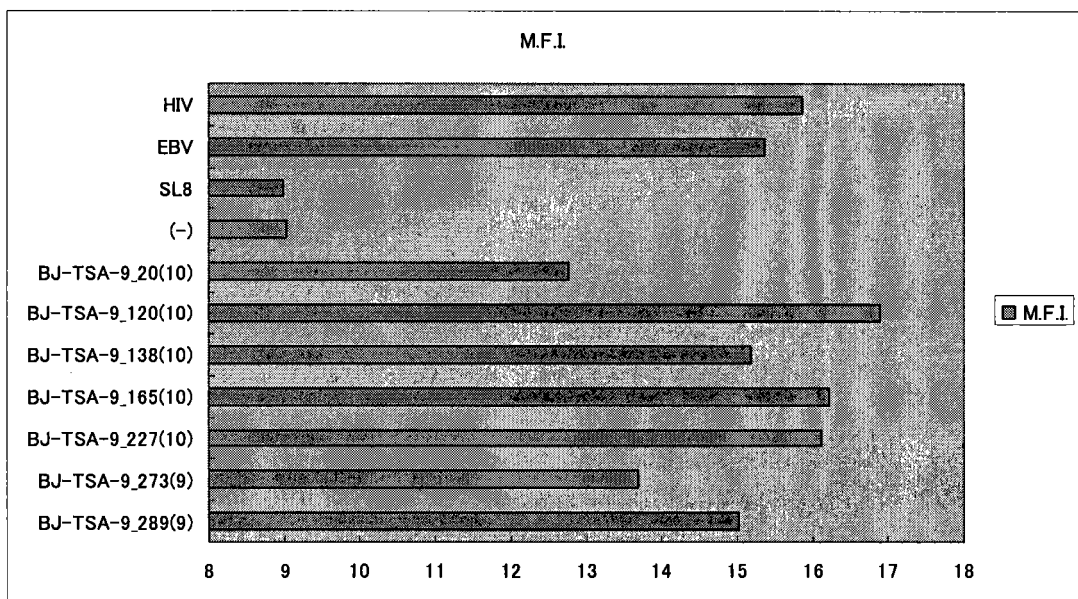


图 7

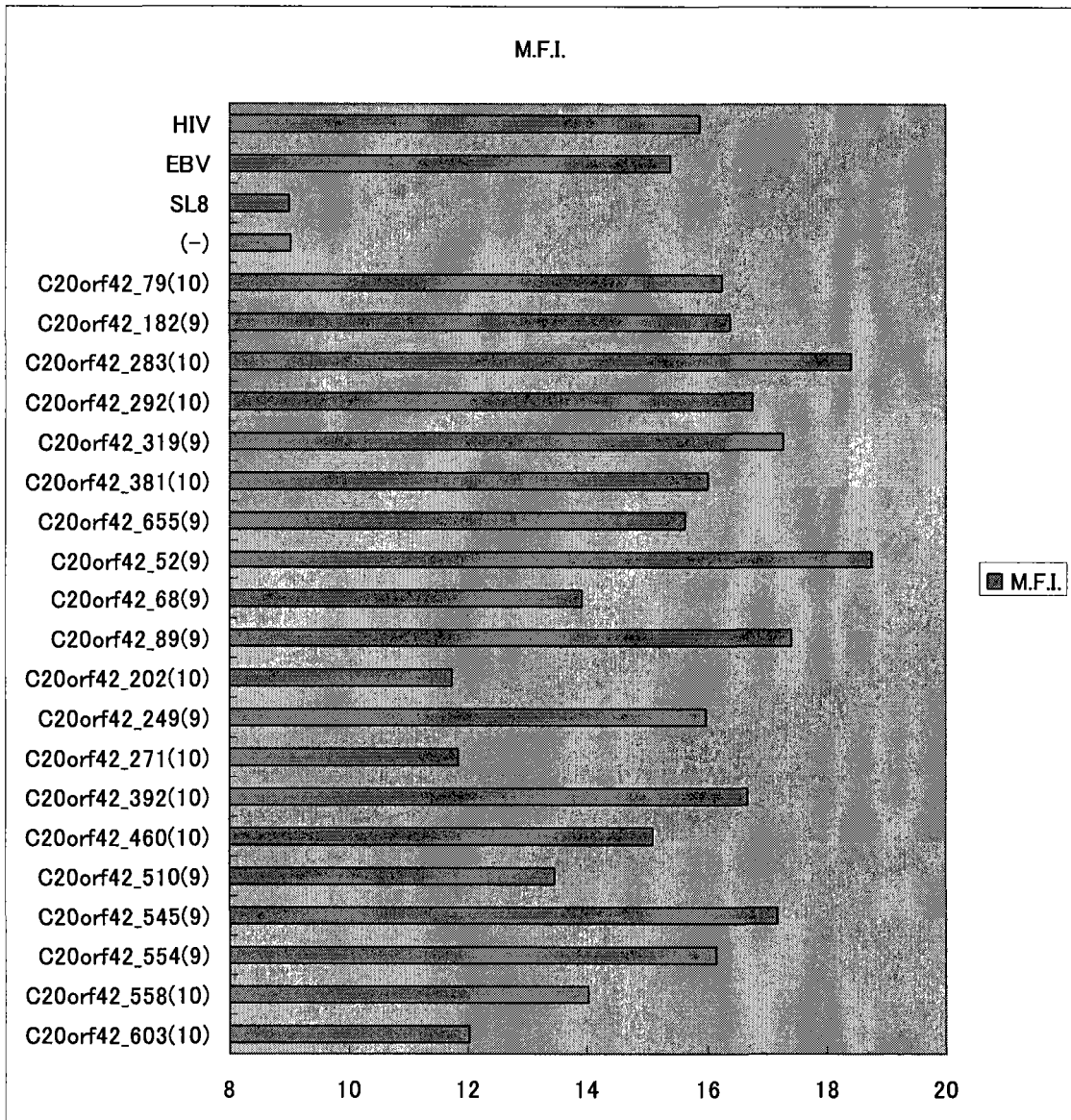


图 8

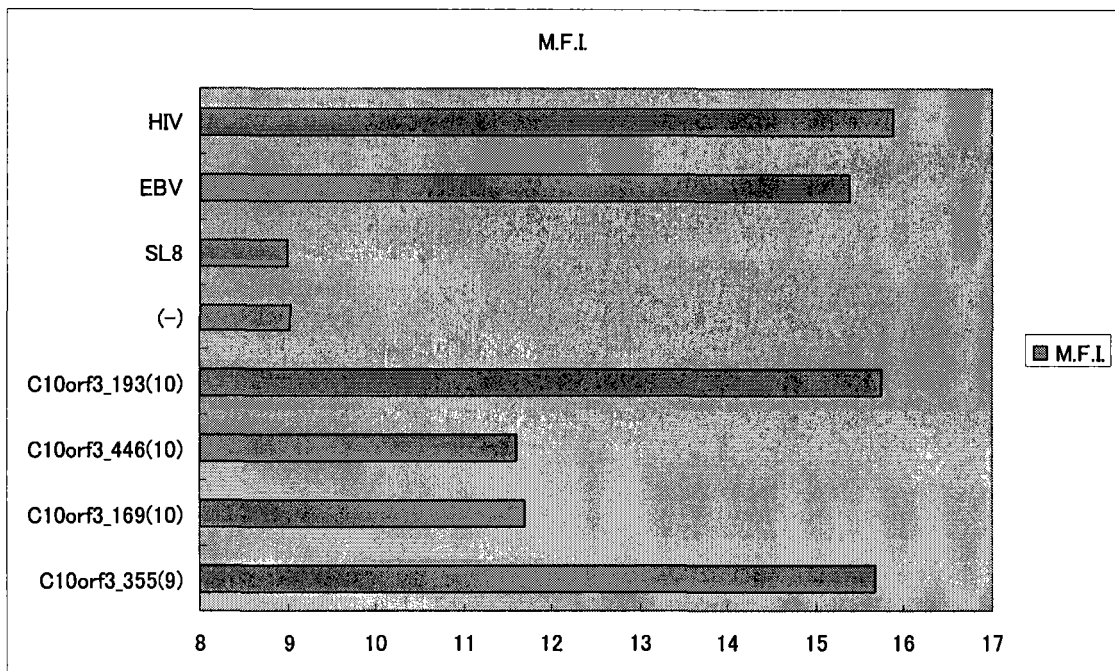


图 9

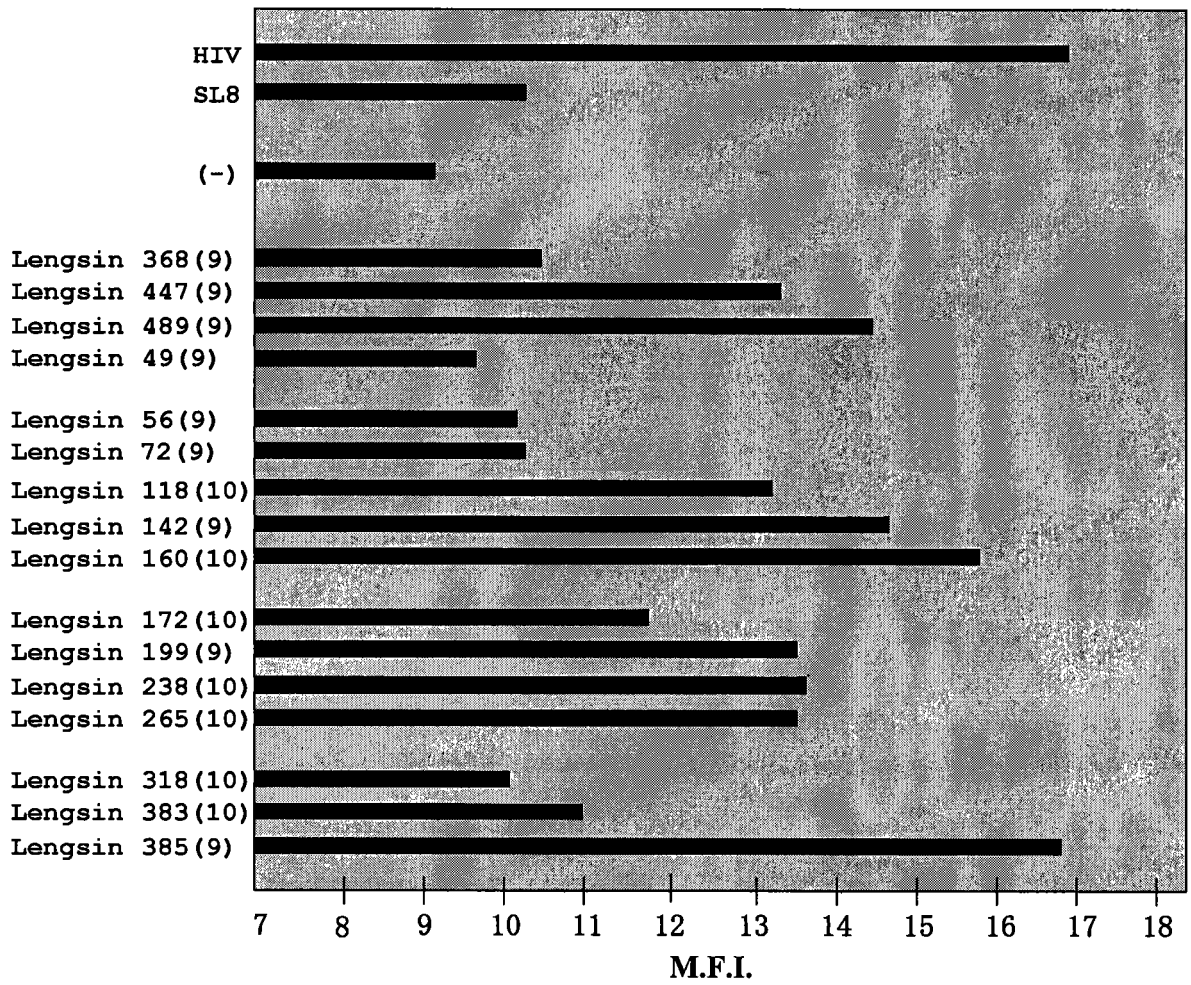


图 10

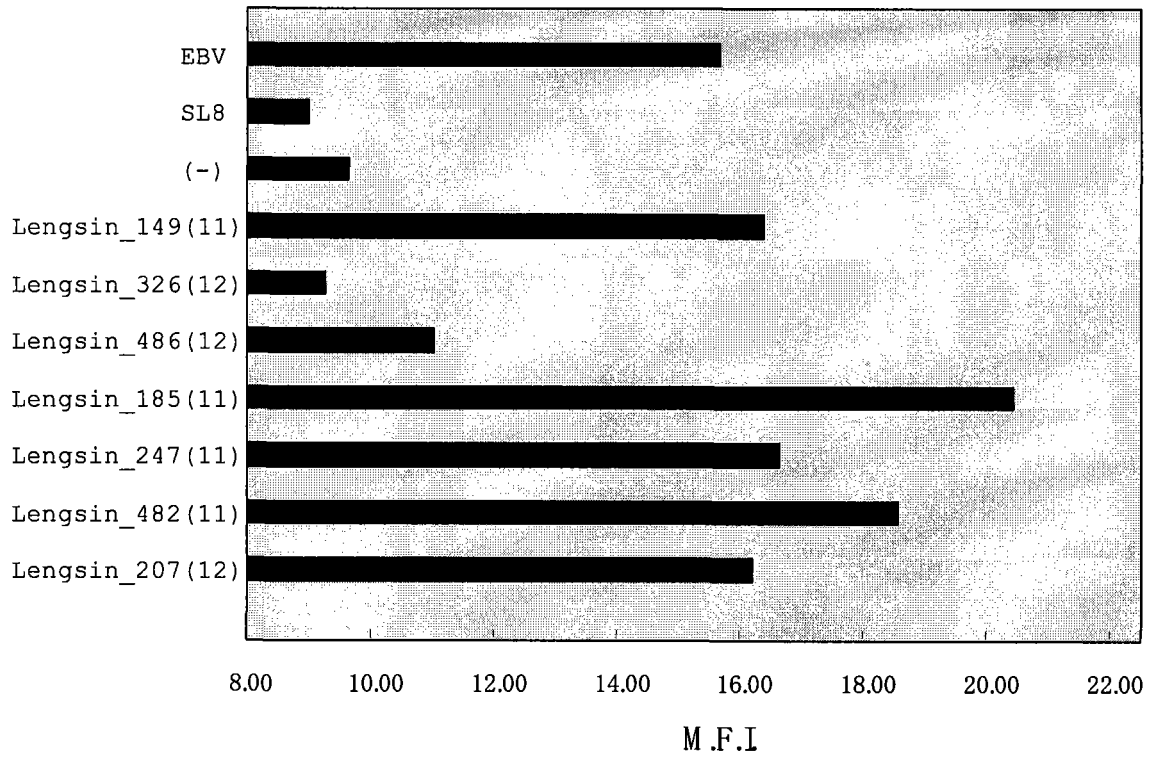


图 11

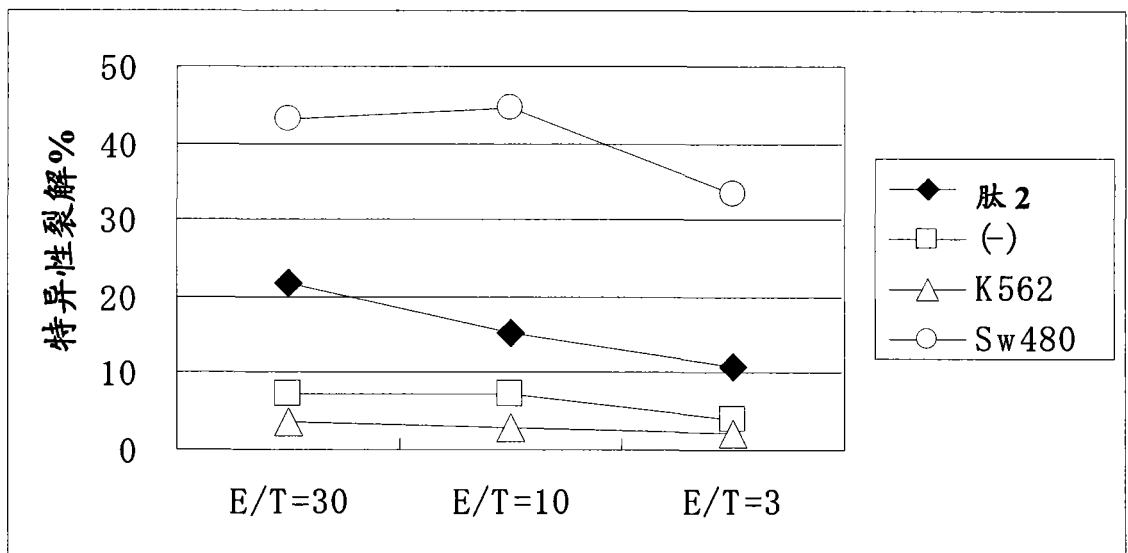


图 12

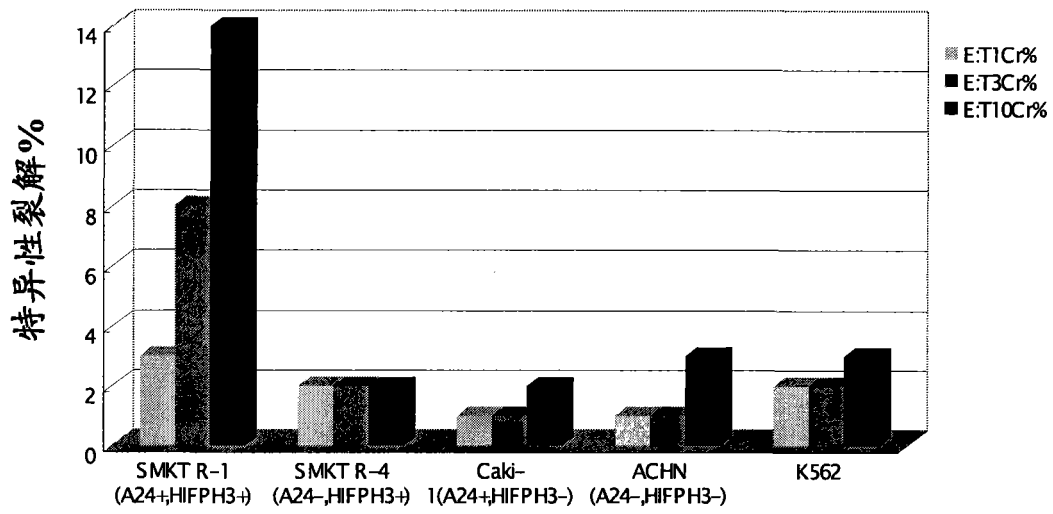


图 15

